PHARMACEUTICAL DOSAGE FORMS: TABLETS

PHARMACEUTICAL DOSAGE FORMS: TABLETS

Third Edition

Volume 3: Manufacture and Process Control

Edited by

Larry L. Augsburger
University of Maryland
Baltimore, Maryland, USA

Stephen W. Hoag
University of Maryland
Baltimore, Maryland, USA

informa
healthcare

New York London

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

© 2008 by Taylor & Francis Group, LLC
CRC Press is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works
Version Date: 20130116

International Standard Book Number-13: 978-1-4200-2030-4 (eBook - PDF)

This book contains information obtained from authentic and highly regarded sources. While all reasonable efforts have been made to publish reliable data and information, neither the author[s] nor the publisher can accept any legal responsibility or liability for any errors or omissions that may be made. The publishers wish to make clear that any views or opinions expressed in this book by individual editors, authors or contributors are personal to them and do not necessarily reflect the views/opinions of the publishers. The information or guidance contained in this book is intended for use by medical, scientific or health-care professionals and is provided strictly as a supplement to the medical or other professional's own judgement, their knowledge of the patient's medical history, relevant manufacturer's instructions and the appropriate best practice guidelines. Because of the rapid advances in medical science, any information or advice on dosages, procedures or diagnoses should be independently verified. The reader is strongly urged to consult the drug companies' printed instructions, and their websites, before administering any of the drugs recommended in this book. This book does not indicate whether a particular treatment is appropriate or suitable for a particular individual. Ultimately it is the sole responsibility of the medical professional to make his or her own professional judgements, so as to advise and treat patients appropriately. The authors and publishers have also attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access www.copyright.com (http://www.copyright.com/) or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

To my loving wife Jeannie,
the light and laughter in my life.

—Larry L. Augsburger

To my dear wife Cathy and my children Elena
and Nina and those who helped me
so much with my education:
My parents Jo Hoag and my late father
Jim Hoag, Don Hoag, and Edward G. Rippie.

—Stephen W. Hoag

Foreword

We are delighted to have the privilege of continuing the tradition begun by Herb Lieberman and Leon Lachman, and later joined by Joseph Schwartz, of providing the only comprehensive treatment of the design, formulation, manufacture and evaluation of the tablet dosage form in *Pharmaceutical Dosage Forms: Tablets*. Today the tablet continues to be the dosage form of choice. Solid dosage forms constitute about two-thirds of all dosage forms, and about half of these are tablets.

Philosophically, we regard the tablet as a drug delivery system. Like any delivery system, the tablet is more than just a practical way to administer drugs to patients. Rather, we view the tablet as a system that is designed to meet specific criteria. The most important design criterion of the tablet is how effectively it gets the drug "delivered" to the site of action in an active form in sufficient quantity and at the correct rate to meet the therapeutic objectives (i.e., immediate release or some form of extended or otherwise modified release). However, the tablet must also meet a number of other design criteria essential to getting the drug to society and the patient. These include physical and chemical stability (to assure potency, safety, and consistent drug delivery performance over the use-life of the product), the ability to be *economically* mass produced in a manner that assures the proper amount of drug in each dosage unit and batch produced (to reduce costs and provide reliable dosing), and, to the extent possible, patient acceptability (i.e., reasonable size and shape, taste, color, etc. to encourage patient compliance with the prescribed dosing regimen). Thus, the ultimate goal of drug product development is to design a system that maximizes the therapeutic potential of the drug substance and facilitates its access to patients. The fact that the tablet can be uniquely designed to meet these criteria accounts for its prevalence as the most popular oral solid dosage form.

Although the majority of tablets are made by compression, intended to be swallowed whole and designed for immediate release, there are many other tablet forms. These include, for example, chewable, orally disintegrating, sublingual, effervescent, and buccal tablets, as well as lozenges or troches. Effervescent tablets are intended to be taken after first dropping them in water. Some modified release tablets may be designed to delay release until the tablet has passed the pyloric sphincter (i.e., enteric). Others may be designed to provide consistent extended or sustained release over an extended period of time, or for pulsed release, colonic delivery, or to provide a unique release profile for a specific drug and its therapeutic objective.

Since the last edition of *Pharmaceutical Dosage Forms: Tablets* in 1990, there have been numerous developments and enhancements in tablet formulation science and technology, as well as product regulation. Science and technology developments include new or updated equipment for manufacture, new excipients, greater understanding of excipient functionality, nanotechnology, innovations in the design of modified release

tablets, the use of artificial intelligence in formulation and process development, new initiatives in real time and on-line process control, and increased use of modeling to understand and optimize formulation and process parameters. New regulatory initiatives include the Food and Drug Administration's SUPAC (scale up and post approval changes) guidances, its risk-based Pharmaceutical cGMPs for the 21st Century plan, and its PAT (process analytical technology) guidance. Also significant is the development, through the International Conference on Harmonization of proposals, for an international plan for a harmonized quality control system.

Significantly, the development of new regulatory policy and new science and technology are not mutually exclusive. Rather, they are inextricably linked. The new regulatory initiatives serve as a stimulus to academia and industry to put formulation design, development, and manufacture on a more scientific basis which, in turn, makes possible science-based policies that can provide substantial regulatory relief and greater flexibility for manufacturers to update and streamline processes for higher efficiency and productivity. The first SUPAC guidance was issued in 1995 for immediate release oral solid dosage forms (SUPAC-IR). That guidance was followed in 1997 with SUPAC-MR which covered scale-up and post approval changes for solid oral modified release dosage forms. These guidances brought much needed consistency to how the Food and Drug Administration deals with post approval changes and provided substantial regulatory relief from unnecessary testing and filing requirements. Major underpinnings of these two regulatory policies were research programs conducted at the University of Maryland under a collaborative agreement with the Food and Drug Administration which identified and linked critical formulation and process variables to bioavailability outcomes in human subjects. The Food and Drug Administration's Pharmaceutical cGMPs for the 21st Century plan seeks to merge science-based management with an integrated quality systems approach and to "create a robust link between process parameters, specifications and clinical performance"[1] The new PAT guidance proposes the use of modern process analyzers or process analytical chemistry tools to achieve real-time control and quality assurance during manufacturing.[2] The Food and Drug Administration's draft guidance on Q8 Pharmaceutical Development[3] addresses the suggested contents of the pharmaceutical development section of a regulatory submission in the ICH M4 Common Technical Document format.

A common thread running through these newer regulatory initiatives is the building in of product quality and the development of meaningful product specifications based on a high level of understanding of how formulation and process factors impact product performance.

Still other developments since 1990 are the advent of the internet as a research and resource tool and a decline in academic study and teaching in solid dosage forms. Together, these developments have led to a situation where there is a vast amount of formulation information widely scattered throughout the literature which is unknown and difficult for researchers new to the tableting field to organize and use. Therefore, another objective to this book to integrate a critical, comprehensive summary of this formulation information with the latest developments in this field.

Thus, the overarching goal of the third edition of *Pharmaceutical Dosage Forms: Tablets* is to provide an in-depth treatment of the science and technology of tableting that

[1] J. Woodcock, "Quality by Design: A Way Forward," September 17, 2003.

[2] http://www.fda.gov/cder/guidance/6419fnl.doc

[3] http://www.fda.gov/cder/guidance/6672dft.doc

acknowledges its traditional, historical database but focuses on modern scientific, technological, and regulatory developments. The common theme of this new edition is DESIGN. That is, tablets are delivery systems that are engineered to meet specific design criteria and that product quality must be built in and is also by design.

No effort of this magnitude and scope could have been accomplished without the commitment of a large number of distinguished experts. We are extremely grateful for their hard work, dedication and patience in helping us complete this new edition.

Larry L. Augsburger
Stephen W. Hoag

Preface

Volume 3 ties the fundamental process principles and the formulation and excipient principles presented in the previous two volumes together and applies these principles, along with additional information, to the commercial production and quality control of tablets. In particular, scale-up and troubleshooting are covered. Chapters 1–4 address the equipment, instrumentation for research and process control, automation in tablet production, and scale-up. In Chapters 5–7, the focus is on postmanufacture testing and evaluation of tablets, and the setting of dissolution specifications. Chapter 8 discusses the regulatory and good manufacturing practices environment in which tablets must be manufactured, with focus on the new paradigms of process analytical technology and quality by design. This volume concludes with chapters discussing the role of near-infrared chemical imaging in testing oral solid dosage forms, surface area and important related physical characteristics of solids, and intellectual property and the patent process.

Larry L. Augsburger
Stephen W. Hoag

Contents

Contributors

Göran Alderborn Department of Pharmacy, Uppsala University, Uppsala, Sweden

Albert Alexander AstraZeneca, Wilmington, Delaware, U.S.A.

Gary E. Bubb Specialty Measurements Inc., Lebanon, New Jersey, U.S.A.

Keith K. H. Chan University of Maryland, Baltimore, Maryland, U.S.A.

Albert W. K. Chan Law Offices of Albert Wai-Kit Chan, PLLC, New York, New York, U.S.A.

Chi-wan Chen Office of New Drug Quality Assessment, Center for Drug Evaluation and Research, U.S. Food and Drug Administration, Silver Spring, Maryland, U.S.A.

Göran Frenning Department of Pharmacy, Uppsala University, Uppsala, Sweden

Jean-Marie Geoffroy TAP Pharmaceuticals Inc., Lake Forest, Illinois, U.S.A.

Vivian A. Gray V. A. Gray Consulting, Inc., Hockessin, Delaware, U.S.A.

Marianthi Ierapetritou Department of Chemical and Biochemical Engineering, Rutgers University, Piscataway, New Jersey, U.S.A.

Linda H. Kidder Malvern Instruments, Columbia, Maryland, U.S.A.

Michael Levin Metropolitan Computing Corporation (MCC), East Hanover, New Jersey, U.S.A.

E. Neil Lewis Malvern Instruments, Columbia, Maryland, U.S.A.

Marcos Llusa Department of Chemical and Biochemical Engineering, Rutgers University, Piscataway, New Jersey, U.S.A.

Donghao (Robert) Lu Office of New Drug Quality Assessment, Center for Drug Evaluation and Research, U.S. Food and Drug Administration, Silver Spring, Maryland, U.S.A.

Patrick J. Marroum Office of Clinical Pharmacology, Center for Drug Evaluation and Research, U.S. Food and Drug Administration, Silver Spring, Maryland, U.S.A.

Ryan J. McCann Department of Industrial and Physical Pharmacy, Purdue University, West Lafayette, Indiana, U.S.A.

Kenneth R. Morris Department of Industrial and Physical Pharmacy, Purdue University, West Lafayette, Indiana, U.S.A.

Fernando J. Muzzio Department of Chemical and Biochemical Engineering, Rutgers University, Piscataway, New Jersey, U.S.A.

Moheb M. Nasr Office of New Drug Quality Assessment, Center for Drug Evaluation and Research, U.S. Food and Drug Administration, Silver Spring, Maryland, U.S.A.

Dale Natoli Natoli Engineering Company, St. Charles, Missouri, U.S.A.

Patricia Portillo Department of Chemical and Biochemical Engineering, Rutgers University, Piscataway, New Jersey, U.S.A.

Denise Rivkees Pfizer, Inc., Morris Plains, New Jersey, U.S.A.

Gerald M. Sando Malvern Instruments, Columbia, Maryland, U.S.A.

Josephine L. P. Soh Department of Industrial and Physical Pharmacy, Purdue University, West Lafayette, Indiana, U.S.A.

Paul A. Webb Micromeritics Instrument Corp., Norcross, Georgia, U.S.A.

1
Tooling for Pharmaceutical Processing

Dale Natoli
Natoli Engineering Company, St. Charles, Missouri, U.S.A.

INTRODUCTION

Compressing powders into a more solid mass dates back thousands of years. It was not until the early 1800s that tablet compression was automated in the sense the hand crank was replaced by a leather belt and a steam driven power bar. These early single station tablet presses were able to produce on an average 100 tablets per minute while meeting the guidelines of tablet uniformity for hardness, thickness, and weight. Soon after, single station presses were fading and making room for new technology, the rotary tablet press. Introduced in the mid-1800s, the rotary tablet press boasted speeds capable of compressing 1200 tablets per minute. Today, tablet presses are able to compress over 24,000 tablets per minute, and at the rate of new technology, it will surely increase (Fig. 1).

Compressing pharmaceutical tablets is the most efficient process for producing a single dose of medication. Tablets are accepted and trusted by professionals and consumers alike, they are easily administered and simple to dose.

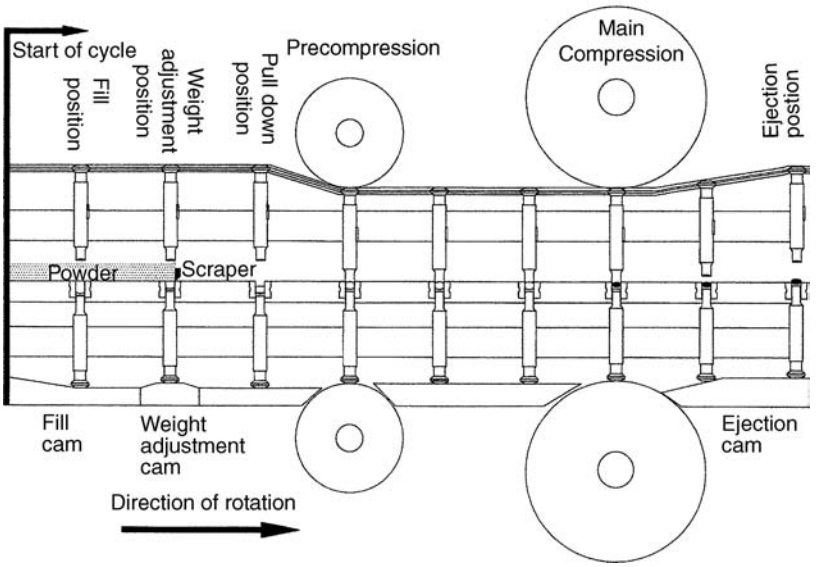


FIGURE 1 Rotary tablet press cycle.

Good granulation is important for compressing quality tablets. If the granulation is poor, the long term results will be too. A proper tablet granulation will have good flow, compressibility, and release properties. Tablet compression tooling is equally responsible for the success of a tableting program. Tooling must be engineered to withstand the stresses associated with tablet compression, provide satisfactory service life and maintain physical tablet uniformity. A proper tablet design is critical as well. Pharmaceutical marketing departments feverishly attempt to design tablets so unique, anticipating the design will quickly become branded and trusted in the eye of the consumer. A proper tool design is essential for putting that innovative design into the eye of the consumer.

The basic knowledge of tablet compression tooling and tablet design can save literally millions of dollars, prevent product loss, reduce unnecessary equipment down-time and help increase market shares. Understanding the basic physics of tablet compression will greatly enhance the ability to compress quality tablets more efficiently and provide better knowledge to troubleshoot and identify potential pitfalls before they happen, and they do!

Communication is important with any tableting campaign. Marketing, R&D, Engineering, Production, and the tooling supplier must be in accord and communicate new product-design and production requirements. The ideas and responsibilities of these departments may vary, but they share the common goal of manufacturing a quality tablet, efficiently, and productively.

TERMINOLOGY

In order to communicate properly and understand the following material it is important to have a basic understanding of the terminology used in this industry (Tables 1 and 2). Although these terms are most common and accepted, some may vary slightly between countries. This chapter deals with the terminology and general information related to the most commonly used rotary press tooling, the "TSM," "B," "D," "Euronorm" 19 and 21 mm configurations.

Common Tooling Standards

Internationally there are two recognized standards for tablet compression tooling, the TSM and the EU standards. Both TSM and EU standards identify the physical tool configuration for B and D type compression tools, their critical dimensions and associated tolerances assuring tablet quality and smooth press operation (Figs. 3 and 4).

The TSM tooling standard is recognized in the Americas and is considered exclusive in the United States. "TSM" is the acronym for the "Tablet Specification Manual" and is published, revised, and distributed by the American Pharmacist Association in Washington DC. The TSM Standards, once known as the IPT standards were originally developed in 1968 by a committee consisting of major pharmaceutical companies in the United States. The motivation was an attempt to maintain standardization for B and D tablet compression tooling which provides interchangeability between tablet presses. The TSM provides engineered drawings that are a valuable reference for troubleshooting and tool inspection. Today, the TSM committee consists of professionals from the tablet press, tooling, and tablet manufacturing industries. The TSM also includes useful information such as standard cup configurations for round tablets and a reference to common bisects for breaking tablets into multiple uniform dosages.

TABLE 1 Punches and Dies Terminology

Term	Definition
Tooling set	A complete set of punches and dies to accommodate all stations in a tablet press
Tooling station	The upper punch, lower punch, and die which accommodate one station in a tablet press
Head	The largest diameter of a common punch which contacts the machines cams and accepts the pressure from the pressure rollers
Head flat	The flat portion of the head which makes contact with the pressure rollers and determines the maximum dwell time for compression
Top head angle	Angle from the outside head diameter to the top head radius; it allows for sufficient head thickness and smoother camming
Top head radius	The radius on the top of the head which blends the top head angle to the head flat. Some head configurations may consist of only the head radius without the head angle. This radius makes the initial contact with the pressure roll and allows a smoother transition into the compression cycle
Head back angle	Sometimes referred to as the inside head angle, located underneath the top head angle or the top head radius which contacts the machine camming for vertical movement of the punch within the punch guides
Neck	Located below the head and provides clearance as the punch cycles through the machine cams
Barrel or shank	The vertical bearing surface of a punch which makes contact with the punch guides in the machine turret for verticle guidance
Barrel chamfer	Chamfers at the ends of the punch barrel, eliminate outside corners
Barrel-to-stem radius	The radius that blends the punch barrel to the stem
Stem	The area from the barrel to the edge of the punch tip
Tip length	The straight portion of the punch stem
Tip straight	The section of the tip that extends from the tip relief to the end of the punch tip; it maintains the punch tip size tolerance
Land	The area between the edge of the punch cup and the outside diameter of the punch tip; this adds strength to the tip to reduce punch tip fracturing
Tip face or cup	The portion of the punch tip that determines the contour of the tablet face; it includes the tablet embossing
Cup depth	The depth of the cup from the highest point of the tip edge to the lowest point of the cavity
Tip relief	The portion of the punch stem which is a undercut or made smaller than the punch tip straight; most common for lower punches to aid in reducing friction from the punch tip and die wall as the punch travels through the compression cycle; the area where the punch tip and relief meet must be sharp to scrape product from the die wall as the lower punch travels down for the fill cycle
Key	A projection normally of mild steel which protrudes above the surface of the punch barrel. It maintains alignment of the upper punch for reentry into the die; mandatory on upper punches with multiple tips and all tablet shapes other than round; commonly used with embossed round tablet shapes when rotation of the punch causes a condition known as double impression

(*Continued*)

TABLE 1 Punches and Dies Terminology (*Continued*)

Term	Definition
Key position	The radial and height position of a key on the punch barrel; not found in all presses
Punch overall length	The total length of a punch, other than flat-face tablet configurations, that is normally a reference dimension which consist of a combination of the working length and the cup depth dimensions
Working length	The dimension from the head flat to the lowest measurable point of the tip face, responsible for the consistency of the tablet overall thickness
Anneal	A heat-treating process used on fragile punch tips to decrease the hardness of the punch cups reducing punch tip fracturing
Bakelite tip relief	An undercut groove between the lower punch tip straight and the relief; it assures a sharp corner to assist in scraping product adhering to the die wall: normally a purchased option for lower punches
Barrel Flutes	Verticle slots machined into the punch barrel to reduce the bearing surface and assist in removing product in the punch guides: a purchased option for upper and lower punches
Die	A component used in conjunction with the upper and lower punches; it accepts the product for compaction and is responsible for the tablet's perimeter size and configuration
Die height or overall length	The entire height or overall length of a die
Die outside diameter	The largest diameter of a die, commonly referred to as the die O.D.
Die bore	The cavity of a die that accepts the product for compaction and determines the tablets size and shape configuration
Die groove	The radial groove around the die O.D. which accepts the die lock to secure the die in position in the die table
Die lock	The mechanism used to lock a die in position after it is installed in the die table
Die chamfer	The angled area between the top of the die and the die bore; it assists in guiding. the upper punch into the die bore
Die taper	A gradual increase in dimension, starting from a given depth in the die bore and increasing to the die chamfer; used normally to release air from the die cavity during the compression cycle

TABLE 2 Tablet Terminology

Term	Definition
Major axis	The largest dimension of a shaped tablet
Minor axis	The smallest dimension of a shaped tablet
End radius	The radius on either end of a capsule or oval-shaped tablet
Side-radius	The radius on either side of an oval or modified shaped tablet
Band	The center section of a tablet between the cup profiles: it is governed by a direct relationship of the die cavity profile.
Compound cup	A cup profile which consist of two or more radii
Embossed	The raised identification on a tablet or a punch face; an embossed punch tip results in a debossed tablet.
Debossed	The depressed identification on a tablet or a punch face: a dehossed punch tip results in a embossed tablet

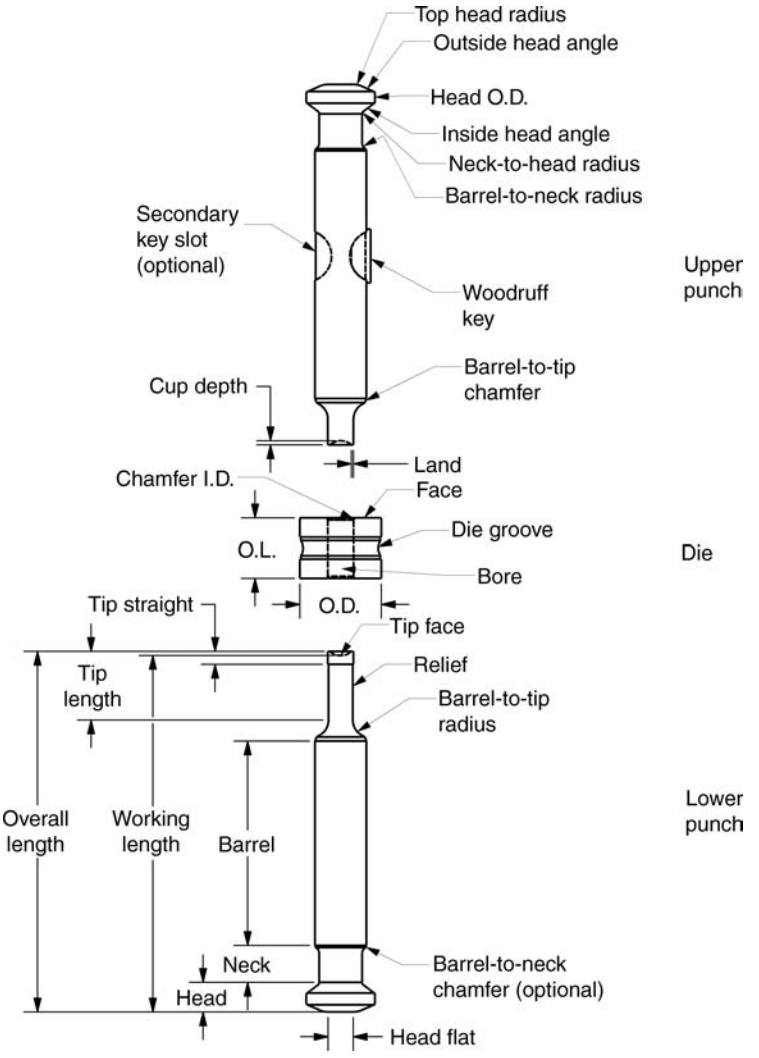


FIGURE 2 Tool drawing.

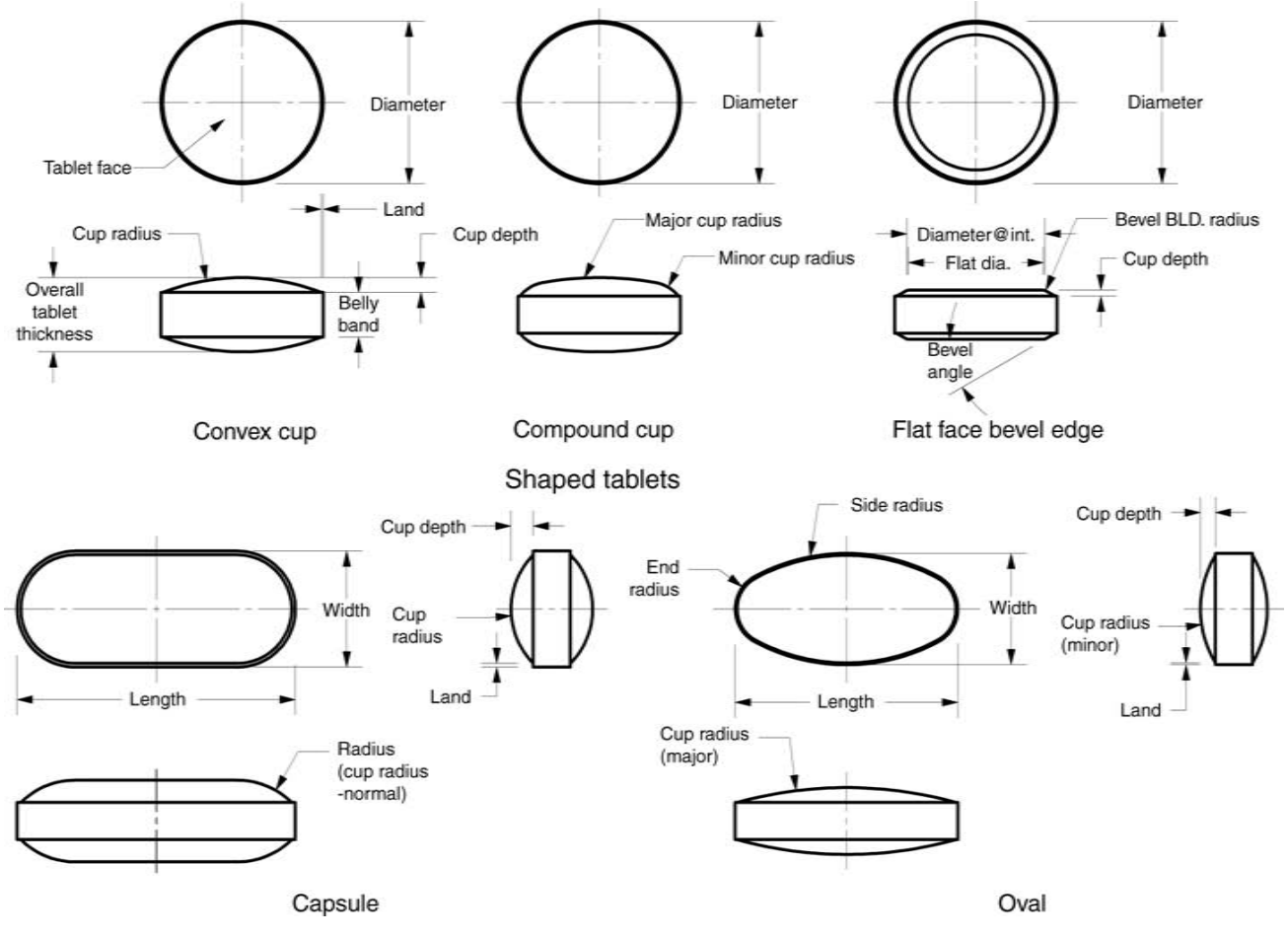


FIGURE 2A Tablet drawing.

The EU tooling standard is internationally recognized and is more widely used than its counterpart, the TSM standard. EU which is the acronym for "Eurostandard" and "Euronorm" is considered the European standard for interchangeable B and D type compression tools. The EU standards are authored by Mr. Trevor Higgins with the attempt to establish a tooling "norm" that provides tool interchangeability with the most common B and D type European tablet presses. The EU standard is printed and distributed by I Holland Ltd, Nottingham, England.

EU, TSM, B AND D TYPE PUNCHES

The TSM and EU standards manuals provide mechanical drawings and technical information for B and D type tools which constitutes a majority of the tool configurations used today. The B type configuration has a nominal punch barrel diameter of 0.750 in./19 mm. The B type has two different die sizes. The larger B dies have a diameter of 1.1875 in. (30.16 mm) and the smaller BB dies have a 0.945 in. (24 mm) diameter. The D type has a larger nominal barrel diameter of 1 in. (25.4 mm) and a die diameter of 1.500 in. (38.10 mm.) The B and D tool designation identifies the physical tooling size and was coined by Engineer Frank J. Stokes in the late 1800s.

Mr. Stokes resided in Philadelphia, Pennsylvania when he developed the first commercially available rotary tablet in the United States, the Stokes B1 Rotary. The B1 rotary press was extremely successful and most wanted by pharmaceutical companies nationwide. Mr. Stokes, realizing the need for compressing larger and heavier tablets, developed the Stokes D3 rotary tablet press. The D3 tablet press uses slightly larger punches and dies, increasing the overall capacity to compress larger and heavier tablets.

During the second industrial revolution, Mr. Stokes expanded manufacturing capabilities and operated a facility in England for international distribution. Stokes soon became the world's leading tablet press manufacture and sold tablet presses and tooling in nearly every industrialized country. The designation B and D quickly became the international standard for identifying a tablet press capacity and a tool configuration, as it still is today.

At the brink of World War II, Stokes left England and focused all manufacturing activities in Pennsylvania. Stokes left behind trained engineers and qualified manufacturing personnel who soon realized the potential of the tablet press market and began manufacturing tablet presses and tooling under the name Manesty. As a marketing strategy, Manesty re-engineered the punches and tablet press cams to enhance tooling life and provide better performance. The Manesty punch is similar to the original Stokes design, but is exclusive to Manesty presses and not interchangeable with the more popular Stokes tablet presses. Manesty called their tablet presses the "Manesty B3B" and the larger "Manesty D3a."

Manesty soon became a major supplier in the compression equipment industry and successfully competed against Stokes in the global market. In the mid-1980s the tablet press industry exploded and press manufactures were competing with tablet press output and innovation. Accommodating newer and high-speed tablet presses, the original Manesty tooling standard was refined to provide better interchangeability with the most common B and D tablet presses, identified by the "Eurostandard," often referred to as the EU standard and the EU norm (Fig. 3).

There are various models of tablet presses that do not conform to the standard B and D tool configurations and are engineered to be exclusive to a particular make and model of tablet press. Some of the more common configurations were designed in the

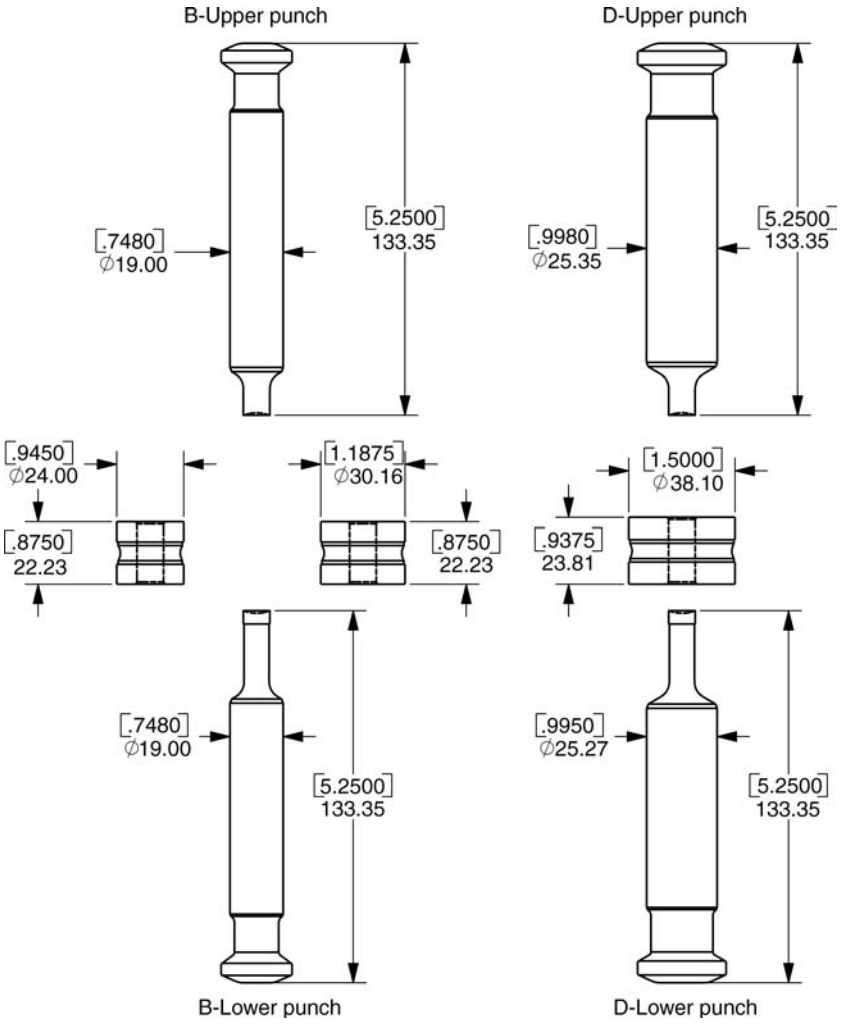


FIGURE 3 Drawing showing the differences between the B and D TSM and EU configurations.

early 1900s and still used on tablet presses today. These unique tablet presses are generally larger and engineered to compress larger tablets more effectively. Kilian Gmbh, a division of IMA in Milan, Italy, is a major European manufacturer of tablet presses using the most common unique tool configuration. The Kilian style upper punch does not use the common punch head configuration to guide the punches through the press cams; instead, the upper punch is guided by a machined cam angle located on the side of the upper barrel. The Kilian design provides a larger head flat, therefore, increasing the compression dwell time over the more popular B and D type tools (Fig. 5).

RECENT INNOVATIONS

New technology continues to introduce innovative tool configurations in the effort to provide better efficiency of tablet press speed, product yield, cleaning, and safety.

In 1997, Ima introduced a line of unique tablet presses called the Ima Comprima. The Ima Comprima models use an innovative approach with tool design and granulation

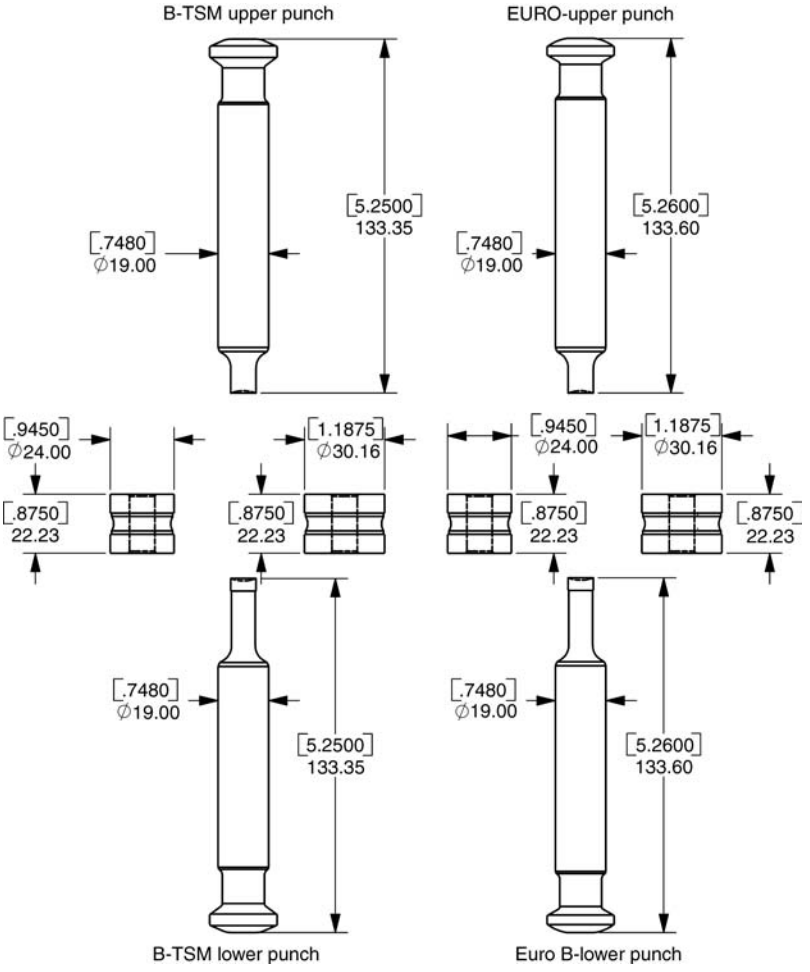


FIGURE 4 Drawing showing the differences between the B and D TSM and EU configurations.

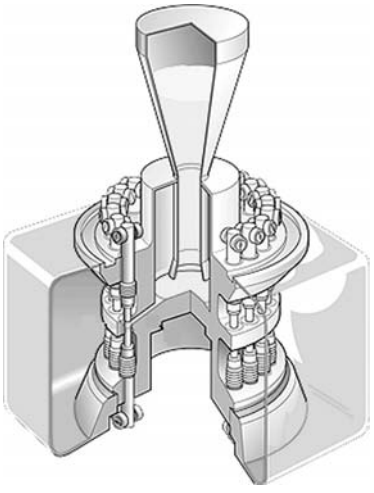


FIGURE 5 Drawing Kilian 27/32.

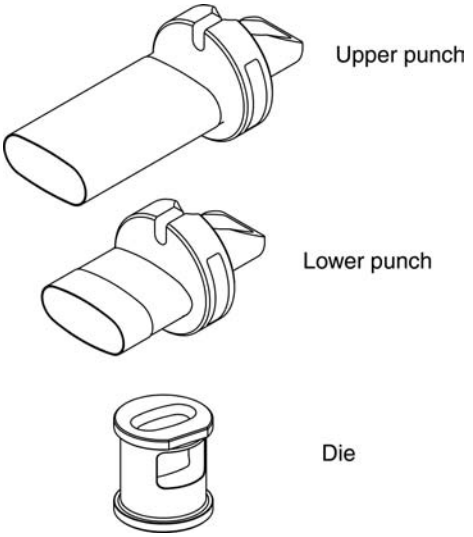


FIGURE 6 IMA press and tools.

delivery. Unlike traditional tablet presses using a gravity feed frame or force feeding mechanism to fill the die with granulation, the Ima Comprima feeds the granulation through the die table taking advantage of the centrifugal force created by the rotating turret for a rapid and uniform die fill. Unlike traditional presses, the Ima Comprima ejects the compressed tablet through the bottom of the die and uses gravity to eject the tablet from the press. Traditional tablet presses eject the tablet at the top of the die, requiring a mechanical stop or a take-off bar to physically contact and knock the tablet from the lower punch face. The Ima Comprima press is engineered to improve product yield, while providing a dust-free environment for a cleaner operation and a safer environment for the operator (Fig. 6).

The most recent innovation with tablet press and tooling technology is developed by Fette GmbH, located in Schwarzenbek, Germany. The new technology was introduced in 2005 and is being favored by high-volume tablet manufactures. The technology does not use traditional compression dies, instead Fette developed die segments. Die segments provide an advantage over traditional dies by combining the tablet press turret die table and dies into 3 or 5 integral segments. Die segments are much easier and quicker to install than individual dies and die locks, reducing tablet press set-up time dramatically. Because the concept does not require the use of dies, more space is available around the turret circumference to increase the number of punches, resulting in more tablets compressed per revolution than traditional presses of the same size (Fig. 7).

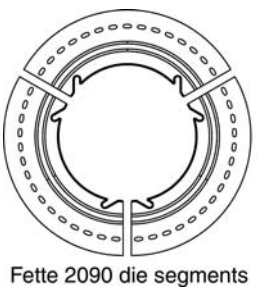


FIGURE 7 Drawing Fette die segment.

Tablet press technology has recently brought attention to the steel used for punches and dies with "wash in place" tablet presses. "Wash in place" tablet presses are becoming more common and available from most major tablet press suppliers. To reduce the possibilities of tool discoloration and corrosion, it is important that the tools are immediately removed and dried, if the tools can not be confirmed dry in the tablet press turret.

Cup Depth, Overall Length, and Working Length

Figure 8 shows these parameters and their corresponding tolerances. These are the most critical dimensions in any tooling program that relate directly to final tablet thickness, weight, and hardness. The overall length (OL), is a reference dimension, therefore, does not have a specified tolerance. A reference dimension is defined by the Machinery's Handbook (2) as:

A dimension, usually without a tolerance and used for information purpose only. It is considered auxiliary information and does not govern production or inspection operations. A reference dimension is the repeat of a dimension or is derived from other values shown on the drawing or on related drawings.

The two dimensions making up the punch OLs are the working length (WL) and the cup depth, with the exception of flat-face tip configuration which does not have a cup and is used to compress a wafer type tablet. The two dimensions are the WL dimension with a tolerance of plus or minus 0.001 in., and the cup depth, tolerance plus or minus 0.003 in. Combining the two tolerances that affect the OL of a punch, the calculated tolerance would be plus or minus 0.004 in. The major concern with these dimensions is to maintain consistency within a set of punches in order to maintain tablet weight, hardness, and thickness. The more critical of the two dimensions is the WL. The WL needs to be inspected as a single dimension and preferably for consistency within the given working-length tolerance, and not for a number formulated from the cup depth subtracted

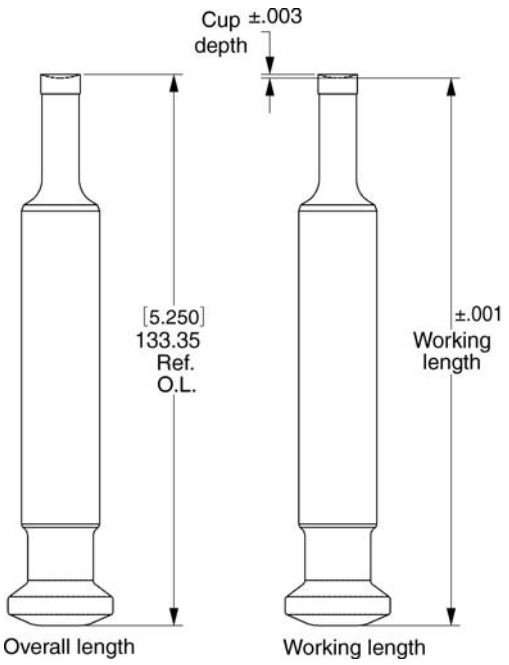


FIGURE 8 Drawing of punch showing CD, OL, and WL.

from the OL. A set of punches should be separated into uppers and lowers and inspected for variances as such. For example, all of the upper punches are checked for length consistency, and then all of the lower punches are checked as a separate unit. As long as both upper and lower punches fall within the desired tolerance range, tablet thickness, hardness, and weight will be consistent.

Although the cup depth is not responsible for tablet thickness, it should be confirmed within the given tolerance to maintain tablet overall consistency; it too should be inspected as single dimension.

Tooling Options

During the 1980s, the tablet compression industry was introduced to higher speed and more automated tablet presses assuring interchangeability with the TSM standard tool configurations. Although the standard tool configuration may be compatible, in some cases was not optimal and required minor modification to achieve expected performance. As well, the standard tool configuration may not be desirable for compressing certain products. All products are different and have unique characteristics, likewise may require slight tooling modifications. Tablet manufactures need to be informed of available options to achieve the best possible performance from the tablet press and tooling. Following is a description of tooling options that can be a benefit on both high-speed and standard presses.

COMMON TOOLING OPTIONS

Domed Heads

The domed head configuration is adaptable to both the upper and lower punch and maintains the identical top head radius and head flat as the "Eurostandard". It is an option only for the TSM head configuration and is compatible with the American TSM cams and should be considered for all high-speed tablet presses. As the speed of the tablet press continues to increase, tablet manufactures are coming to realize the advantage of the domed-style head with the larger top radius. The domed head style has several advantages over the standard TSM head profile. The larger 5/8 in. radius on the domed head reduces the enormous stress which is more common with the smaller 5/16 in. radius on a standard head when the punch makes initial contact with the pressure roller. This stress can cause a condition called head pitting which is identified by voids on the head flat. The impact of the pressure roller and head radius at high-speeds and heavy forces can cause a work-hardening effect, contributing to the pitting of the head flat. This form of pitting is detrimental to the life of the punches and pressure rollers. The domed head configuration provides a smoother transition into the compression cycle of the tablet press, reducing stress, and premature wear of the pressure rollers (Fig. 9).

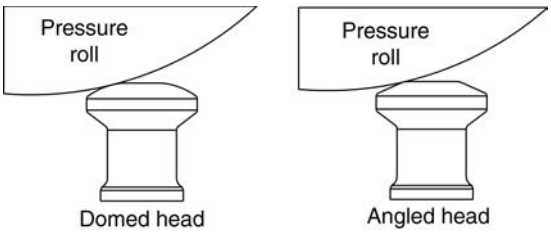


FIGURE 9 Differences between TSM and TSM Domed.

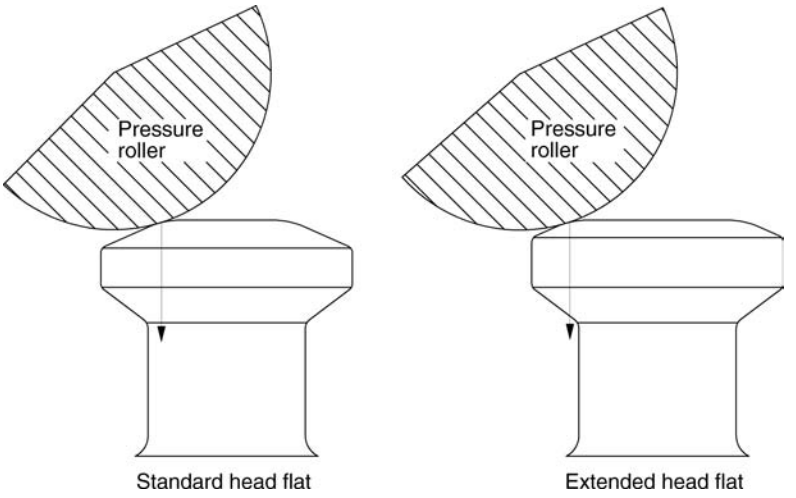


FIGURE 10 Drawing extended head flat and downward pressure on the head.

Extended Head Flat

Some tool manufactures will provide a head profile with a larger head flat. The advantage of the larger head flat is to increase the tablet press output and/or to increase the dwell time of compression. The disadvantage of the extended head flat is the possibility of head fracturing. Head fracturing can occur if the pressure roller makes contact to the head outside of the neck diameter. The initial contact of the pressure roller to the head should always be within the diameter of the neck to provide support (Fig. 10).

Rotating Heads

The rotating punch head is a two part punch configuration, the head is separate from the punch barrel and tip allowing the head to be removed and replaced as the head wears. When compressing round tablets, the punches will rotate as they are pulled around the cam track through the various stages of the tablet compression. As the punches rotate the wear and stresses on the back angle of the head is distributed around the entire back angle bearing surface. When compressing tablet shapes other than round the punches do not rotate, causing the wear to be concentrated at a single point, resulting in premature head wear. Because the rotating head configuration allows the head to rotate when compressing non-round tablet shapes, the wear is distributed along the entire surface of the back angle. This helps to decrease head wear and prolong the life of the punches (Fig. 11).

Mirror Finished Heads

Some high-speed tablet presses use heavy metal cams such as bronze and bronze alloys. This material is good for eliminating premature head wear and prolonging tool life, but it has a negative effect by contaminating the lubrication and turning it to a black, dark green color. The typical finish of a punch head is done with fine emery or fine abrasive pads. This finish leaves fine radial lines on the contact surfaces of the heads and has a filing effect on the softer cams, causing discoloration of the lubrication and premature cam wear. Polishing the punch heads with a soft cotton wheel and fine polishing compound to a mirror finish, helps to keep the lubrication cleaner and prolongs cam life.

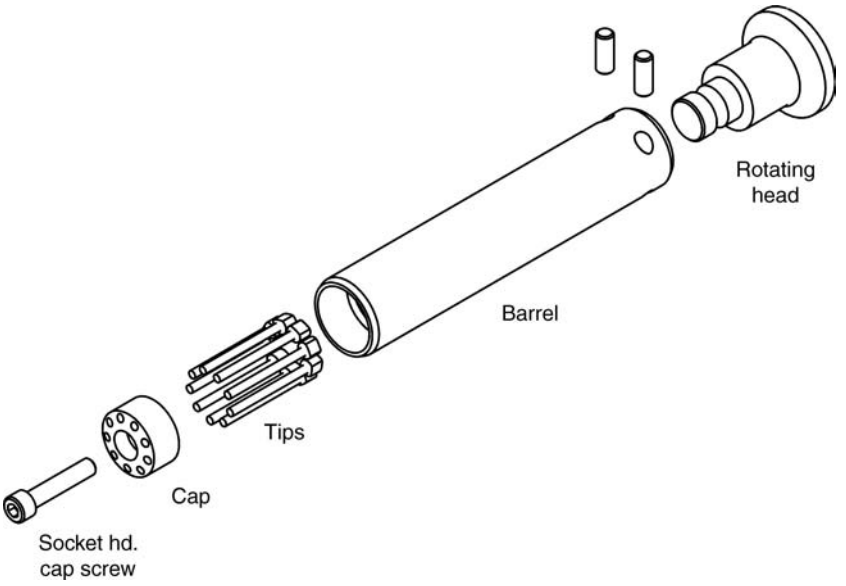


FIGURE 11 Exploded view of rotating head.

Bakelite Relief and Double Deep Relief

It is important to maintain a sharp edge around the lower punch tip relief. A sharp edge assists with the pull down cycle of the lower punch after tablet ejection. If residual product is adhered to the die wall, the sharp lower punch tip relief will help scrape the die clean as well as cutting through the product to reduce the possibility of product wedged and re-compressed between the punch tip and die wall. Product wedged between the punch tip and die wall may cause excessive heat and thermal expansion of the punch tip. This could result in punch binding and/or seizure, premature head wear, tablet discoloration or burning and dark specs contaminating the tablet. A bakelite relief assures a sharp edge to assist with removing product adhered to the die wall allowing the punch tip to move freely in the die. A "double deep relief" increases the depth of the lower punch relief and provides the same results as the bakelite relief; both designs are to assure a sharp edge at the punch tip. The bakelite relief is an added cost option for punches, whereas the double deep relief is generally a no charge option (Fig. 12).

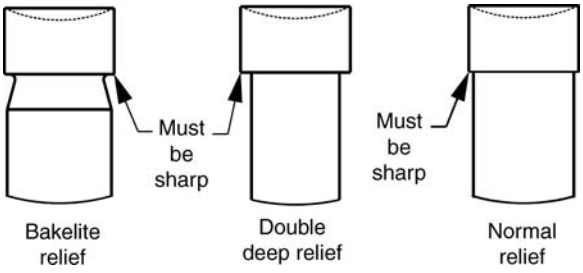


FIGURE 12 Drawing of bakelite relief and double deep relief

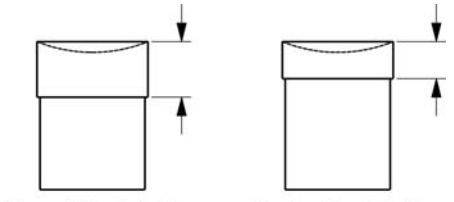


FIGURE 13 Drawing short tip straight.

Short Lower Punch Tip Straight

The lower punch tip creates a tremendous amount of friction as it travels the full length of the die through the various stages of tablet compression. When compressing sticky products or products with a low melting point, the friction created by the lower punch tip can cause lower punch binding. Reducing the bearing surface of the lower punch tip will reduce friction allowing the punch to travel easier in the die and reduce operating temperatures (Fig. 13).

Punch-Barrel Chamfers

Punch-barrel chamfers are required on punches used with presses fitted with rubber or plastic guide seals. The barrel chamfer has an advantage over the common break edge for these press models. The absence of a chamfer on the tip end of the punch can create difficulties while installing punches. Forcing the punch past the seal can cause damage to the seals, resulting in seepage of lubrication from the upper-punch guides, inherently causing product contamination. Damaged lower guide seals can allow product seepage into the lower-punch guides and mixing with the lubrication, causing tight punches, and possibly press seizure. A barrel chamfer on the head end of the punch can reduce wear of the punch guides caused from the punches being cocked from the torque of rotation as the punch travels vertically in the guides.

KEY TYPES AND POSITIONS

Punch barrel keys are mandatory for upper punches when compressing non-round tablets. The upper punch key maintains alignment of the tip for re-entry into the die for compression. Keys are not generally required for lower punches as the lowers do not leave the die during the compression cycle, so maintaining alignment is not required. Keys may also be required when compressing round tablets with embossing to eliminate the punch from spinning after compression, causing damage to the embossed tablet and reducing the likelihood of a "double impression" on the tablet face. The punches may also require keys when the orientation of the embossing for the top and bottom of the tablet is required to be constant.

Keys fitted to the upper punches are available in two configurations: (*i*) the standard Woodruff key, sometimes referred to as the pressed-in key; and (*ii*) the feather or flat key, often referred as the European key.

The Woodruff key, often referred to as the half moon key because of it's shape, is available in two styles, standard and the Hi-Pro. The Hi-Pro key has a tab on each side of the exposed top section and rests on the barrel. The taps keep the key secure by eliminating the rocking action common to the standard Woodruff. To obtain maximum security for high-speed presses, the Woodruff key is fastened into the barrel using screws. Because the

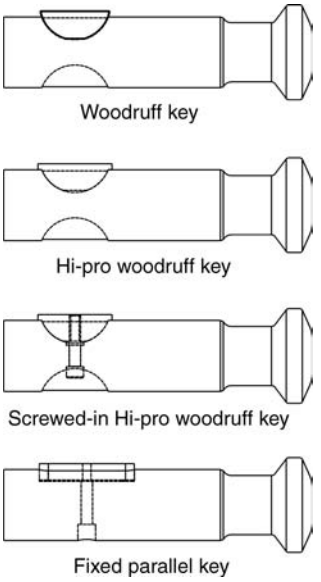


FIGURE 14 Drawing of key types.

Woodruff key is pressed into position, it can swell the barrel at the position of the key slot, causing excessive drag and sometimes galling of the upper punch and punch guide.

The feather key is a longer flat key, and comes in a variety of lengths, depending on the tablet press. Unlike the pressed in woodruff key, the feather keys fits into a milled slot and are secured into position using machine screws.

The height and radial position of a key is critical to obtain maximum press performance. Unfortunately no standard has been established due to the particular requirements of the many styles of tablet presses. If the key is placed too low or is too long, it can interfere with the upper punch guide seal and cause damage and/or seepage of lubrication, resulting in product contamination. If the key is too high, it can travel out of the key slot at the top of the punch guide, resulting in severe damage to the punches and press (Fig. 14).

TOOL CONFIGURATION FOR SMALL AND MICRO TABLETS

It is common to experience difficulties maintaining tablet hardness, thickness, and weight while compressing small and micro tablets. Compression force is sensitive and will generally require minimum forces. In some cases the tablet is compressed by the weight of the punch. Excessive tonnage can distort the punch tip and alter the critical WL, making tablet consistency virtually impossible. Tip breakage is also frequent and can damage additional punches and the tablet press, most commonly the feed frame. A special tool configuration is recommended for compressing tablets smaller than 0.125 in. (3 mm). This configuration modifies the punches and dies and is used in conjunction with a shallow fill cam that is fitted on the press to minimize lower punch travel in the die. The punch modification involves shortening the punch tips and eliminating the lower punch relief. Shortening the tip straights to their minimum length will strengthen the tip increasing the maximum compression force considerably. The lower punch tip relief is removed to reduce the clearance between the tip stem and the die bore, providing

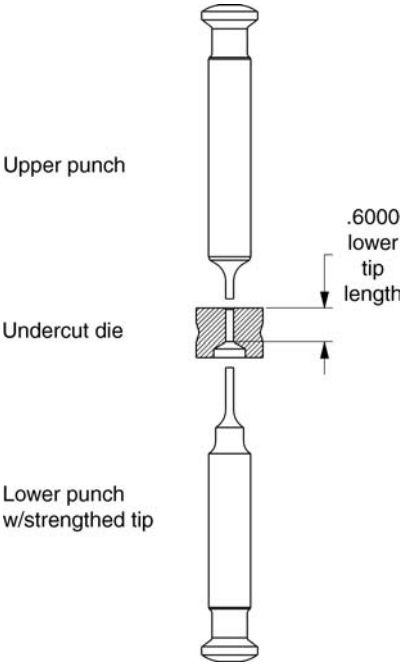


FIGURE 15 Exploded view of single tip punches with strengthened lower tip and undercut die.

additional support to the tip stem, decreasing distortion. Reducing the tip length increases the barrel length; therefore the bottom of the die is undercut to accept the longer barrel for tablet ejection (Fig. 15).

Tapered Dies

A tapered die has numerous advantages. A die can be tapered on one side or on both sides, with the advantage of turning the die over and doubling its life. The biggest advantage of a tapered die is to exhaust trapped air in the product as the upper punch enters the die at the beginning of the compression cycle. This is especially helpful for deep-cup punches, fluffy granulation, and high-speed presses. A tapered die provides the ability to compress a harder tablet with the same amount of pressure as required with a straight die. It is helpful in reducing capping and laminating. Taper will allow the tablet to expand at a slower rate as it is being ejected from the die, reducing stress that can cause lamination and capping. Taper decreases the ejection force, prolonging the life of the lower punch heads and ejection cam, thus reducing friction and allowing the press to operate at a lower temperature. Tapered dies help align the upper-punch tip upon entering the die, eliminating premature tip wear; this is especially helpful for presses with worn upper-punch guides. A standard taper on a BB or D die is 0.003 in. by 3/16 in. deep. Die taper can be tailored to meet special requirements. Although there are numerous advantages with using taper there are disadvantages as well. Because the taper is conical with the largest area at the top, the upper punch can wedge in-between the punch tip and die wall as it is pressed into the die. Excess product can migrate between the punch tip and die bore due to the additional punch tip to die bore clearance as a result of the taper. If the upper punch is wedged and sticks in the die it will be evident by spotty tablets and/or premature wear at the back angle of the upper punch (Fig. 16).

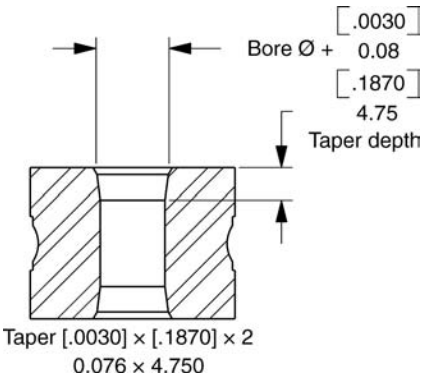


FIGURE 16 Drawing of taper dies.

Tablet Designs

Proper punch face contour is essential for tooling life and tablet quality. The compression force should be determined during the R&D phase of a new product. If heavy compaction forces are required then a shallow or standard cup configurations should be considered to assure satisfactory tooling life and tablet quality. If the compaction force is to remain light to standard, a variety of configurations may be considered. Compression force has a lateral force that can expand the sides of the punch cup outward toward the die wall. Figure 17 shows the flexing w arrows in the cup. Excessive pressure can permanently distort and cause premature failure of the punch tip. For a high-compaction force the cup may be strengthened by:

1. Increasing the land area on the punch tip to provide additional strength;
2. Reducing the hardness of the punch tip, allowing the tip to flex without breaking;
3. Increasing the cup radius or decreasing the cup depth to eliminate the damaging effect of flexing and abrasion to the inside of the cup.

The flat-face bevel edge (FFBE) tablet configuration is subjected to the same lateral force. These edges can be strengthened by steps 1 and 2 and by increasing the radius between the flat and the bevel which is normally 0.010–0.015 in. The flat-face radius-edge (FFRE) configuration provides a stronger punch tip than the FFBE and can eliminate edge chipping by reducing sharp corners on the tablet face. Another common cup configuration is the compound cup. The compound cup has two radii which makes the

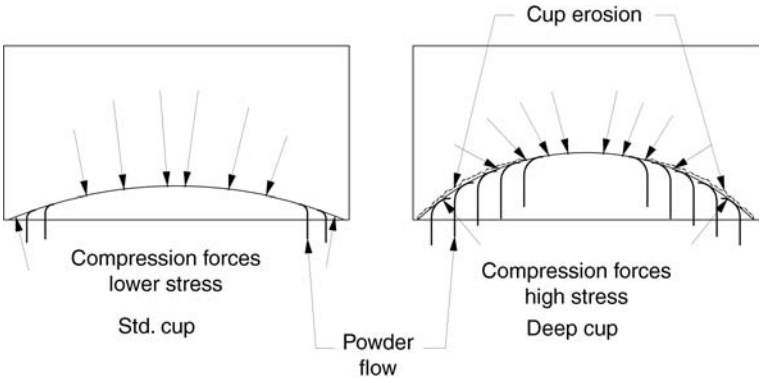


FIGURE 17 The flexing w arrows in the cup.

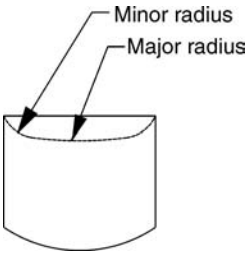


FIGURE 18 Detail of CC cup.

tablet roll better during the coating process, eliminating tablet edge erosion. The compound cup design generally has more cup volume and is the optimum tablet design for heavy tablets, as it generally reduces the tablet band giving the tablet a thinner appearance. However, the compound cup is one of the weakest tablet designs due to the stresses created at the intersection of the two cup radii and the steep cup which causes excessive abrasion during compression, shortening the tool life (Fig. 18).

Elaborate three-dimensional cup configurations are becoming more common in the candy and vitamin industry. Because of the high and low cup designs, it is critical that compaction forces are determined during the R&D phase and results provided to the tooling manufacture.

The concavity standards for round punch tips are published in the TSM. These standards (Table 3) include cup depths for shallow, standard, deep, extra deep, modified ball, FFBE, and FFRE. For radius cup designs, the TSM identifies the cup by the cup depth, whereas the European tableting industry identifies the cup by the cup radius. Figure 19 shows a TSM standard cup and an EU standard cup identifying the radius.

Tablet Shapes

There are as many tablet shapes as there are applications, which are endless. Tablets are used in automobile air bags, batteries, soaps, fertilizers, desiccants, and buttons just to name a few. Historically, round tablets were most common, uncomplicated and easy to set-up and to maintain. Special-shape tablets are tablet shapes other than round and include shapes such as capsule, oval, square, triangle. etc. Exotic shape tablets are more unique than round or special shapes. Exotic shaped tablets include animal and heart shaped tablets and other unique tablet shapes requiring an internal radii or angle. A unique tablet shape will provide better tablet identification helping to maintain consumer interest and loyalty (Fig. 20).

The most common special shapes in the pharmaceutical industry are the capsule, modified capsule, and oval shapes. These shapes typically accommodate more volume and are more unique than standard rounds. A film-coated tablet is better to use with a

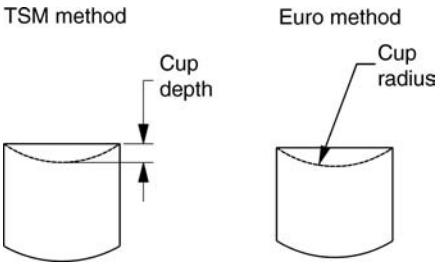
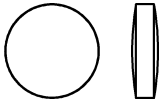
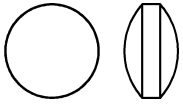
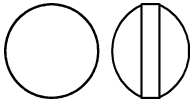


FIGURE 19 TSM standard cup and an EU standard cup identifying the radius.

TABLE 3 TSM Cup Depth of Single Radius Tablet Configurations

Tablet diameter						
Inches [millimeters]	Shallow cup depth	Standard cup depth	Deep cup depth	Extra deep cup Depth	Mod. ball cup depth	F.F.B.E./ F.F.R.E. cup depth
1/8 [3.175]	0.005 [0.127]	0.017 [0.432]	0.024 [0.610]	0.030 [0.762]	0.040 [1.016]	0.007 [0.178]
5/32 [3.970]	0.007 [0.178]	0.021 [0.533]	0.030 [0.762]	0.036 [0.914]	0.049 [1.245]	0.008 [0.203]
3/16 [4.763]	0.008 [0.203]	0.029 [0.737]	0.036 [0.914]	0.042 [1.067]	0.059 [1.499]	0.009 [0.229]
7/32 [5.555]	0.009 [0.229]	0.026 [0.635]	0.042 [1.067]	0.048 [1.219]	0.069 [1.753]	0.010 [0.254]
1/4 [6.350]	0.010 [0.254]	0.031 [0.787]	0.045 [1.143]	0.050 [1.270]	0.079 [2.007]	0.011 [0.279]
9/32 [7.142]	0.012 [0.305]	0.033 [0.838]	0.046 [1.168]	0.054 [1.372]	0.089 [2.261]	0.012 [0.305]
5/16 [7.938]	0.013 [0.330]	0.034 [0.864]	0.047 [1.194]	0.060 [1.524]	0.099 [2.515]	0.013 [0.330]
11/32 [8.730]	0.014 [0.356]	0.035 [0.899]	0.049 [1.245]	0.066 [1.676]	0.109 [2.769]	0.014 [0.356]
3/8 [9.525]	0.016 [0.406]	0.036 [0.914]	0.050 [1.270]	0.072 [1.829]	0.119 [3.023]	0.015 [0.381]
13/32 [10.318]	0.017 [0.432]	0.038 [0.965]	0.052 [1.321]	0.078 [1.981]	0.128 [3.251]	0.016 [0.406]
7/16 [11.113]	0.018 [0.457]	0.040 [1.016]	0.054 [1.372]	0.084 [2.134]	0.133 [3.378]	0.016 [0.406]
15/32 [11.905]	0.020 [0.508]	0.041 [1.041]	0.056 [1.422]	0.090 [2.286]	0.148 [3.759]	0.016 [0.406]
1/2 [12.700]	0.021 [0.533]	0.043 [1.092]	0.059 [1.499]	0.095 [2.413]	0.158 [4.013]	0.016 [0.406]
17/32 [13.493]	0.022 [0.559]	0.045 [1.143]	0.061 [1.549]	0.101 [2.565]	0.168 [4.267]	0.016 [0.406]
9/16 [14.288]	0.024 [0.610]	0.046 [1.168]	0.063 [1.600]	0.107 [2.718]	0.178 [4.521]	0.016 [0.406]
19/32 [15.080]	0.025 [0.635]	0.048 [1.219]	0.066 [1.676]	0.113 [2.870]	0.188 [4.775]	0.016 [0.406]
5/8 [15.875]	0.026 [0.660]	0.050 [1.270]	0.068 [1.727]	0.119 [3.023]	0.198 [5.029]	0.016 [0.406]
11/16 [17.463]	0.029 [0.737]	0.054 [1.372]	0.073 [1.854]	0.131 [3.327]	0.217 [5.512]	0.020 [0.508]
3/4 [19.050]	0.031 [0.787]	0.058 [1.473]	0.078 [1.981]	0.143 [3.632]	0.237 [6.020]	0.020 [0.508]
13/16 [20.638]	0.034 [0.864]	0.061 [1.549]	0.083 [2.108]	0.155 [3.937]	0.257 [6.528]	0.020 [0.508]
7/8 [22.225]	0.037 [0.940]	0.065 [1.651]	0.089 [2.260]	0.167 [4.242]	0.277 [7.036]	0.020 [0.508]
15/16 [23.813]	0.039 [0.991]	0.069 [1.753]	0.094 [2.388]	0.179 [4.547]	0.296 [7.518]	0.020 [0.508]
1 [25.400]	0.042 [1.067]	0.073 [1.854]	0.099 [2.515]	0.191 [4.851]	0.316 [8.026]	0.025 [0.635]

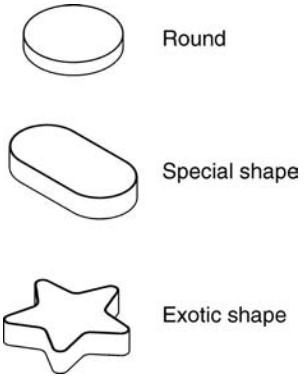


FIGURE 20 Drawing of round, special and exotic shaped tablets.

modified capsule rather than a capsule shape, to eliminate twinning during the coating process. A modified capsule shape can be designed to have the appearance of a capsule shape with the advantage of a radius on the major axis, reducing the contact surface area during the coating process (Fig. 21).

Tablet Face Configurations

Tablet shapes are virtually infinite as are tablet face configurations. The tablet face configuration is commonly referred to as the "cup" of the punch. The cup is the area at the tip end of the punch that is responsible for the configuration of the top and bottom of a tablet. The TSM provides cup depth standards for the six most common cup configurations for round tablets.

The TSM defines the cup depth of single radius tablet configurations by the depth of the concavity and is differs from the EU configurations which uses the cup radius value. The cup radius is more difficult to check and to set internal limits for reworking.

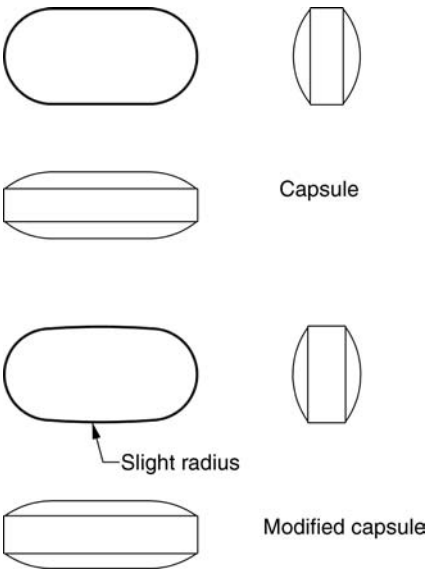


FIGURE 21 Capsule and modified capsule.

A single radius cup is the strongest cup configuration and is the most common configuration for round tablets. Adding another radius to the cup changes the cup configuration to a compound cup or a dual radius cup. The compound cup has an advantage of having more volume than the single radius cup. Increased volume to the cup will reduce the size of the "Belly Band" making the tablet appear to be thinner and easier to swallow. The configuration of compound cup is better for film coating. The rounded edges tend to roll better in the coating pan reducing the possibilities of edge erosion. There are several disadvantages to using the compound cup design. The intersection of the two cup radii becomes a high-stress point which is prone to failure under extreme loading, therefore has a much lower maximum compression force rating than the single radius shallow and standard cup. Extreme loading is not uncommon with the compound cup configuration. The compound cup has more volume; therefore as the upper punch cup enters the die, it fills the die with air, and then must be extracted before compression. Because of this, the compound cup commonly requires slower press speeds or higher compression force than a single radius shallow or standard cup. The compound cup sidewall is steep and receives high-abrasion as the tablet is being compressed, wearing the tip and weakening the cup. The tip land is critical to the punch tip strength and should be checked often for wear. If the land wears thin it will cause a condition known as "J hook" which is a common cause of capping and laminating. The land is easily refurbished using 400 grit sharpening stones and a large cotton buff wheel. The compound cup design has a smaller window or available space for engraving and printing than the single radius shallow and standard cup.

Three-dimensional cup configurations are common with vitamins and candies. The three-dimensional cup configuration provides raised features on the tablet surface providing the opportunity to sculpt features and character details.

Undesirable Shapes

A tablet shape too close to round may cause a condition known as punch-to-die binding or self-locking. These shapes need to be avoided in order to provide maximum tablet output and satisfactory tool life (Fig. 22).

The corner radius of a special shape such as a square and triangle is critical for maintaining the strength and integrity of the die. A corner radius less than 0.032 in. can cause excessive stress and failure as the die is locked into position with the die lock and subjected to the shock of tablet compression (Fig. 23).

TABLET IDENTIFICATION

There are two basic methods for identifying a tablet, printing and engraving; the latter is the most common. There are two styles of engraving, embossed and debossed. With debossing, the identification is raised on the cup face and engraved into the tablet, while embossed identification is cut into the cup face and raised on the tablet (Fig. 24). These two styles can be used in conjunction with each other.

To ensure product identification many companies engrave their corporate logos on their product line. As tablet size decreases, the legibility of the identification tends to diminish, eventually reaching the point at which it is no longer legible. For this reason, tablet manufactures should consider the entire range of tablet size when considering the format of a logo for better legibility. As a tablet decreases in size,

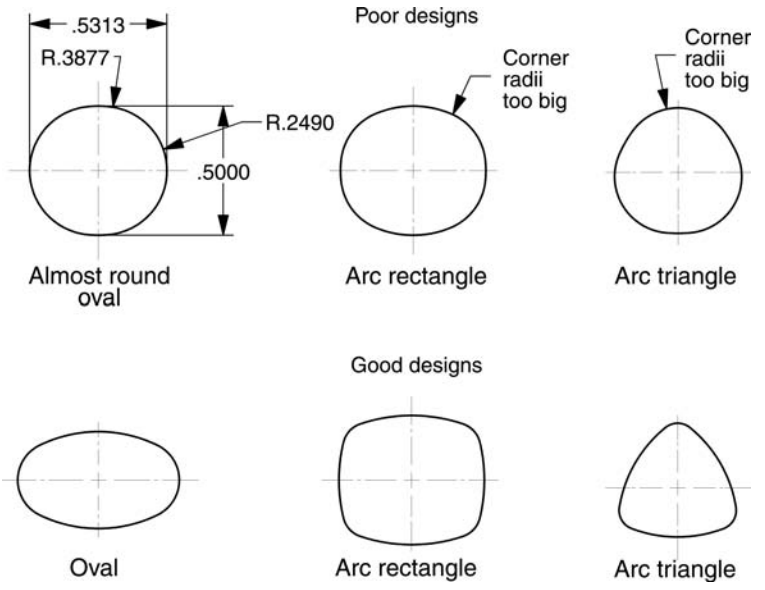


FIGURE 22 Undesirable shapes.

the logo and drug code are subject to picking (product sticking in or around the identification). Because some products are more prone to picking than others, formulation data and product history, if available, should be provided to the tooling manufacturer so that they may engineer an engraving style and format to help minimize picking and sticking.

A company that engraves or embosses most or all of their tablets should consider maintaining a character font. The font should be designed to eliminate sharp corners

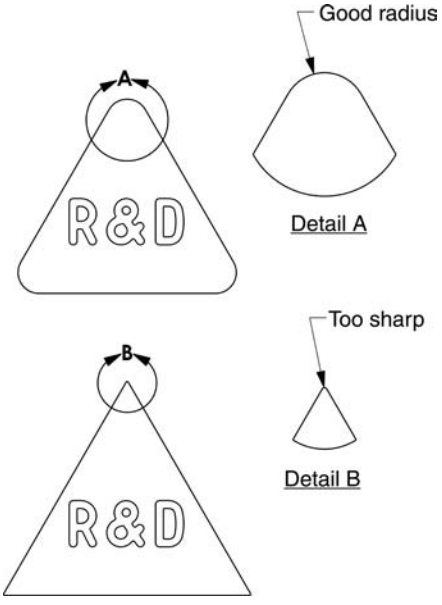


FIGURE 23 Drawing showing good and bad corner radius.

FIGURE 24 Raised embossing in a panel.

BAD FONT

GOOD FONT **FIGURE 25** Sample fonts good and bad.

whenever possible and opening closed-in areas of a character as much as possible (Fig. 25).

For sticky products, the engraving style can be designed to pre-pick the islands of a character, for example, filling in the centers of the B, R, 0, 8, etc. The pre-pick character can be difficult to film coat and is prone to fill in and bridging therefore for film coated tablets the characters can be partially pre-picked. A partial pre-pick is generally preferred and only removes a percentage of the island instead of removing the island completely (Fig. 26). A ramped engraving style, also referred to as a tapered peninsula, provides the same advantage as a pre-picked style and used at the outside corners and open areas of a character. It provides a lower depth of these areas and then tapers the tablet surface (Fig. 26).

The radius at the top of an engraving cut at the tablet surface can be a main contributor to picking and tablet erosion. A general guide for the value of the radius is approximately one third of the engraving cut depth. For example, if the engraving cut depth is 0.012 in. then the radius at the top of the engraving should be 0.003 in./0.004 in.

The angle of a standard engraving cut for a non-coated tablet is 30°. If sticking occurs, it is recommended to increase the angle to 35°–40° which is the angle recommended for film-coated tablets. The wider engraving angle may diminish legibility of the engraving cut by allowing more light into the bottom of the cut, but has a better draft angle which provides improved product release (Fig. 27).

Incorrectly placing an engraving cut too close to the tablet edge or to close to the secondary radius for compound cups can result in punch tip fracturing. Although tooling manufacturers generally maintain certain guidelines for the layout and configuration of the engraving, they must consider the amount of engraving in relation to the tablet size, tablet configuration, and product characteristics before releasing the final tablet design for approval.

Bisects

Bisects, commonly known as a score or break line, are available in a variety of styles (Fig. 28). The purpose of a bisect is to break the tablet into a predetermined dosage, most commonly two equal parts. Breaking a tablet into prescribed dosages should give the

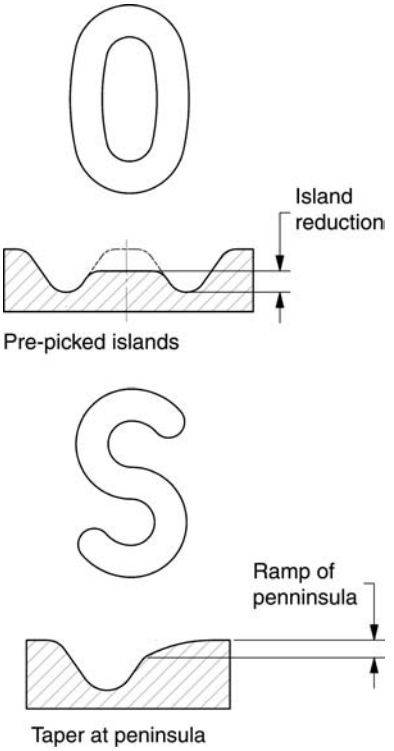


FIGURE 26 Pre-picking and tapered peninsula.

consumer a certain degree of confidence that they are receiving the proper dosage. Bisects should be placed on the upper punch whenever possible. Placing the bisect on the lower punch can create problems when the take-off bar removes the tablet from the lower punch. The depth of the bisect is generally deeper than the engraving cut, therefore making it difficult to slide the tablet across the punch face at the ejection cycle. The standard TSM bisect has two different configurations for concave tablets, protruding and cut flush. The protruding bisect style follows the curvature of the cup and extends past the tip edge of the punch. This style helps break the tablet into equal parts, because the extended bisect is pressed into the tablet band. The problem with this style is that the protruding bisect may run into the tip edge of the lower punch if they become too close during tablet press set-up or if the tablet press continues to cycle after the hopper has been emptied. Hitting the bisect into the lower punch edge will leave deep impressions while smashing and swelling the protrusion of the bisect on the upper punch. This is the reason the standard cut-flush bisect has become more popular (Fig. 28).

A cut-through bisect, also known as a European style bisect, can only be used on radius cup designs. It has an advantage over the standard bisect by allowing the consumer to easily break the tablet into equal dosages. The cut-through bisect is wider at the center of the tablet than the standard bisect, which reduces the available engraving space on the tablet face. The height of the cut-through bisect is generally the same as the cup depth.

Steel Types

Choosing a steel type is generally left up to the tooling manufacture, unless a specific type has been requested. The criteria for selecting a steel type includes the quantity of

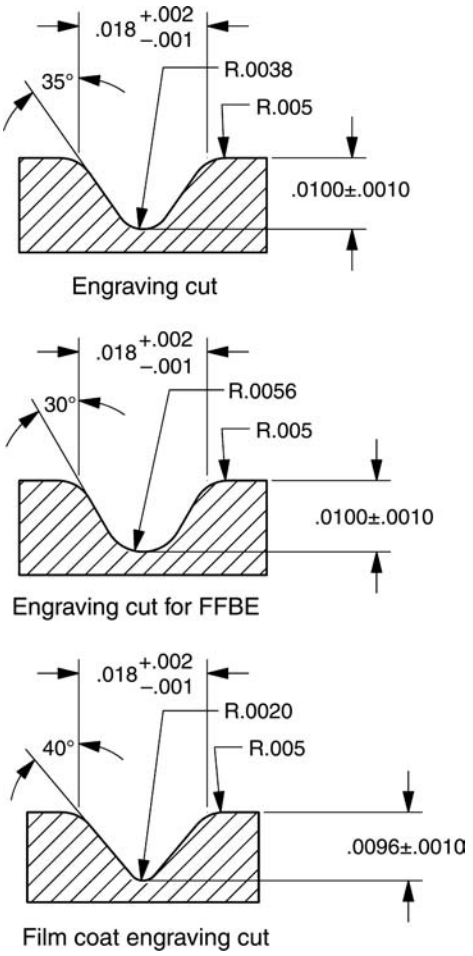


FIGURE 27 Engraving cut angles.

tablets to be produced, the abrasiveness or corrosiveness of the granulation, the pressure required for compression and the cup configuration.

There are two categories of steel common to this industry, standard and premium. Although the category names may imply that one is superior in quality to the other, this is not the case. Standard steels are the most common grades used and premium steels are for special applications. The cost is generally higher for premium steels due to the quality of the steel purchased by the tooling manufacturer and the steel composition. Premium steels tend to be harder, but at the same time more brittle than standard steels, prone to fracturing under excessive pressure and may not be suitable for deep cup configurations. Standard steels are available of the following grades: S-5, S-7, S-1, and 408. Premium steels are available in D-2, D-3, 440-C stainless steel and 0-1. Table 4 shows the toughness-wear relationship:

Inserted Dies

Dies are usually manufactured from D-3 premium steel. This grade does not provide toughness, but is superior for wear. Dies are not subjected to the same pressures or shock as the punches, and therefore can be manufactured from a larger selection of materials.

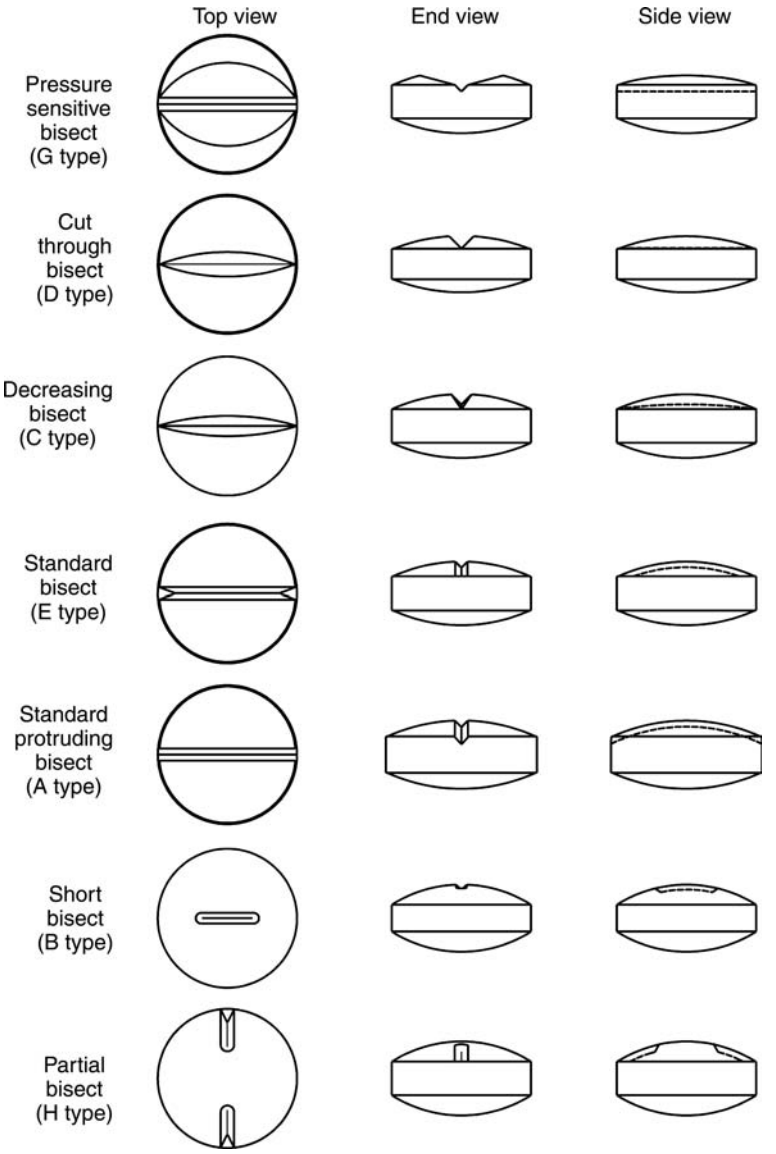
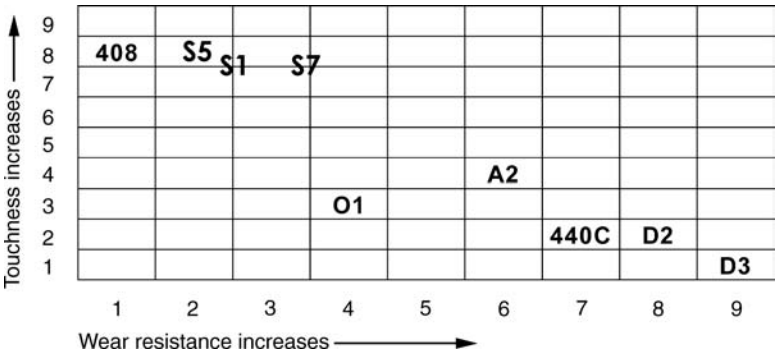


FIGURE 28 Bisects.

The most common die for abrasive formulations is the carbide-lined die. The carbide insert is heat shrunk into a softer steel sleeve which provides a cushion for the brittle carbide. These sleeves, fitting of the die O.D. and the die groove, are normally made of S-5 and A-2 tool steel. Carbide dies demand a much higher investment which is justified by superior die wear and tablet quality; die life is easily increased by 10 times in most cases. Because the carbide die is much harder, it is more brittle and subject to fracturing under excessively heavy loading. If the carbide liner is too thin at its narrowest point, it can fracture due to die lock pressure and stresses of compression. This is also true for the steel sleeve. The tooling manufacture should be consulted to determine if a tablet size is acceptable for a carbide liner.

TABLE 4 The Toughness-wear Relationship



When inserting carbide dies into the die pocket, a die driving rod fitted with a nylon tip should be used to prevent carbide fracturing. Die lock pressure should also be reduced by 10%. Ceramic-lined dies are becoming more widely used as tougher grades become available. The most common ceramic grade used in compression dies is currently partially stabilized zirconia (PSZ). Dies lined with PSZ have the same general wear characteristics and require the same precautions as carbide-lined dies but have an advantage in reducing the friction coefficient during the fill and ejection cycles. The ceramic liner is commonly a light cream or white color and is quickly gaining in popularity over carbide.

MULTI-TIP TOOLING

Normally one punch compresses one tablet, the exception is using multi-tip tooling. Multi-tip tools are more common in Europe and only recently accepted in the United States. The multi-tip tool configuration is engineered to compress more than one tablet at a time with the total number of tablets dictated by the punch size, tablet size, compression and ejection force, and the characteristics of the granulation.

There is a tremendous advantage using multi-tip tooling when considering production, operating efficiency, and overall capacity. Operator safety, multiplying the number of tablets produced in a given area, eliminating the need for additional room and tablet presses are only a few of the advantages. Increasing production by the multiple of punch tips can be achieved but should not be expected. Using the formula, Tablets currently produced $\times$ number of punch tips $\times 0.9 =$ number of tablets expected, will provide a more accurate estimate.

Multi-tip punches are available in two configurations, as a solid punch or an assembly with multiple parts. The solid punch configuration is easier to clean and assures alignment of the punch tips in the die; unfortunately if only one tip is damaged the entire punch is unusable and discarded. The solid configuration is more difficult to polish individual punch faces using a soft cotton wheel. The punch assembly separates the punch tips from the punch body and are secured using a cap and/or set screws. If a punch tip is damaged, it is simply removed and replaced, putting the punch back into service. To properly clean the assembly it must be disassembled, cleaned, dried thoroughly, and reassembled which can require substantial labor.

Tablet compression and ejection force becomes greater as does operating temperature and should be monitored closely to reduce premature wear and tablet sticking

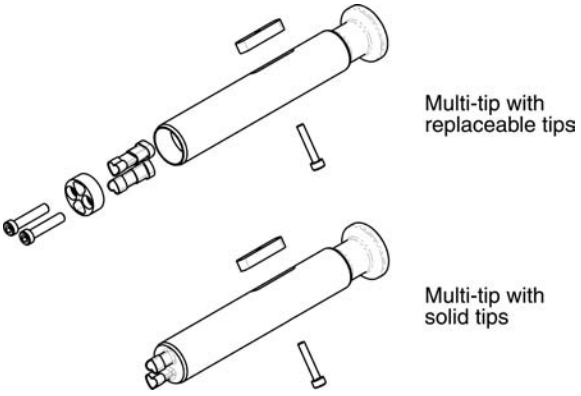


FIGURE 29 The solid punch and multiple piece punch exploded view.

and/or discoloration. Premature tooling wear will be evident by excessive wear on the punch head and tablet press cams.

It is recommended to use the rotating head option for the lower punch. The torque of the rotating turret tries to spin the punch in the guide. The rotating head will reduce the stress by spinning, thus taking pressure from the punch tips allowing the punch tip to travel the length of the die without binding (Fig. 29).

Punch-Tip Pressure Guide

Punch tip pressure guides, originally calculated by tablet press manufactures, are available and based on the tablet configuration and steel type. With the assistance of computer aided designing and finite element analysis (FEA) software, tooling manufactures have become more accurate with the maximum tonnage for round and shaped punch tablet designs.

Table 5 gives the cup configurations with the corresponding maximum tonnage force for round punch tips. This guide has been calculated from the computer-generated procedure FEA and is the most accurate guide available.

Calculating the maximum compression force for shaped tablets (i.e., capsule oval, etc.) can be difficult and confusing. It is recommended to contact the tooling supplier and request these values. The maximum tonnage for round and shaped tablets should be provided on the engineered tablet drawing provided by the tooling supplier along with the cup volume and surface area. It is important that these values have a strong presence with R&D and are used when formulating a new product. The tonnage requirement should be acceptable before the product reaches the production phase. If tool failure is experienced at the R&D phase, the tablet can be redesigned to accept the required tonnage.

Care of Punches and Dies

Punches and dies are precision instruments and can damage easily, so great care must be taken when cleaning, transporting, and storing. Upon receiving punches they should be cleaned and dried thoroughly prior to use. If standard operating procedures require incoming inspection, then the tools should be inspected immediately and any concerns or discrepancies reported to the supplier before the tools are used and/or put into storage for future use. Following inspection, the tooling should be lightly oiled, carefully packed in a protective container, and stored in a dry place.

TABLE 5 Maximum Compression Force by Cup Depth (Kilonewtons)

Punch tip diameter	Shallow concave	Standard concave	Deep concave	Extra-deep concave	Modified ball	F.F. B.E.	F.F. R.E.
1/8	12.5	4.4	2.7	1.8	1.0	3.7	4.9
5/32	18.0	7.0	4.2	3.1	1.6	5.3	7.6
3/16	27.0	9.6	6.1	4.7	2.2	7.2	11.0
7/32	37.0	14.0	8.3	6.7	3.0	9.3	14.9
1/4	49.0	20.0	12.5	10.5	3.9	11.5	19.5
9/32	60.0	27.0	18.5	14.5	5.0	14.0	25.0
5/16	75.0	37.0	26.0	18.0	6.1	16.5	30.0
11/32	92.0	48.0	34.0	22.0	7.4	19.0	37.0
3/8	107.0	61.0	44.0	26.0	8.8	22.0	44.0
13/32	127.2	73.0	55.0	30.0	10.5	25.0	51.0
7/16	149.0	87.0	67.0	35.0	13.5	29.0	60.0
15/32	168.0	104.0	79.0	40.0	14.0	33.0	68.0
1/2	192.0	120.0	92.0	47.0	16.0	38.0	78.0
17/32	219.0	137.0	107.0	53.0	18.0	43.0	88.0
9/16	242.0	159.0	123.0	59.0	20.0	48.0	99.0
19/32	271.0	179.0	139.0	66.0	22.0	53.0	110.0
5/8	302.0	200.0	157.0	73.0	24.0	59.0	122.0
11/16	363.0	246.0	195.0	88.0	30.0	63.0	147.0
3/4	436.0	296.0	238.0	104.0	36.0	75.0	175.0
13/16	509.0	356.0	284.0	122.0	42.0	89.0	206.0
7/8	587.0	417.0	331.0	142.0	48.0	103.0	238.0
15/16	679.0	482.0	286.0	163.0	56.0	118.0	274.0
1	770.0	552.0	445.0	185.0	63.0	119.0	311.0

When tooling is required to be shipped, they should not be shipped in storage containers. Most storage containers are not designed to support the weight of the tooling through the handling practices of commercial shipping companies. Tooling should be returned in their original individual plastic or cardboard shipping containers and packed tightly to avoid movement. Because punch tips are extremely fragile they should be protected at all times from hitting each other or hard surfaces. A dent or nick on a punch tip can keep the punch from fitting properly into the die. To avoid damage to the die during set-up, a proper driving rod should be used when inserting the die in the die table. A mild steel rod with the same diameter as the punch guide fitted with a nylon tip is recommended. To prevent damage to the die, die table, and die lock, the die lock pressures indicated by the tablet press manufacturer's operator's manual should be observed. Excessive die lock pressure can distort the die bore and cause punch tightness, fracture the die, and even crack the die table costing thousands of dollars to repair.

TOOLING INSPECTION

Tooling inspection programs are becoming more common and performed as a precautionary measure to reassure critical dimensions and embossing details. Confirming critical dimensions will also confirm proper clearances between the punch and mating parts of the tablet press to eliminate tool binding and premature wear. Most tooling suppliers will provide a detailed inspection report or a Certificate of Conformance to assure tablet

manufacturers that a specific set of tooling is within the specified tolerance and will produce consistent and quality tablets. The inspection area should be a controlled environment, well lit for visual inspection and equipped with properly calibrated inspection instruments and gauges.

The tooling inspection program should be divided into two sections, incoming inspection and in-process inspection.

The incoming inspection program is for new tools and confirms adherence of critical dimensions. Tools that are supplied with a detailed inspection report should be verified by checking a small percentage of tooling to qualify the suppliers inspection report. A confirmation of the checked dimensions should be recorded and maintained for future reference.

The in-process inspection procedures are recommended for determining wear subjected on critical dimensions responsible for tablet quality and press operation. A visual examination will disclose tableting deficiencies which are easily identified by excessive and premature wear and overall tooling condition. The most important dimension affecting tablet hardness, weight and thickness consistency is the WL of the punches. It is not critical to inspect the WL for a calculated dimension, but to inspect for consistency within the set. During the inspection process it is good practice to determine if the punches and dies are in need of polishing and/or light reworking.

The punch tip is also critical for inspection and examination. Unfortunately, the worn punch tip is difficult or nearly impossible to inspect using traditional measuring instruments such as a micrometer or an indicator. The punch tip wears at the edge of the cup and can only be measured accurately using an optical comparator. Dies should be visually checked for wear rings in the compression zone, and replaced if worn. The severity of a die wear ring can be checked with an expanding indicator. The expanding indicator will not provide the actual die size, only the depth of the wear ring. The expanding indicator is also capable of measuring the amount and depth of the die taper.

The results of the WL inspection should be documented as well as noting tool wear and polishing or reworking if performed. When tooling wear exceeds the new tool specification, it is not generally considered unusable or out of new punch specification.

Reworking

If considerable reconditioning of the punches and dies is necessary they should be returned to the manufacture for evaluation. Extensive reworking of the tooling should be performed only by skilled personnel to assure conformance to strict tolerances providing tablet consistency and proper press operation.

Polishing the cup is the most common procedure of punch reworking performed by the tablet manufacturer and is easily achieved with proper training. Excessive polishing can reduce the cup depth and diminish the height of the embossing, thus reducing legibility and the ability to film coat. There are three common procedures of polishing the cup, (*i*) large soft cotton wheel fitted to a bench grinder motor, (*ii*) small nylon brushes or hard cotton bobs and polishing paste using a dremmel tool, and (*iii*) a process called drag finishing which drags the punch through walnut shells infused with polishing compound. The most effective of the methods is using the large cotton wheel. Polishing the cup with a large soft cotton wheel is the only method that polishes the cup and restores the critical land at the same time. Restoring the land can increase tool life, strengthen the punch tip and reduce the likelihood of capping and laminating. Polishing the punch cups with nylon brushes or using a drag finisher is the simplest method of polishing but does not restore

the tip edge or the land to eliminate hooked edge commonly referred to as a "J hook" that is common to capping and laminating. It is not advised to polish or restore the head flat; as this can alter the critical WL resulting in inconstant tablet hardness, thickness, and weight.

Troubleshooting

Learning to troubleshoot tableting problems is necessary to operate an efficient tableting program. Understanding the cycle of the press and the normal tooling wear associated with each cycle will greatly enhance the ability to identify deficiencies. Knowledge of the granulation and how it acts and reacts during compression is equally important. Tables 6 and 7 provide a useful troubleshooting guide for tooling and tablets.

Press Wear

Tablet press wear can sometimes be the reason for tooling failure and is often overlooked. As the tolerances of punches and dies are constantly monitored, so should the critical tolerances of a tablet press. For example, if tablet overall thickness is inconsistent the WL of the punches should be checked first; in most cases this dimension is the easiest to check. If the WL of the punches is acceptable, the tools are usually put back into service to frequently experience a reoccurrence of the initial problem. If the pressure roller is out of round, out of concentricity, or worn with severe pitting or flat spots, the result will be inconsistent tablet thickness as would be expected with improper punch WLs. Tables 6 and 7 show some of the critical press areas that should be monitored and how the wear affects the tooling and tablet production.

Figure 30 shows the correct way to check the turret guide for wear. A new turret may have an approximately 0.003 in. tip deflection. A turret guide considered worn has a tip deflection of 0.012–0.014 in. and should be sleeved or replaced.

Problems in tableting often have a domino effect. It is important to identify and remedy a problem before it affects other areas of the press, the tooling and tablet quality.

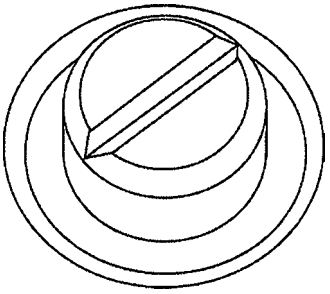
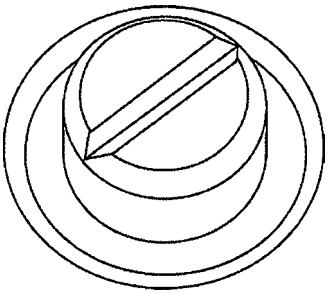
Purchasing Tablet Compression Tooling

To expedite a tooling order, it is important to provide the tooling supplier with the following details:

The size, shape, and cup depth of the tablet to be compressed (a sample tablet or sample tools would be sufficient if the information is not readily available).

1. Drawing number of the tablet if a drawing exists, if not, request a drawing for future reference.
2. Hob number, if the order is a replacement.
3. Press type, model number, and number of stations required.
4. Steel type if other than standard.
5. Historical data concerning capping, sticking, picking, high-ejection forces, etc.
6. If the tablet requires film or sugar coating.
7. Special options such as tapered dies, domed heads, key type, etc.
8. Special shipping instructions.

TABLE 6 Production Problems with Tablet Quality

Tablet problem	Possible cause(s)/corrective action(s)
A. Nonuniform tablet weight 250.00 mg 	1. Erratic punch flight *Check for/action* a. Free movement of punch barrels in guides (Guides must be clean and well lubricated) b. Excessive press vibration c. Worn or loose weight-adjustment ramp d. Proper operation of lower-punch control devices e. Limit cam on weight-adjustment head missing, worn, or incorrectly fitted f. Check dust seals g. Check that antiturning device is set correctly h. Reduce press speed 2. Granulation lost or gained after proper filling of die *Check for/action* a. Tail over die missing or not lying flat on die table b. Recirculation band leaking c. Excessive vacuum pressure, or nozzle improperly located 3. Feeders starved or choked *Check for/action* a. Incorrect setting of hopper spout adjustment b. Granulation bridging in hopper c. Wrong fill cam in use d. Excessive recirculation of granulation
243.75 mg 	4. Dies not filling *Check for/action* a. Excessive press speed b. See A3 and A5 c. Check speed or shape of feeder paddle 5. Lower punch pulled down before die is filled *Check for/action* a. Inadequate recirculation of granulation b. Recirculation scraper missing or bent

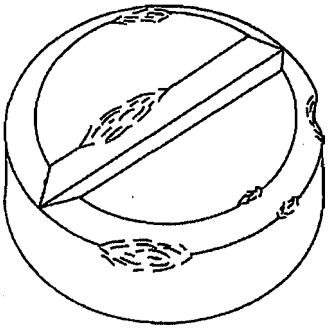
(*Continued*)

TABLE 6 Production Problems with Tablet Quality (*Continued*)

Tablet problem	Possible cause(s)/corrective action(s)
	6. Poor scrape-off of granulation *Check for/action* a. Scraper blade bent, worn, or not lying flat; bad spring action
	7. Nonuniform punch length *Check for/action* a. Check that working length is within ±.001 inch [.025 millimeter] of TSM specification
	8. Projection of die(s) above die table *Check for/action* a. Clean die pocket or check die dimension
	9. Automatic weight-control system not working correctly *Check for/action* a. Check that system's settings and operation are correct; see manufacturer's handbook
	10. Wide variation in thickness of lower punch heads *Check for/action* a. Check that head thickness of lower punches is within ±.010 inch [.025 millimeter] of TSM specification
B. Nonuniform tablet thickness (Not pictured)	1. Nonuniform tablet weight *Check for/action* a. See A
	2. Bouncing of pressure rollers *Check for/action* a. Improper setting for overload release b. Press operating near maximum density point of granulation; increase thickness and/or reduce weight within allowable tablet tolerances c. Pressure rollers not moving freely; punch faces in poor condition d. Air trapped in hydraulic overload system e. Worn pivot pins on roller carriers

(*Continued*)

TABLE 6 Production Problems with Tablet Quality (*Continued*)

Tablet problem	Possible cause(s)/corrective action(s)
	3. Nonuniform punch lengths *Check for/action* a. Check that working length is within ±.001 inch [.025 millimeter] of TSM specification
C. Nonuniform tablet density (friability) 	1. Nonuniform tablet weight and thickness *Check for/action* a. See A and B b. See capping in G
	2. Unequal distribution of granulation in die bores *Check for/action* a. Stratification or separation of granulation in hopper b. Excessive recirculation (This causes classification of granulation because only finer mesh material escapes the rotary feeders.)
	3. Particle segregation or stratification in hopper *Check for/action* a. Reduce variations in particle size; reduce machine vibration; reduce machine speed
	4. Low moisture content *Check for/action* a. Add moisture to aid bonding
D. Excessive vibration of press (Not pictured)	1. Worn drive belt *Check for/action* a. Inspect drive belt
	2. Mismatched punch lengths *Check for/action* a. See A-7
	3. Press operating near maximum density point of granulation *Check for/action* a. Increase tablet thickness and/or reduce its weight within allowable tablet tolerances

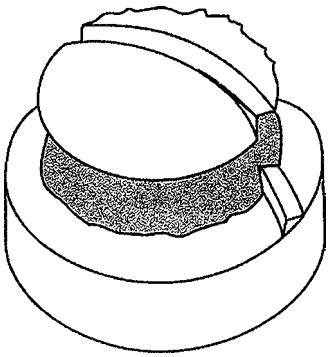
(*Continued*)

TABLE 6 Production Problems with Tablet Quality (*Continued*)

Tablet problem	Possible cause(s)/corrective action(s)
	4. High ejection pressure *Check for/action* a. Worn ejection cam b. Add more lubrication to granulation, or taper dies c. Barrel-shaped die bores
	5. Improper pressure-release setting *Check for/action* a. Increase pressure to the tooling's limit
E. Dirt in product (black specks) (Not pictured)	1. Dust, dirt, or press lubrication in the granulation *Check for/action* a. Clean press more frequently b. Excessive or wrong press lubrication c. Use proper punch dust cups and keyway fillers d. Rubbing of feeder components e. Punch-to-die binding
F. Excessive loss of granulation (Not pictured)	1. Incorrect fit of feeder to die table *Check for/action* a. Feeder base set incorrectly (i.e, too high or not level) b. Bottom of feeder pans worn due to previous incorrect settings; relap pans, if necessary
	2. Incorrect action of recirculation band *Check for/action* a. Gaps between band's bottom edge and die table b. Binding in mounting screw c. Inadequate pressure on hold-down spring
	3. Insufficient scraping of die table *Check for/action* a. Worn or binding scraper blade b. Outboard scraper edge allowing granulation to escape
	4. Granulation lost from die prior to upper punch entry *Check for/action*

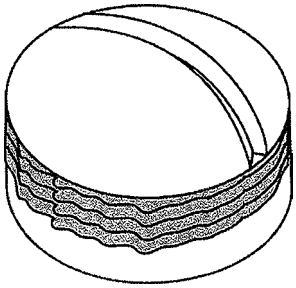
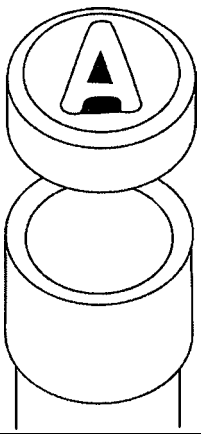
(*Continued*)

TABLE 6 Production Problems with Tablet Quality (*Continued*)

Tablet problem	Possible cause(s)/corrective action(s)
	a. Tail over die not lying flat on table
	5. Granulation lost at compression point *Check for/action* a. Compression occurring too high in the die b. Excessive suction or misdirected exhaust nozzle
	6. Excessive sifting *Check for/action* a. Excessive clearance between lower punch tip and die bore b. Excessive fine particles in the granulation c. Tapered dies installed upside down
G. Capping and lamination 	1. Air entrapment *Check for/action* a. Compress granulation higher in the die b. Reduce press speed c. Precompress granulation d. Reduce quantity of fine particles in the granulation e. Reduce cup depth on punches f. Taper dies g. Ensure that punch-to-die clearance is correct
	2. Excessive pressure *Check for/action* a. Reduce tablet weight and/or increase its thickness within allowable tolerances b. Adjust pressure
	3. Ringed or barrel-shaped die bore *Check for/action* a. Reverse dies b. Hone or lap bores c. Compress granulation higher in the die
	4. Too rapid expansion of tablet upon ejection *Check for/action* a. Taper dies
	5. Weak granulation *Check for/action*

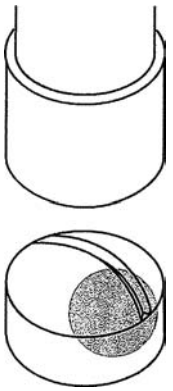
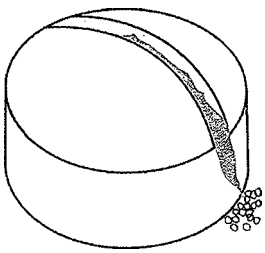
(*Continued*)

TABLE 6 Production Problems with Tablet Quality (*Continued*)

Tablet problem	Possible cause(s)/corrective action(s)
	a. Increase quantity of binder; use stronger binder
	6. Excessively dry granulation *Check for/action* a. Increase level of lubricant
	7. Excessive lubrication of granulation *Check for/action* a. Decrease level of lubricant; blend all ingredients fully before adding lubricant
	8. Punch cavity too deep *Check for/action* a. Use punches with less concave depth
	9. Punch tips worn or burred *Check for/action* a. Refurbish or replace punches
	10. Lower punch set too low at tablet take-off (Reworking or refurbishing punches can cause this.) *Check for/action* a. Set lower punch tip flush with top of die
	11. Tablet take-off bar set too high *Check for/action* a. Adjust take-off bar
H. Picking and sticking	1. Excessive moisture *Check for/action* a. Check moisture content of granulation b. Check room humidity
	2. Punch face condition *Check for/action* a. Pits on punch faces and/or improper draft on embossing; try repolishing punch faces b. Try chrome-plating punch faces
	3. Insufficient compaction force *Check for/action*

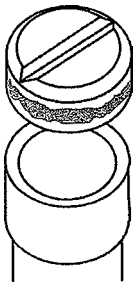
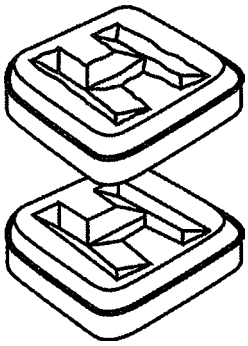
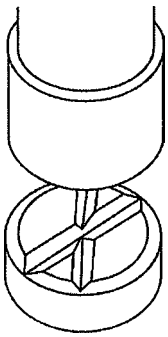
(*Continued*)

TABLE 6 Production Problems with Tablet Quality (*Continued*)

Tablet problem	Possible cause(s)/corrective action(s)
	a. Increase tablet weight and/or reduce its thickness within allowable tolerances
	4. Inadequate lubrication of granulation *Check for/action* a. Check and/or adjust level of lubricant used
	5. Poor embossing design *Check for/action* a. Redesign embossing per TSM guidelines, or consult tooling supplier
I. Mottled or marked tablets 	1. Contamination of granulation, usually by grease or oil *Check for/action* a. Check oil seals on upper punch guides b. Reduce lubrication of upper punches to an acceptable level c. Fit oil/dust cups to upper punches
	2. Contamination of granulation from chutes, feed hoppers, take-off bar, or rubbing together of feed paddles *Check for/action* a. Clean and reset components correctly
	3. High moisture content of granulation *Check for/action* a. Re-dry granulation
	4. Oversized granulation particles *Check for/action* a. Reduce particle size
J. Chipping or splitting 	1. Poor surface finish on punch tips; worn punches and dies *Check for/action* a. Polish punch tips; replace punches and dies
	2. Poor tooling design (e.g., sharp embossing or bisect lines) *Check for/action* a. Polish punch tips; replace punches and dies

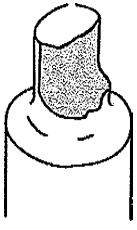
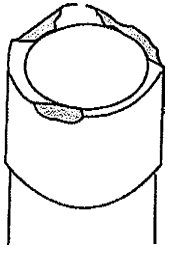
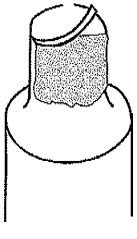
(*Continued*)

TABLE 6 Production Problems with Tablet Quality (*Continued*)

Tablet problem	Possible cause(s)/corrective action(s)
K. Splitting of layered tablet 	1. Excessive pressure *Check for*/*action* a. Decrease pressure 2. Excessive lubrication of granulation *Check for*/*action* a. Reduce amount of lubricant
L. Indistinct breakline or embossing 	1. Incorrect embossing design *Check for*/*action* a. Redesign embossing per TSM guidelines, or consult tooling supplier 2. Worn punch tips *Check for*/*action* a. Replace punches 3. Excessively coarse granulation *Check for*/*action* a. Reduce particle size 4. Inadequate binder *Check for*/*action* a. Increase binder strength 5. Picking *Check for*/*action* a. Compress granulation at a lower pressure
M. Double impression of embossing 	1. Rotation of punches *Check for*/*action* a. Adjust antiturning device b. Use keyed punches

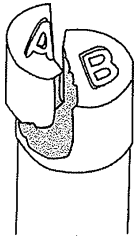
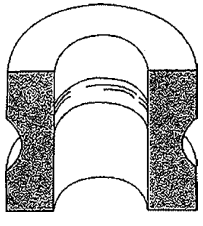
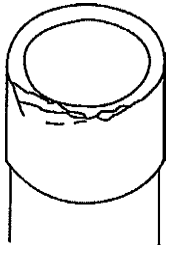
Note: Table reprinted with permission from *Pharmaceutical Dosage Forms*. Vol. 2. 2nd ed. New York: Marcel Dekker, Inc.; 1989: 603–607.

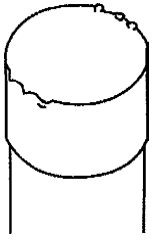
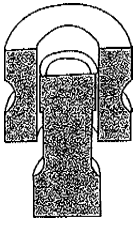
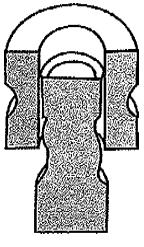
TABLE 7 Production Problems with Tooling

	Tooling problem	Cause(s)	Corrective action(S)	Comments
	(1) *The tip has cracked across the face of the concave and then broken away.*	1. Excessive hardness for application. Excessive pressure	one: discard tool; consult tooling manufacturer.	Tools should always be run at the minimum pressure required to achieve a satisfactory tablet.
	(2) *The tip has cracked and broken away along the angle between the bevel and tip face.*	2. See cause for 1.	See action for 1.	A crack will always follow the line of least resistance, which may be the sharp angle between the punch face and the embossing.
	(3) *The tip has cracked and broken away along the angle between a breakline and a concave tip face.*	3. Excessive hardness. Areas of concentrated stress near breakline or on embossing (i.e., abrupt change of surface contour). Excessive pressure.	See action for 1.	See comments for 2.

(*Continued*)

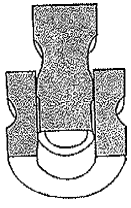
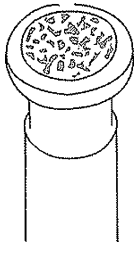
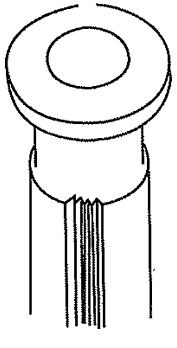
TABLE 7 Production Problems with Tooling (*Continued*)

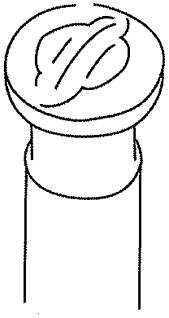
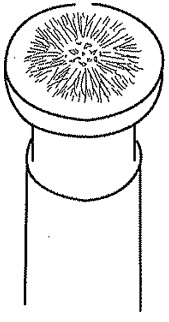
	Tooling problem	Cause(s)	Corrective action(S)	Comments
	(4) *The tip has cracked and broken away along the embossed lettering.*	4. See cause for 3.	See action for 1.	See comments for 2.
	(5) *This die show a typical wear pattern in the bore.*	5. Normal die wear caused by continuous pressure at the compression area in the bore.	Examine dies with magnifying glass and monitor tablet ejection. When possible, compress tablets in different areas of the die to spread wear, and reverse the die when one end is worn. Check that correct steel was chosen. If wear is a serious problem, consult tooling manufacturer.	If allowed to go too far, the die wear can lead to ejection problems and other problems associated with punch tightness. If a known abrasive granulation is to be compressed, the tooling manufacturer can possibly offer a more wear-resistant material for tooling.
	(6) *The edge of the tip has been damaged outside the press.*	6. Mishandling of punch (punch has collided with or been dropped onto a hard surface). Accidental damage occurred during fitting of punches to the press.	Carefully remove damage by blending and polishing. Exercise extreme care when handling tools; the tips are very fragile. Train personnel to handle tools properly.	Careful examination of this type of damage will reveal clues to its cause, (a) If the damage has caused the tip to spread beyond its diameter, the damage most likely occurred out of the press, (b) The texture of

				the surface causing the damage will be transferred to the damaged part.
	(7) *The punches have met in the press; damage occurred where the opposing punch has a breakline.*	7. Contact between upper and lower punches in the press.	Carefully remove dents by blending and polishing. Do not run the press without granulation at setup; manually turn over the dies until all are filled with granulation.	In some presses, if tools are run or even turned without granulation, the punches can meet, causing damage.
	(8) *Again, the punches have met in the press, but the opposing punch has no breakline.*	8. See cause for 7.	See action for 7.	See comments for 7.
	(9) *Pressure has started to spread the punch tip; working length may not yet be affected. The spreading will probably occur on both upper and lower punches.*	9. Excessive pressure (first stage for upper and lower punch).	In the early stages before working length is affected, punch damage can be removed by blending or polishing. Check all punch lengths before reusing the set; other punches may have been damaged.	This type of damage can be checked by measuring the tip diameter at the extreme edge and at the tower end. If these dimensions vary, damage has occurred.
	(10) *Lower punch is over-pressured to the point where the stem is distorted and the working length is reduced.*	10. Excessive pressure (final stage for lower punch).	None: the final stage of over-pressure cannot be rectified; the punch is permanently distorted.	Rolling the punch barrel on a flat surface is a simple way to check for this type of damage: the punch tip will be seen to rotate out of true.

(*Continued*)

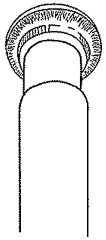
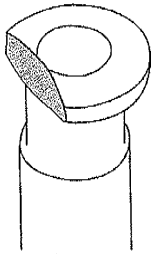
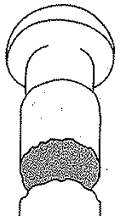
TABLE 7 Production Problems with Tooling (*Continued*)

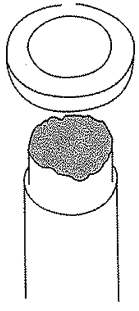
	Tooling problem	Cause(s)	Corrective action(S)	Comments
	(11) *Excessive pressure will have the same effect on the upper punch as on the lower; see (10).*	11. Excessive pressure (final stage for upper punch).	See action for 10.	See comments for 10.
	(12) *The head flat has worn to the point where fragments of metal are being removed from the punch head.*	12. Excessive pressure and damaged or worn pressure roller. Foreign matter between pressure roller and punch head.	Reduce pressure; replace lubricant; repair pressure roller. Spalling of the head deposits metal particles in the press: clean press throughout. Consult tooling manufacturer.	If not tackled early, this type of damage can lead to serious wear and damage to the tools and the press.
	(13) *Scoring of the punch barrel is caused by a lack of lubrication and/or the presence of foreign matter in the punch guides.*	13. Tightness of the punch barrel in the turret leading to possible scoring and pick up of metal, which leads to increased tightness. Poor lubrication.	If possible, polish punch to restore original condition. Check that guides are clear of granulation and metal particles. Pay particular attention to the punch sockets in the turret. Check working length before reworking	Many tooling problems are caused by tightness; marking of the barrel is a definite indication of trouble. If the lubrication becomes contaminated with the granulation, its lubricating properties are

			punch. Ensure that the lubrication system is clean, correct, and operative.	destroyed and excessive wear occurs.
	(14) *The punch is not rotating, and the pressure roller may be running tight, causing wearing of the head in only one spot, (Shaped punches do not rotate.)*	14. Excessive pressure. Lack of lubrication. Tight punches or pressure rollers.	Check that head flat is not too small to achieve satisfactory dwell time during compression. Check underside of head for damage. If warranted, polish head. Resolve pressure problem; ensure that punch and pressure roller can move freely; ensure adequate lubrication.	Press damage is possible.
	(15) *The ejection cam is causing wear on the lower punch head.*	15. A rotating punch is running very tight on ejection, causing a radial pattern of wear. Insufficient head flat. Excessive pressure. Damaged, bruised, or scored compression roller.	Polish head or increase size of head flat. Ensure that punches can operate freely at all times. Resolve ejection problem; to ease ejection loads, taper dies. Always use minimum pressure needed to compress tablets. Ensure that surface of compression roller is clean and free of burrs or bruising. Check cam for excessive wear; clean and remove any metallic particles from the cam track and pressure rollers.	If the head flat is too small, the compression force is concentrated on a small area and ultimately will cause the center of the head to fail. Tooling is subjected to continuous high pressure and eventually the structure of the steel will break down. If punches are tight, unnecessary pressure is applied to tooling, cams, and compression rollers. If not corrected, damage to punch heads or compression rollers will

(*Continued*)

TABLE 7 Production Problems with Tooling (*Continued*)

	Tooling problem	Cause(s)	Corrective action(S)	Comments
				transfer rapidly to all the punches in the press.
	(16) *Tight punches have caused excessive wear to the inside head angle, (Damage to press cams is likely.)*	16. Punch has become tight in the die or press turret due to lack of lubrication. Incorrect cam angle on punch heads. Bruised or scored press cams.	None: discard the punch. Determine cause and ensure that replacement punch moves freely (i.e., punch should fall freely under its own weight when antiturning device is loosened). Clean the press to remove metal particles. Ensure that punch guides are clean and correct lubrication is applied. Check that cam angle is compatible with the press cams. Inspect cams for bruises and scores; if needed, repolish or replace cams.	The top of the punch head may also be damaged. This kind of damage leaves metal particles in the press.
	(17) *This damage is similar to (16), but the punch was not allowed to rotate, resulting in part of the head breaking off.*	17. This problem is similar to 16, but the punch is not rotating due to the use of a keyed punch or tightening in the turret.	None: discard the punch. Determine cause of problem, and ensure that replacement punch is loose (i.e., punch should fall freely under its own weight when the antiturning device is loosened). Clean the press to remove metal particles.	See comments for 16.
	(18) *The punch barrel has snapped in the press.*	18. Upper punch is possibly being prevented from entering the die due to tip breakage (see 1, 2, 3 or 4); the head then strikes part of	Discard tool; monitor condition of tooling at all times to avoid tightness and excessive pressure.	With unenclosed presses, the broken part may be ejected from the press with considerable force,

		the punch guide system and breaks the barrel. Excessive tightness.		endangering personnel and equipment.
	(19) *The punch snapped in the press, but this time the head has broken off.*	19. Due to wear and refurbishing, head flat has become larger than the neck diameter. When compression force is applied, the punch is unsupported at the neck and breakage results.	None: discard tool and monitor the condition of tools in use, especially after refurbishing. Ensure that all metal fragments are removed from the press.	Severe damage to the press is almost certain.
	(20) *Burrs are present inside the punch tip (clawing). (Not pictured)*	20. Misalignment of punch tips in die bore. Worn punch guides or die sockets. Eccentricity of punch tips to punch body. Extrusion of product between punch tips and die bores. Excessive feather edge on punch tips, especially deep concave cups.	Ensure that internal chamfer of die bores is sufficient. Check for wear and rectify; check concentricity of punch tips. Ensure that tip-to-die bore clearance is correct. Increase land or flat on tip edge; ensure that land is blended.	
	(21) *The surface finish of the punch face is deteriorated (i.e., pitted or discolored). (Not pictured)*	21. Compression of an abrasive or corrosive granulation.	Ensure that the correct steel has been chosen. Check for sufficient lubrication of the granulation.	

Source: Reprinted with permission from *Tooling Problems,* Holland Educational Series, No. 4. Nottingham, England: I Holland Limited; 1988.

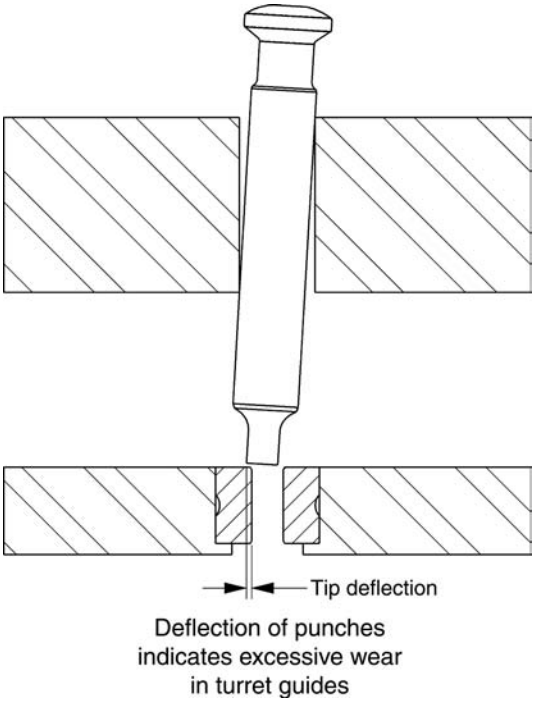


FIGURE 30 Checking turret guide wear.

If the tablet will be a new design or new shape then a sketch or a reference to an existing product and tablet weight should be submitted. From this information the tooling supplier will generate a tablet drawing for further approval. After the drawing is approved, the tablet manufacturer has the option to request a placebo tablet or a sample of the punch tip for further review and approval, there is normally a fee for this service. After approval of the sample punch tip or placebo tablet, the process of tool manufacturing will begin.

CONCLUSION

Choosing the current options for a tableting operation is normally accompanies by trail and error, therefore accurate record keeping is essential. It is recommended to utilize all available industry resources such as tablet press and tooling manufacturers for assistance with these choices. Chances are they have resolved similar difficulties for other customers and have the expertise to recommend the correct options for most tableting operations.

Tablet press and tooling manuals should be located for easy access to the press setup, compression, and tooling personnel. The three basic rules of tableting are:

1. Keep compression forces as low as possible.
2. Clean and lubricate the press and tooling properly.
3. Keep punches and dies in good condition.

This along with strong communications will result in an efficient tableting operation, producing high-quality tablets.

2
Tablet Press Instrumentation in the Research and Development Environment

Gary E. Bubb
Specialty Measurements Inc., Lebanon, New Jersey, U.S.A.

> *If you can measure that of which you speak and express it in numbers, you know something about your subject; but if your cannot measure it, your knowledge is of a very meager and unsatisfactory kind.*
>
> William Thomson (Lord Kelvin) (1824–1907)

INTRODUCTION

When asked to write a chapter on tablet press instrumentation, the challenge was not what to write, but rather, how much should be left out. Covering the topic in sufficient detail as to provide a roadmap on how to properly instrument a tablet press including the design of the sensors, electronics and analysis software would require an entire volume, not just a chapter. On the other hand, it is desirable that the reader have a sufficient knowledge of the topic to be an educated consumer. The objective of this chapter, therefore, is to give the reader an appreciation of what is involved in the makeup of a data acquisition system and what is important to fulfill their requirements.

Tablet press instrumentation discussed in this chapter will be limited to that of force and displacement. Other parameters, such as vibration, noise, and temperature can be meaningful, but are not commonly used in the research and development arena. The same is true for the measurement of punch pull up and pull down forces and tablet press control systems.

This chapter will deal with current practices of instrumentation and not offer any significant historical perspective unless it has a bearing on today.

OVERVIEW OF A DATA ACQUISITION SYSTEM

Although there are many components that make up an instrumentation system they will be grouped into six major categories for the purpose of this discussion. Though calibration is technically not a component of the system, its importance is so significant that it has been included.

1. Sensor types:
   a. Piezoelectric
   b. Strain gauge:

   i. Wheatstone Bridge
   ii. Temperature compensation
   iii. Bridge balance
   c. Displacement
2. Signal conditioning:
   a. Power supply
   b. Differential amplifier
3. Analog to digital conversion:
   a. Resolution
   b. Aliasing filters
4. Representative tablet press sensors for compression, ejection and take off
5. Calibration:
   a. Precision; accuracy; and repeatability
6. Analysis software

Sensor Definition

In the broad sense, a sensor or transducer is a device that transforms one type of energy into another. By this definition, a battery is a transducer (the conversion of chemical energy into electrical). Narrowing the definition to a specific class of transducers, electromechanical, a transducer is a device that converts a physical parameter into an electrical signal that can be measured and or recorded.

Examples of a sensor or transducer are given in the following chart:

Force		Electrical Signal
Pressure		Voltage
Torque	⇨	Current
Acceleration		Pulses
Displacement		
Temperature		

DISCUSSION OF SENSORS FOR FORCE MEASUREMENTS ON A TABLET PRESS

There are two generic types of sensors that have been used for the measurement of compression and ejection forces, piezoelectric and strain gauge-based. Piezoelectric were the early favorite because of their small size, large self-generating output and high frequency response. A drawback to this type of sensor is the low frequency response allowing its use only in dynamic events. Signal changes as a result of cable movement and contamination within connectors are also problematic. These could be overcome by carefully routing and anchoring cables, but the low frequency response presents a challenge for calibration. Typically, calibrations are performed by gradually applying a force, holding it for several seconds to allow the signal to decay to zero, and then rapidly removing the force. This procedure actually performs a negative force calibration relying on the belief that a positive and negative calibration were equivalent.

The strain gauge-based transducer offers the advantage of a static or DC response. That is to say an applied force will continue to be displayed properly independent of the application time. A piezoelectric sensor will "bleed down" to a zero reading in some seconds, even if the force is still being applied. Additionally, a well-designed stain gauge-based transducer is an order of magnitude more accurate. For these reasons, the strain gauge-based transducer has dominated the measurement of forces in the pharmaceutical industry.

Piezoelectric Load Cells

Piezoelectric force transducers are generally constructed of quartz or piezoceramic elements. The quartz crystal is cut in a precise orientation to the crystal axes depending on the application and design of the transducer. The crystal produces an electrical output when experiencing a change in load. The general belief is that they cannot be used for static measurements, their use being limited to dynamic events only. However, this is a misconception. Quartz transducers, paired with appropriate signal conditioners can offer excellent quasi-static measuring capability (1,2).

Anyone wishing to utilize a piezoelectric force transducer should contact the manufacturer of the device for directions. Mounting is extremely important as off center loading can cause great errors. Time constants must be considered. If the load application is slow the peak value will be understated and the return to zero will overshoot the baseline. The signal conditioning must match the sensor impedance (see below) and should be tailored to the application. Used properly, piezoelectric force transducers are rugged, accurate devices that are small in size and generally easy to install.

There are two basic types of piezoelectric force transducers, low impedance and high impedance.

- *High impedance*. The piezoelectric effect was first discovered by Pierre and Jacques Curie in 1880. When the element was distorted a current was produced. In order to relate the current to the deformation a special amplifier is required; a charge amplifier. This system offers the user the most flexibility. Time constants can be made longer allowing easy short-term static calibration. Because they contain no built-in electronics, they have a wider operating temperature range. They do come with some significant disadvantages, however. Because of the high impedance, any changes in the resistance or capacitance of the connections between the quartz element and the charge amplifier will likely cause a false signal. Special impedance cables must be used and all connectors need to be free on any contamination. Even the oil from ones fingers is sufficient to cause problems.
- *Low impedance*. Transducers of this type are the same in their construction with the addition of a built in amplifier. This will increase the size of the transducer and limit the temperature range because of the internal electronics, but will eliminate the concerns with cable movement and connector contamination. Low impedance transducers can be used with general purpose cables in environments where high humidity/contamination could be detrimental to the high insulation resistance required for high impedance transducers. In addition, longer cable lengths, between transducer and signal conditioner and compatibility with a wide range of signal display devices are further advantages of low impedance transducers.

Strain Gauge

The strain gauge is the basic element in the construction of a strain gauge load cell or transducer. There is a common misconception that a quality strain gauge load cell is merely installing four strain gauges into a Wheatstone bridge and performing a calibration. This is far from the truth. A proper load cell consists of a designed spring element, proper installation of strain gauges onto the mechanical spring element, temperature compensation for no load and full load conditions along with a calibration performed after installation into the machine.

Strain gauge-based load cell are used by the NIST as primary standards for force measurements because of their accuracy, repeatability, and robustness. With today's

technology, the life expectancy of strain gauge-based load cell should approach 25–50 years depending on the environment.

There have been many in-house designed instrumentation systems that served the pharmaceutical industry well in the past, some better than others. Because the strain gauge-based load cells are by far the dominant sensor on modern tablet presses, and because the quality of the installations varies widely, there will be a significant discussion on this area.

Strain, the Definition:

There are two definitions of strain, true strain and engineering strain. For all practical purposes in the design of load cells, they are identical as the deformations are so small (Fig. 1).

True Strain = Change in length divided by the current length.
Engineering Strain = Change in length divided by the original length.

When any item undergoes stress there is a resulting strain, the magnitude varies with the elastic modulus or Young's modulus of elasticity.

Picture the image on the left as a length of copper wire. When stretched, the wire becomes longer and smaller in diameter, both contribute to an increase in the resistance of the wire (Fig. 2).

Strain Gauges, the History

The exact discovery of the strain-induced resistance change of electrical wires is not clear; Lord Kelvin did report on the effect in the 1800s. The initial wire strain gauge utilized small holes drilled into the part under test at a given distance apart. Small posts were then inserted into the holes and a wire wrapped around the posts. As the part underwent strain, the resistance change of the wire was measured and correlated to the strain.

In 1944, Simmons was awarded a patent for a bondable wire strain gauge pressure transducer. During the same time period Ruge, an MIT professor was using the bonded

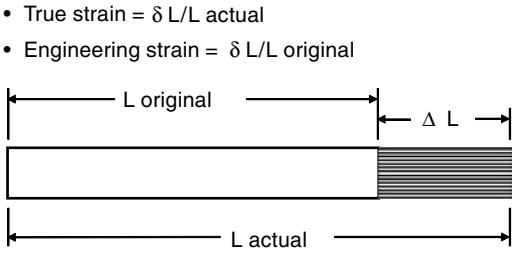


FIGURE 1 Definition of strain.

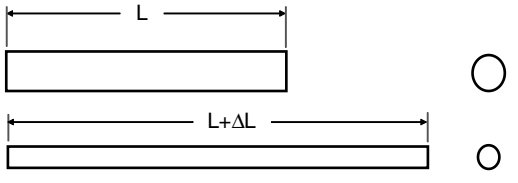


FIGURE 2 Strain and resistance change.

wire strain gauge for early force transducers. Simmons and Ruge are generally credited as co-inventors of the bonded wire strain gauge. Ruge is credited as being instrumental in advancing the applications of this emerging technology (3).

In the 1950s printed circuit technology gave birth to the bonded foil strain gauge. The foil quickly supplanted the wire with better heat dissipation, reduced creep, and much greater design flexibility. Today there are more than 20,000 different patterns using specialized alloys and shapes to assist the strain gauge transducer designer.

There are two other strain gauge types that deserve attention:

Sputtered or Deposited Metallic Strain Gauges

Metal films can be vaporized and sprayed onto an electrically insolated surface and used as strain gauges. By proper masking the desired strain gauge pattern can be deposited directly onto the surface. In this manner, multiple gauge patterns can be sprayed at once (3). There are several advantages to this approach; elimination of an organic adhesive and low cost high production rates. The disadvantage at this time is high set-up cost and generally lower performance than achievable with rolled alloy foils.

Semiconductor Strain Gauges

Semiconductor strain gauges are generally small silicon chips that have been preferentially cut on a specific silicon crystal axis. Depending on the cut direction the sensitivity can be up to 80 times higher than a typical foil gauge. The small size and high sensitivity make them ideal for miniature high output transducers.

The disadvantages are a high sensitivity to temperature, inability to dissipate heat produced from the excitation voltage and a reduced linearity, especially at higher strain levels. One of these negative factors can actually be turned into an advantage as designing a spring element for a lower strain means a stronger part or greater overload rating before structural failure would occur. This also makes for a stiffer component with a resulting higher frequency response. An overload will result in a permanent offset in the strain circuit, however, not likely to cause structural failure of the component part and possibility taking a machine out of service.

Semiconductor strain gauges are ideal for tablet press transducers, such as take-off, scrape off, knock off or whatever name you apply to the tablet being removed from the lower punch tip after ejection.

WHEATSTONE BRIDGE

The Wheatstone bridge is not the only strain gauge circuit available, but is certainly the most commonly accepted for use in industry. It is excellent for use with multiple gauge installations and measurements of both static and dynamic events.

The Wheatstone bridge was first described by Samuel Hunter Christie in 1833, but it was Sir Charles Wheatstone who found practical applications for the circuit that carries his name today. Wheatstone called the circuit a "Differential Resistance Measurer." This is still the best description today for this simple but elegant circuit.

In simple terms, and as applied to strain gauges, there are four closely matched resistors (strain gauges) arranged in the following geometry.

In Figure 3 +E is the positive excitation voltage to the circuit, − E is the negative excitation voltage to the circuit, + signal is the positive voltage output from the circuit, and − signal is the negative voltage output from the circuit.

Based on Figure 3 below and making the initial assumption that all four resistors, wire and wire connections are exactly the same resistance values within each arm or leg of the Wheatstone bridge; the voltage potential at the signal corners would be zero. The beauty of this simple circuit is that even with a large applied excitation voltage the differential voltage at the signal corners is still zero. Therefore, even very small signal changes can be amplified without bias from the excitation voltage. Amplifier gains in excess of 10,000 today show excellent linearity and frequency response making this circuit extremely sensitive to minute changes in resistor values.

Let us say that the resistors are strain gauges. As pointed out earlier a wire or foil under a positive strain (tension) will increase in length and decrease in diameter, resulting in an increase in resistance. A compressive force will decrease the wire length, increase the diameter, and lower the resistance. Let us assume for the moment that the strain gauge in arm 1 goes into tension resulting in an increase in resistance. The current in the circuit will always take the path of least resistance, therefore, more current will flow through arm 2 and less through arm 1, causing a higher voltage potential at the junction between arms 2 and 3 than the junction of arms 1 and 4. For that reason, the junction between arms 2 and 3 is called the positive signal for this arrangement. Following the same logic if the strain gauge in arm 2 went into compression, it would produce the same positive potential as arm 1 going into tension. The same discussion can be offered for arms 3 and 4.

The conclusion to all of this is that an increase in resistance of either arm 1 or 3 will cause a positive output in the circuit while a decrease in resistance in arms 2 and 4 will also cause a positive signal. For this reason, arms 1 and 3 are referred to as the positive arms while arms 2 and 4 are called the negative arms. The term bridge factor is an expression of the number of equivalent active arms in the circuit. For example, if only

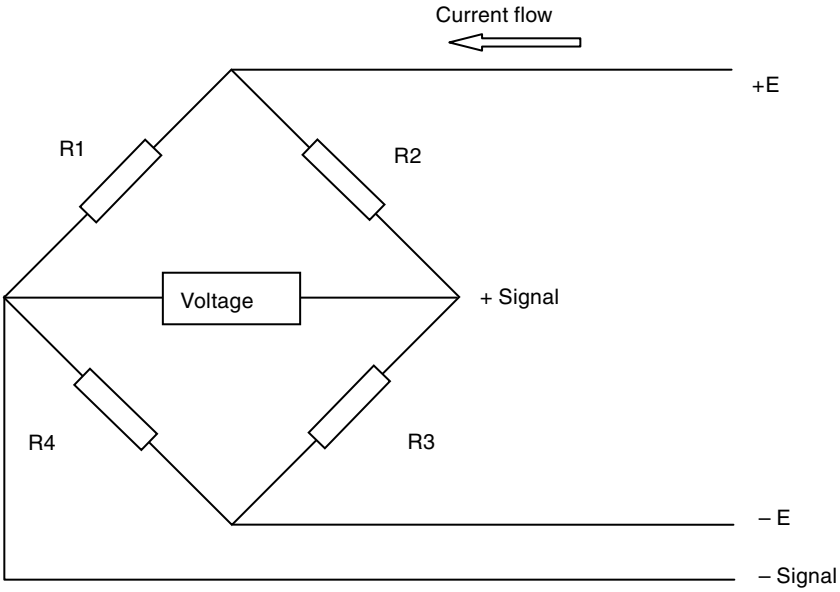


FIGURE 3 Wheatstone bridge. *Abbreviations*: +E, positive excitation voltage to the circuit; −E, negative excitation voltage to the circuit; + Signal, positive voltage output from the circuit; − Signal, negative voltage output from the circuit.

arm 1 contained a strain gauge that actually saw a strain the bridge factor would be 1. If the strain gauges in arms 1 and 3 saw tension and the strain gauges in arms 2 and 4 saw an equal amount of compression, the bridge factor would be 4.

STRAIN GAUGE TRANSDUCER CONCEPTS

The well designed transducer needs to be linear with minimal hysterias, sensitive, exhibit good thermal stability, and have a good return to zero under a no load condition. Additionally, the *transducer should only respond to the force to be measured* and not to any other force or physical parameter. The choice of materials to manufacture the transducer from will be a consideration as well as the design of the spring element, the area where the strain gauges will be attached. If the physical design of the transducer is not well thought out, the sensor will not perform as hoped. The following simple examples are shown to demonstrate the principle, not an actual design concept.

Cantilever Beam

The two gauges on the top will experience tension as the beam is deflected, therefore, one gauge should be installed in arm 1; the other in arm 3 of the Wheatstone bridge (Fig. 4). Provided that the other two arms contained only resistors and not strain gauges the bridge factor would be 2.0. However, if two additional strain gauges were installed on top surface perpendicular to the other two, they would see only Poisson's ratio of the full strain, or 0.3. Therefore, the bridge factor would be $1 + 0.3 + 1 + 0.3$ or 2.6. Now if the two strain gauges on the bottom that see compression were installed in arms 2 and 4, the bridge factor would be 4. In order to make a proper transducer, the length and thickness of the beam would be designed to provide the desired stress and resulting strain for the material the beam is made of.

There are hundreds of unique transducer concepts that have been utilized for force applications. The roll pin concept for compression force was introduced into the pharmaceutical industry in the early 1980s (4). Prior to that time compression forces on a rotary tablet press were measured with strain gauges installed on structural tie rods or eye bolts. Wheatstone bridges were applied but no additional consideration was given to the spring element design or temperature compensation. To this day many transducers manufactured for the Pharmaceutical Industry are not properly temperature compensated. The load cell roll pin is a good example of a proper design (Fig. 5). The sensor is

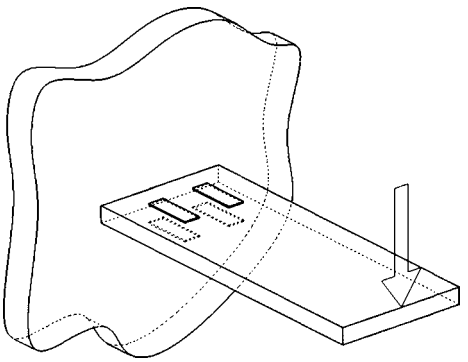


FIGURE 4 Cantilever beam.

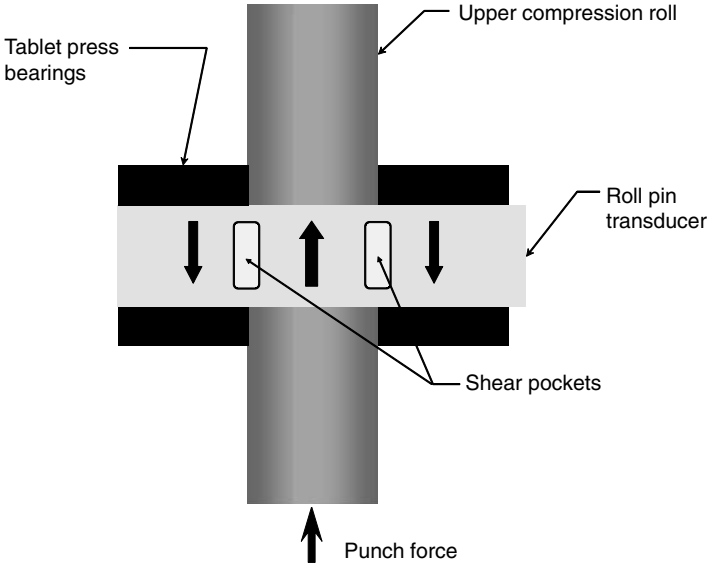


FIGURE 5 Roll pin shear load cell.

physically close to the force to be measured, the action line of the force is coincident with the load cell, the bridge factor is 4, and it can easily be temperature compensated.

Roll Pin Shear Load Cell

The roll pin load cell replaces the existing roll pin in this application while keeping all of the original functionality, including lubrication. Shown above is a representation of an upper roll load cell. The upper punch is exerting a force on the compression wheel that is being transferred to the center of the roll pin. The pin then transfers the force through the shear pockets to the ends of the pin and finally into the structural support of the machine. In this instance, a compression force is converted into a shear force for the purpose of making a transducer. The shear pocket geometry is conceived to produce the desired sensitivity for the anticipated forces (Fig. 6).

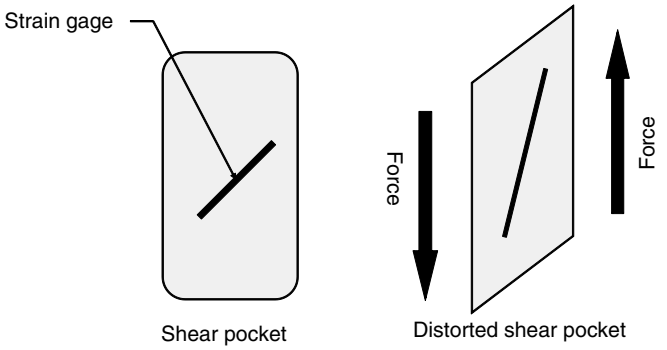


FIGURE 6 Strain in roll pin transducer.

The shear pocket on the left is not under load. The shear pocket on the right is an exaggerated picture of how the real distortion would look. With the strain gauge mounted at a 45° angle the strain is positive in this pocket. By carefully choosing the correct strain gauge orientation for each of the four pockets a bridge factor of 4 is obtained and the roll pin responds only to the desired force. One must be careful here as there are three possibilities on how the gauges are positioned and only one is correct.

1. The load pin reacts only to the compression force.
2. The load pin reacts only to the torque in the pin from the compression wheel turning.
3. The load pin reacts to both the torque and compression force.

Number three is the most insidious as it will not show up during a calibration with only an axial load applied, however, will yield incorrect information during operation due to the tensional component. A check is to try to rotate the compression quickly without applying an upward force and see if the load cell produces any output (Figs. 7 and 8). Remember that the torsion affect will be much greater under a compressive force so any output observed no matter how small is a good indication of an improperly installed or wired set of strain gauges.

Temperature Compensation

The basic strain gauge and Wheatstone bridge circuit is generally adequate for low-accuracy do it yourself transducers. These types of systems have, in fact, served the pharmaceutical industry very well over the past several decades and much benefit has come from these homegrown systems. Even today, some companies promoting themselves as experts are in reality offering transducers only at this quality. This level of thermal compensation, however, is not nearly adequate for a large class of commercial transducers available over the last 20 years.

There are two thermal considerations to account for:

1. Zero shift with change in temperature.
2. Span or sensitivity change with change in temperature.

Zero Shift

There are four orders of temperature compensation for zero shifts that can be achieved on a strain gauged load cell.

1. Select the proper alloy coefficient of expansion.
2. Use strain gauges from the same manufacturing lot for a load cell.

FIGURE 7 Ungauged Piccola pin.

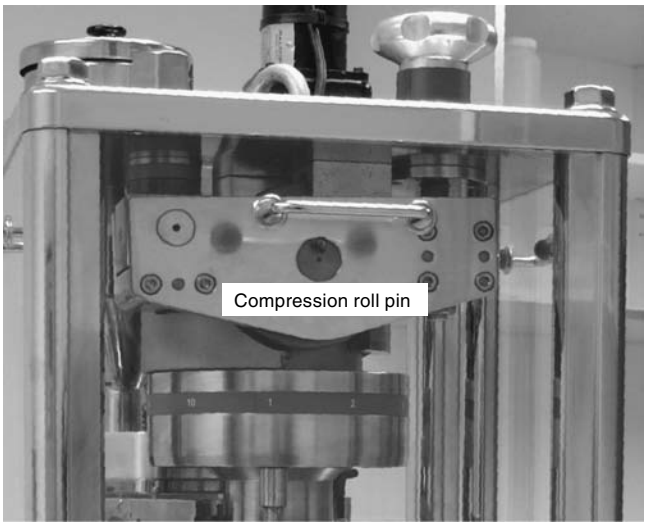


FIGURE 8 Roll pin transducer in tablet press.

3. Perform an oven temperature test and make corrections.
4. Install active circuitry to correct imperfections from step 3.

Alloy STC Coefficient (Self-Temperature Compensating)

The strain gauge manufacture can supply strain gauges where the thermal expansion of the alloy closely matches the thermal expansion of the parent material the strain gauge is adhered to. Strain output because of a temperature change under no load is referred to as apparent strain. Strain that is apparently there but not the result of a load change.

Strain Gauges from the Same Manufacturing Lot

Residual apparent strain from a proper alloy selection can be reduced by using four strain gauges from the same manufacturing lot and the use of a full Wheatstone bridge. Provided that an identical apparent strain resulted from each strain gauge installation, the undesired output from each gauge would be the same, and the positive and negative arms of the Wheatstone bridge would correct the problem. There would be two negative apparent strains and two positive values, the sum of which would be zero leaving only the desired signal as a result of force. The problem is the strain gauges do not react perfectly alike. There may be slight differences in the alloy or adhesive thickness under the gauge, resulting in a change in signal with no change in loading. The telltale sign here is a nonreturn to a zero signal when there is no longer any applied load.

The technology in most strain gauge applications include the above two methods of temperature compensation, but that may not be sufficient for more demanding applications. A tablet press used in research may only be run for short durations at a time and not see any appreciable change in temperature near the load cell. Machines that are run for extended periods of time do get warmer and require additional temperature compensation to maintain their reputed accuracy.

Wheatstone Bridge Third Order Corrections

Now the professionals step in. This is the step that separates the home grown systems from the professional manufacturer. A system should not be promoted as temperature compensated until this step is completed. Two additional temperature-sensitive foil adjustable resistors are installed in each adjacent arm of a Wheatstone bridge. The load cell is slowly heated in a controlled oven to observe the apparent strain of the load cell under a no load but increasing temperature environment. The results are recorded and a calculation performed to determine which resistor needs to be adjusted and to what value. This extra step is time consuming but necessary as it will improve the zero stability by an order of magnitude. In addition, it serves as a quality control check.

Active Circuitry

This degree of temperature compensation is required only if extreme accuracy or unusual temperatures are to be encountered. They are routinely not performed nor need they be as part of a tablet press operation. Basically, an accurate temperature sensor is attached as part of the strain gauge installation and correction made to the data accordingly.

Span or Sensitivity Change with Temperature

The normalized output of a transducer, referred to as mv/v at full scale, will change with temperature. This fact is ignored by the do it yourself crowd but not by commercial manufacturers of quality load cells. Whether or not this is important or trivial for the pharmaceutical industry is questionable. The change occurs because both the gauge factor (sensitivity) of the strain gauges and the modulus of elasticity of the spring element are functions of temperature. As an example, for a typical installation, at an increase in temperature of say 50°F (38°C), the increase in the sensitivity of the strain gauges is about ¼%, while the decrease in modulus of steel is approximately ¾%, a 1% total error if left uncorrected.

Span shifts with temperature can be corrected by inserting a temperature-sensitive resistor in the bridge excitation supply line. With a resistor of the proper value and temperature sensitivity, the voltage to the Wheatstone bridge will vary to offset the span error. In other words, as the full-scale sensitivity of the bridge increases with temperature, the temperature-sensitive resistor will also increase in value, lowering the voltage to the bridge, thereby reducing its output. If performed correctly, the net result is a zero change in full-scale output.

The proof that span shift compensation has been performed correctly is difficult as the transducer must be calibrated at two different temperatures. The nominal value of a selected temperature-sensitive resistor, however, can easily be calculated that will be proper for the material of the spring element. Doing so is not perfect, but will reduce the span error by an order of magnitude making a 1% error discussed above a 0.1% error, one that can easily be ignored for use with a tablet press even in a production environment.

Wheatstone Bridge Balance

Bridge balance means zero output when there is no applied load to the transducer. Installation of four strain gauges into a Wheatstone bridge will need some method of making the output read zero at zero load. This can be accomplished with external signal conditioning or within the bridge itself. Some external techniques distort the geometry of the Wheatstone and introduce system errors, so it is beneficial to perform

this task within the confines of the bridge. This is easily accomplished by installing two adjustable, small but identical values, non-temperature-sensitive resistors, one in each adjacent leg of the bridge. By adjusting the proper resistor, the output of the bridge can easily be made to be zero.

Summary of the Wheatstone Bridge

The simple circuit shown in Figure 1 has now taken on a different appearance. Installation of additional resistors, both temperature-sensitive and non-temperature-sensitive for bridge balance, zero shift with temperature, and span change with temperature makes the Wheatstone appear as in Figure 9.

DISPLACEMENT SENSOR

There are sensors which measure angular (rotational) and linear position.

Linear displacement sensors are widely used in tablet presses. Single station tablet presses use them to determine the position of the upper and lower punches and to correct for tooling and machine compliance. Production tablet presses use displacement sensors to define, control or limit the position of weight cams and roll positions. These types of sensors are available in many forms, from strain gauge, linear variable differential transformers (LVDT) to magnetic and optical (3,5,6).

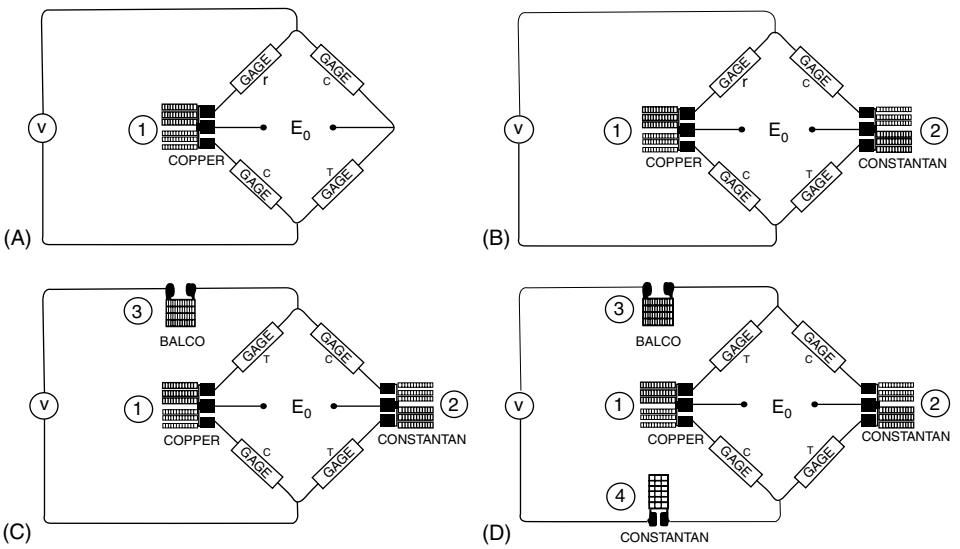


FIGURE 9 Summary of the Wheatstone bridge. (**A**) High-TCR copper resistor (*1*) inserted in corner of bridge circuit, and adjusted to maintain bridge balance over the opening temperature range. (**B**) Low-TCR constantan resistor (*2*) inserted in second corner of bridge circuit, and adjusted for initial zero balance. (**C**) High-TCR Balco resistor (*3*) inserted in bridge excitation supply line, and adjusted to maintain essentially constant transducer sensitivity (span) over the operating temperature range. (**D**) Low-TCR constantan resistor (*4*) inserted in bridge power supply line, and adjusted to set the initial span at the desired calibration level.

An LVDT Displacement Transducer comprises three coils; a primary and two secondary coils. The transfer of current between the primary and the secondary coils of the LVDT displacement transducer is controlled by the position of a magnetic core called an armature. At the center of the position measurement stroke, the two secondary voltages of the displacement transducer are equal but because they are connected in opposition the resulting output from the sensor is zero. As the LVDT's armature moves away from center, the result is an increase in one of the position sensor secondary and a decrease in the other. This results in an output from the measurement sensor. With LVDTs, the phase of the output (compared with the excitation phase) enables the electronics to know which half of the coil the armature is in. The strength of the LVDT sensor's principle is that there is no electrical or mechanical contact across the transducer position sensing element which, for the user of the sensor, means clean data, infinite resolution and a very long life. There is a slight variation of this concept that is called a gauging head whereby a mechanical spring extends the armature to the fully extended position to come in contact with the moving part to be measured without a mechanical connection as with the free style armature. Some designs also contain electronics so that only a DC voltage needs to be applied from a power supply.

LVDT sensors are very robust with nonlinearity from 0.1% to 1% depending on the model. Measurement ranges are generally from 0.5 mm full scale to 40 plus mm full scale. Frequency response is generally greater than 100 Hertz which is more than adequate for even high-speed tablet press or compaction simulator applications.

Noncontact displacement sensors are rarely used as the range is typically limited to less than 5 mm. One application is to determine if a part is in place for safety considerations.

Rotary displacement sensors are being used more on rotary tablet presses today than in the past to accurately define the exact angular position of the turret on a rotary tablet press. Resolvers or their digital counterpart, rotary encoders can resolve an angular change as small as 0.006 °. This is useful to determining the exact punch location relative to a compression roll and the resulting force to evaluate the compact relaxation under the constant strain period known as dwell time.

SIGNAL CONDITIONING

Power Supplies

The power supply is the source of excitation to the sensor. Historically, power supplies were notoriously noisy electrically and tended to drift or change their output voltage values. Today, they are much more stable and smaller in size. That being said it is still prudent to measure the voltage output from the power supply before sampling the voltage from the sensor. In the case of most sensors the output is directly proportional to the applied voltage, noise included.

Ratiometric measurements are the most accurate method to assure that the reading of the signal is independent of the applied voltage. The output from the load cell is normalized by dividing the output from the sensor by the applied voltage from the power supply. This is expressed as mv/v or so many millivolts out per applied voltage in. All quality load cells are supplied with calibration certificates in mv/v and a good data acquisition system should do the same by measuring the power supply and dividing the output signal by this value. All in situ calibrations should also be performed in mv/v and not just as a number in the final units.

Power supplies generally produce either a constant voltage or a constant current and there are advantages to each. Lead wire resistance, for example, is not of concern with a constant current system as it is with a constant voltage where extended lead wire length adds an effective resistance in series with the sensor, reducing the voltage to the sensor. Lead wire lengths are generally minimal around tablet presses but the proper calibration should be performed at the point where the lead wires terminate at the input to an amplifier.

The critical item is that the load cell needs to be matched to the power supply or all of the efforts to temperature compensate the transducer will be incorrect. In the United States the standard is for constant voltage power supplies and load cell manufactures assume that to be true. *If you plan on using a constant current power supply you must order your load cells accordingly*. They will work fine either way but they will not be properly temperature compensated.

What Excitation Voltage Should I Use?

Typical excitation levels used for powering strain gauge circuits range from a high of 15 VDC to a low of 3 VDC. Why the large range and what is appropriate? The answer is it depends on the physical size of the strain gauge, the gauge resistance, the desired accuracy and what material the gauge is bonded to. A strain gauge is like a toaster grid. Current flowing through the grid produces heat that must be dissipated into the material that the strain gauge is bonded to. A strain gauge bonded to copper or aluminum will be capable of dissipating much more heat than one bonded to stainless steel and therefore allow much more excitation voltage. Excessive heating will cause a thermal drift causing a shift in the zero base line of the transducer.

So, if too much voltage is applied the transducer will drift, too little and the output will be too small. For a desired moderate to high accuracy transducers with the strain gauges bonded to steel the power dissipation should be kept to 2 W/in^2 (3 kW/m^2). The correct excitation level is easy to calculate. Using basic Ohm's law relationships, the following equation is easily derived (3):

$$E = \sqrt{RAP},$$

where E is the voltage for the Wheatstone bridge, R is the resistance of the strain gauge, A is the grid area of the strain gauge, and P is the power dissipation of the strain gauge discussed above.

A typical strain used in roll pins for precompression and main compression is a shear pattern from the Measurements Group J2A-06-SO91K-350. This is a 350-Ω gauge resistance with a grid size of 0.125 by 0.105 in. (3.18 by 2.67 mm). Inserting these values into the above equation results in an optimal bridge excitation of 6 V. Some wireless systems apply only 3 V to the bridge; this lower value is in consideration for conserving battery power, not for optimizing performance of the strain gauge circuit.

Strain Gauge Amplifiers

The small millivolt signals from the strain gauge Wheatstone bridge need to be amplified to a higher level voltage for conversion into a digital signal for subsequent analysis. This is generally performed in two steps, each with a purpose. The first amplifier is called a differential or instrumentation amplifier and may only have a gain of one. A second amplifier will usually perform the actual amplification and may have a programmable gain from 100 to 1000 times.

The purpose of the differential amplifier is to remove electrical noise from the environment carried into the amplifier by the electrical cables. The signal cables are in a sense like an antenna with a resistance at the end (the strain gauge bridge). The cable from the strain gauges should be shielded and the wires within the cable twisted and not parallel to each other. Nonshielded exposed wires should be minimized as they will be excellent antennas. The positive signal wire should carry the signal from the Wheatstone bridge; the negative signal should remain at zero volts. If the negative lead were to be attached to an electrical ground this would be referred to as a single-ended input.

For a single-ended input, the positive input to the amplifier would see the signal from the strain gauges as well as any electrical noise which in turn would be amplified by the high gain second stage amplifier. Provided that the negative lead is not attached to an electrical ground but to the negative side of the differential amplifier, this is called a differential input. Since both wires (positive and negative signal) are run within the same cable, and in fact, twisted together both should see the same electrical noise. The purpose of the differential amplifier is to take the difference between the two signal leads, which should eliminate the cable noise and allow only the data through to the high gain amplifier. The common mode rejection (CMR) of an amplifier is a measure of how well this is performed. The higher the CMR, the better the noise canceling and subsequent signal-to-noise ratio.

ANALOG TO DIGITAL CONVERSION

The advent of high speed, high resolution analog to digital conversion (A/D) has enabled large quantities of data to be analyzed and displayed in a meaningful way so that either a person or a feed back control system can respond to the data. The purpose of the A/D converter is to change the incoming analog signal to a series of digital numbers. The rate at which this is performed and the resolution of the conversion will have a lot to do with the overall accuracy of the data acquisition system. Although there are many factors that need to be considered, such as amplifier settling time, switching rates, programmable amplifiers only the major three items will be covered:

- resolution,
- sample rate,
- aliasing and the need for aliasing filters.

Resolution Sample Rate

Resolution is the number of parts that an analog signal is represented by and is described by the number of bits for the conversion process. Mathematically, it is expressed as 2^x where x is the number of bits. A single bit conversion ($x = 1$) with a 5 V DC input can be thought of as any value between 0 and 2.5 V will be put into one bin and any value between 2.5 and 5 will go into a second bin. The greater the number of bins, the greater the resolution. Table 1 shows the relationship between resolution and bits. The last two columns are based on a bi-polar setup that is plus and minus the stated amount. The last column is the resolution for a bi-polar signal where full scale is 50 kN.

TABLE 1 Analog to Digital Resolution vs. Number of Cuts

Bits	Equation	Resolution (one part in)	Percent of full scale	N resolution*
1	Resolution $= 2^1$	2	100	50,000
2	Resolution $= 2^2$	4	50	25,000
3	Resolution $= 2^3$	8	25	12,500
4	Resolution $= 2^4$	16	12.5	6,250
5	Resolution $= 2^5$	32	6.25	3,125
6	Resolution $= 2^6$	64	3.125	1,562
7	Resolution $= 2^7$	128	1.56	781
8	Resolution $= 2^8$	256	0.78	391
9	Resolution $= 2^9$	512	0.39	195
10	Resolution $= 2^{10}$	1,024	0.20	98
11	Resolution $= 2^{11}$	2,048	0.10	49
12	Resolution $= 2^{12}$	4,096	0.05	24
13	Resolution $= 2^{13}$	8,192	0.024	12
14	Resolution $= 2^{14}$	16,384	0.012	6
15	Resolution $= 2^{15}$	32,768	0.006	3
16	Resolution $= 2^{16}$	65,536	0.003	1.5

Looking at the table above it would appear that the 10 or 12 bit resolution would be more than adequate for the acquisition of data on a rotary tablet press, and that would be the case provided that an amplifier gain was unique for each channel that raised the milli-volt signal to the full scale of A/D converter. Typical amplifier gains are fixed, however, and not optimized, letting the resolution of the A/D converter solve the shortcomings. Let us take two realistic examples.

Example 1: A transducer with a 2.0 mv/v output; excitation voltage of 3V, a fixed gain amplifier of 64 and a 12 bit A/D. Determine the percent resolution and equivalent number of Newton's with a full scale of 50 kN at 5V.

Transducer output of 6 mv is amplified to 0.384V with the fixed gain of 64 amplifier. A 12 bit bi-polar A/D can measure 1 part in 2048 out of 5V or 2.4 mV. 2.4 mV resolution with a 0.384V signal represents 0.64%. Therefore, what appeared as a resolution of 0.05% quickly became 0.64% or 320N on a 50 kN transducer.

Example 2: A transducer with a 2.0 mv/v output; excitation voltage of 5 V, a fixed gain amplifier of 64 and a 14 bit A/D.

The transducer output is 10 mv amplified to 640 mV with the amplifier. The 14 bit bi-polar A/D can measure 1 part in 8192 out of 5 V or 0.61 mV for a resolution of 0.095% or 47.5 kN on a 50 kN transducer. By using a higher excitation and a 14 bit A/D, the resolution became close to 7 times better and more in line with the requirements for a tablet press transducer system.

Resolution Summary

High resolution analog to digital converters are commonplace today and at reasonable prices and performance. Common practice in the past was to use adjustable amplifier gains to optimize the transducer full scale to that of the input of the A/D converter. For instance, a 10mV signal would be amplified with an amplifier gain of 500 to produce a 5V signal for a 5V input to the A/D converter. Today programmable gain amplifiers are used that cannot be adjusted so the full scale input signal to the A/D is less than optimal.

Sample Rate

Frequency response, sampling rate, and Nyquist theory are commonly misunderstood. Sample rate is easy; it is the number of times a digital reading is taken over a period of time, usually one second. This is sometimes expressed in Hertz. Therefore, a 100Hz digital sample rate is 100 equally time spaced samples taken for each second.

The confusion is the word Hertz. In the analog world Hertz refers to the number of cycles per second. Therefore, in analog speak; a 1Hz sine wave or one cycle per second may require 10 samples per second to represent the sine wave. In digital speak this is a 10-Hz rate. In other words, for this example, it takes a 10 Hertz digital sample rate to define a 1Hz analog signal.

Nyquist theory states that the frequency content of any analog signal can be determined with a sample rate of only twice that of the analog frequency. The common misconception is that the analog frequency need only be doubled with the digital sample rate to reproduce the original data. That is not what the Nyquist states and it is very misleading. Nyquist states you can obtain correct frequency information this way but says nothing about reproducing the shape of the data. There is a relationship between the number of samples required to define a cycle and the statistical error of missing the peak value of the cycle. The graphic below clearly shows the problem. The analog sign wave is being sampled at a rate of 5 samples per cycle. The computer would basically connect the dots, making a pseudo square from this sine wave.

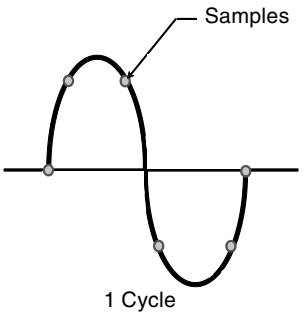
Provided that you wish to limit your peak detection error to 0.25% you must sample digitally 100 times the analog frequency contained within the data. Such high sample rates are generally not used and the user is never aware of what is being missed. For tablet press instrumentation, a digital sample rate (Hertz) of at least 10,000 is required to cover all presses and transducers (Fig. 10).

Aliasing Errors

Nyquist states as follows:

If frequencies greater than ½, the sampling rate are allowed to the input of the A/D converter, the higher frequency *will* erroneously be represented by a lower frequency that *cannot be separated from the real data*.

The only way to eliminate this error is to use an anti-aliasing filter prior to digitizing the input signals (Fig. 11). Therefore, if a sample rate of 10,000 Hz is to be used a



Samples/Cycle	% Error
5	5.00
10	2.50
20	1.25
50	0.50
100	0.25
200	0.10

FIGURE 10 Sample rate vs. error. For example: If the frequency of your data is 100 Hz and you desire a maximum error of 0.25%, you must sample the 100 Hz at 100 samples per cycle or 10,000 samples per second.

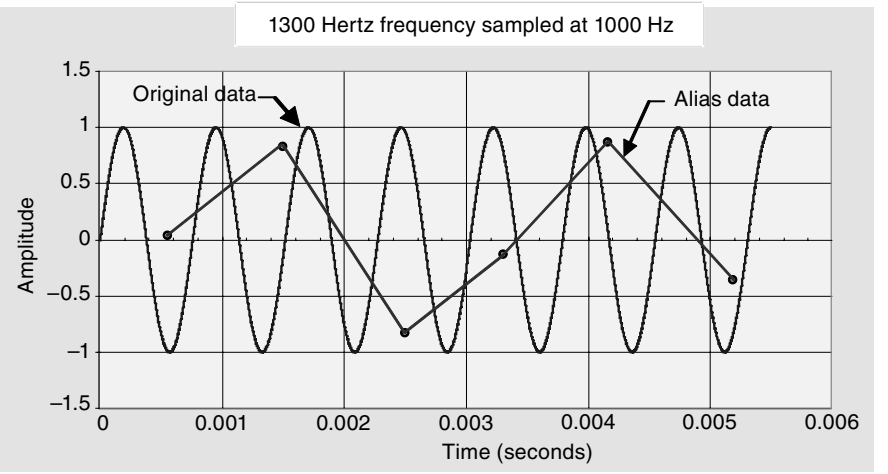


FIGURE 11 Aliasing error.

low pass analog filter of < 5,000 Hz must be used to prevent aliasing errors. This filter will prevent analog frequencies of greater than 5,000 Hz from being digitized. Just because the higher frequencies are not present when the system is installed does not mean they will never be present. Changes in equipment in the facility, use of hand held radios or even new utilities can be the source of high frequency noise.

Any good data acquisition system must incorporate such protection into the design or the user will someday receive incorrect information and never even know that his system is creating new data to superimpose on the actual data.

A classic example that most of us can relate to is the wagon wheel in a western movie. The camera is taking pictures at a fixed rate, say 60 frames per second. If the wagon wheel makes 90% of a rotation between frames the wheel will appear to have rotated backwards by 10%. Wrong in both magnitude and direction! The same phenomena will occur will your data acquisition system if it is left unprotected without the use of an anti-aliasing filter.

REPRESENTATIVE TABLET PRESS TRANSDUCER CALIBRATIONS

Examples of Tablet Press Transducers

Instrumented compression roll pin for a Piccola bi-layer tablet press (Fig. 12) (4). Instrumented ejection ramp for a Riva Piccola tablet press (Fig. 13). Back side of a not yet strain gauged ejection ramp for the Piccola tablet press showing the pockets where the strain gauges will be placed (Fig. 14). The two spring elements are differential bending beams on each end with a relief in the middle (4).

Calibration

Calibration is the comparison of a component or group of components against a known and recognized standard under a specific set of conditions. A system is considered within calibration if it complies or can be adjusted to comply with the acceptable uncertainties.

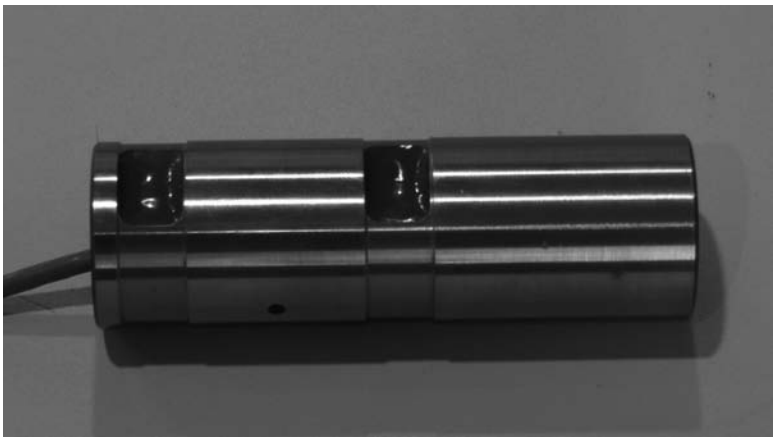


FIGURE 12 Representative tablet press transducer.

Validation in the sense of measurement systems is a set of calibrations over the environmental conditions the system must perform within. This implies that if a measurement system is to operate over a specified temperature and humidity range; it must be calibrated over the extremes to be validated.

In the United States, the National Institute of Standards and Technology (NIST) maintains standards and is considered the arbiter and ultimate U.S. authority for values of SI units and industrial standards. NIST also provides traceability to its standards by calibration, by which an instrument's accuracy is established by comparing, in an unbroken chain, to the standards maintained by NIST. For each step in the process, the measurement uncertainty is evaluated.

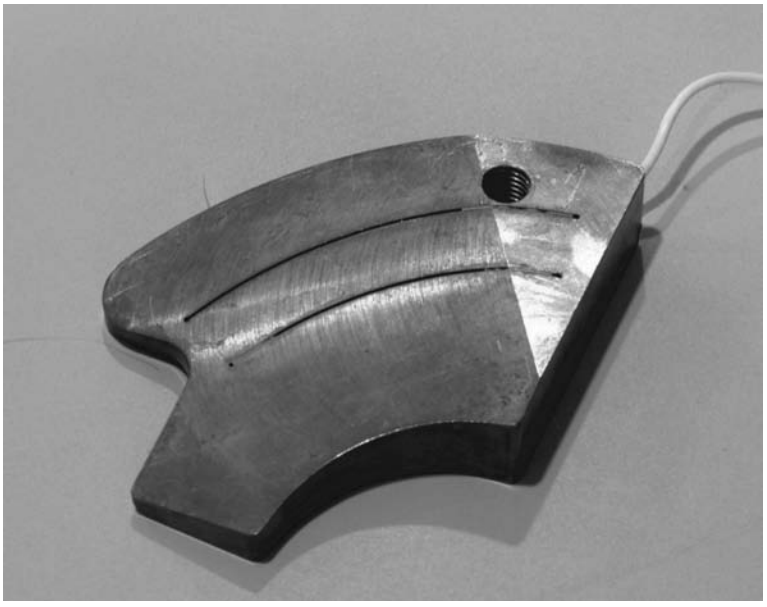


FIGURE 13 Instrumented ejection ramp.

FIGURE 14 Back side of piccola ejection cam showing strain gauge pockets

Traceability is the property of a standard whereby it can be related to stated references, usually national or international standards, through an unbroken chain of comparisons, all having stated uncertainties. The level of traceability establishes the level of comparability of the measurement: was the measurement compared to the previous one? was it compare to a measurement from the day before? was it compared to a measurement from a year ago? or was it compared to the result of a measurement performed somewhere else in the world?

Figure 15 shows the organizational chart for the standards in the United States. It is a Federal offense for one to misrepresent their facility and may well result in time spent in jail and a personal meeting in front of the Senate. Most in-house calibration facilities fall into instrument maintenance while companies specializing in calibration services are secondary laboratories. Secondary laboratories rely on a primary laboratory for their internal standard to be calibrated that will in turn rely on a direct NIST calibration for their standards. Therefore, the calibration performed by a process application technician must have an unbroken chain of traceability directly to NIST.

The level of uncertainty increases the longer the chain from NIST. A secondary laboratory will rely on the standards of the primary laboratory to be in compliance with the requirements of the NIST.

Calibration of Tablet Presses

Calibration of a rotary tablet press needs to be done with caution as it is easy to make an incorrect calibration. Calibrated punches can become misaligned, causing excessive friction resulting in a loss of applied force to the machine load cell. The calibrated punches should have at least two standards, one each in the upper and lower punches with

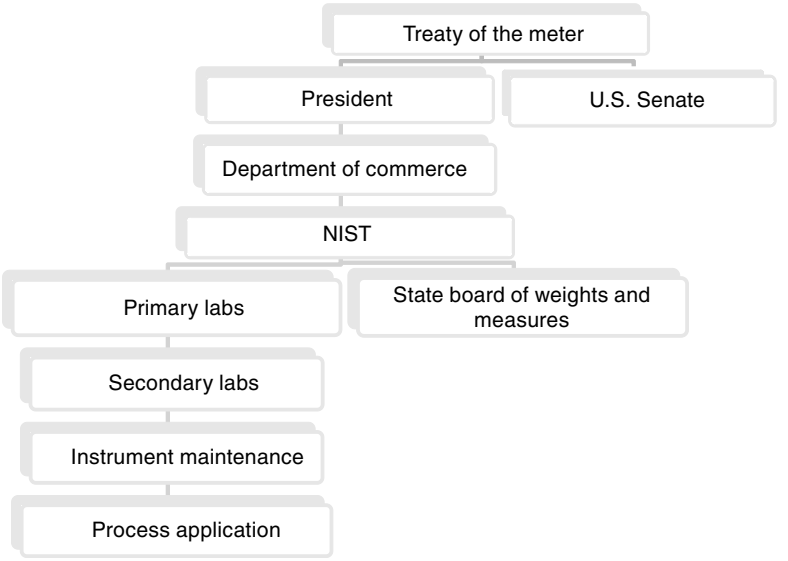


FIGURE 15 United States standards structure.

procedures to make sure the standards agree with each other before the results can be accepted. One vendor of calibration services uses three standards in line to ensure that none of the applied load is being lost due to friction from misalignment. It is interesting to note that misalignment is not obvious to the eye, and there is no method of knowing that it had occurred if only one reference is used, the resulting calibration will look completely normal, just with incorrect values.

There are two basic methods of performing a static calibration on a rotary tablet press. One is to perfectly align the modified punches between the rolls and apply the load with a hydraulic ram while acquiring data from the standards and the machine load cell. The second method is to install the modified punches prior to the rolls and using the machine hand wheel, roll the punches through the compression cycle. The first method can apply a higher force smoothly and with more control, and is easier to ensure the modified punches are properly aligned. The second method is quicker and does not involve hydraulic rams, pumps, and hoses; however, the load cannot be controlled as well. Both methods produce acceptable results.

Figure 16 shows a field hydraulic loading system with two different capacity jacks. Figure 17 shows a calibrated punch that will be rotated under the compression roll by the machine hand wheel.

Calibrated Punches

The design of a custom punch to be used as a standard or reference must follow the general rules of transducer design (4):

1. The mechanical design of the punch must be such that it has excellent sensitivity in the direction of the desired force to be measured and low sensitivity to all undesired forces.
2. The placement of the strain gauges should be such to electrically cancel any residual stress from all other undesirable forces, such as side loads.

FIGURE 16 Calibration kit view 1.

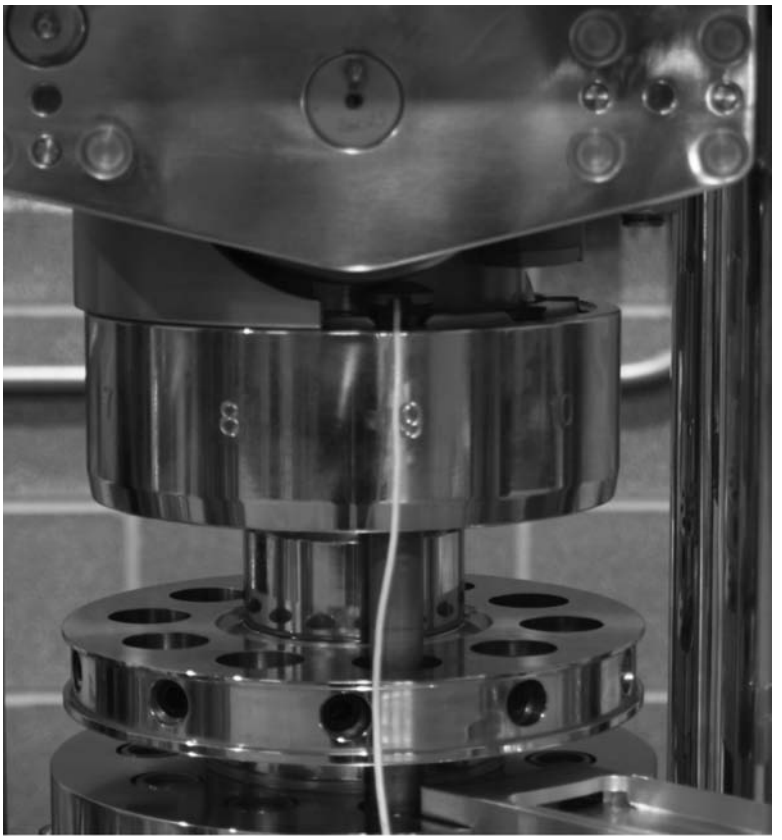


FIGURE 17 Calibrated punch in tablet press.

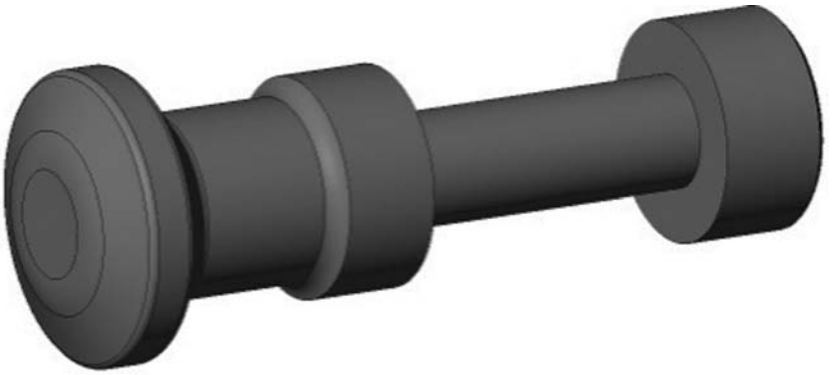


FIGURE 18 Calibrated punch reduced cross section in design.

3. Placement of the strain gauges within the Wheatstone bridge to cancel unwanted forces and respond only to the desired force.

Let us compare three potential mechanical designs for a 50-kN calibrated punch spring element.

Design 1

Machine a smaller diameter on the punch barrel and install a Poisson full bridge set of four strain gauges (Fig. 18).

Reducing the outside diameter to 14 mm from the original 19 mm to allow room for the strain gauges and yield a correct sensitivity for calibration purposes results in a cross sectional area of 154 mm^2.

The axial stress on the reduced area is:

$$\text{Stress} = \frac{\text{Force}}{\text{Area}}$$

The equation for bending because of an offset load such as when the punch contacts the roll is:

$$\text{Stress} = mc/I$$

where c is the distance from the punch centerline to the position of the strain gauges and m is the bending moment. I is the moment of inertia which is $\pi d\ 4/64$ for a circular cross section.

Using the above equations and geometry, the axial and transverse sensitivity can be computed.

Design 2

Machine flats on the punch barrel to install strain gauges (Fig. 19).

Design 3

Machine pockets in the punch to install the strain gauges. This results in a cross-sectional area resembling a structural member used in building and bridge construction called an I beam. As expected this design offers many advantages. In fact, this design is five times more resistant to undesirable bending forces than the other two (Figs. 20 and 21).

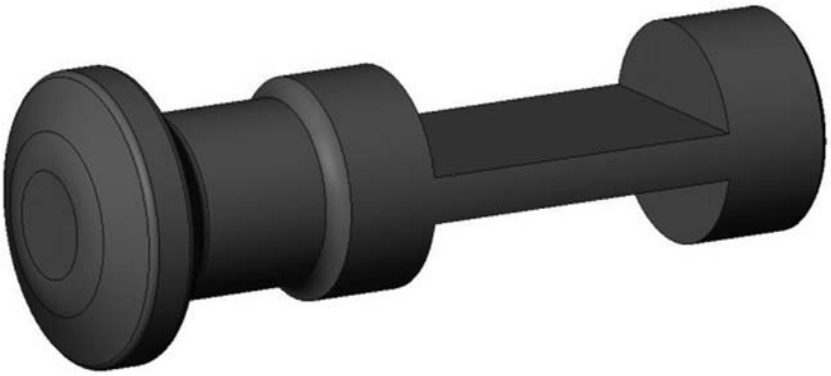


FIGURE 19 Calibrated punch rectangular.

Using the Calibrated Punches

The strain gauged punch must be calibrated against a recognized standard to be used as a calibration standard. It must be calibrated on a regular interval as dictated by Company SOP. The SOP at SMI is that the punch must be calibrated against a standard every three months and the standard must be sent to an independent agency for certification within the last 12 months. This policy prevents in-house propagation of errors. Another part of the SMI procedure is that one set of strain gauges will be installed in each of three pockets, one in the upper punch and two in the lower punch, in essence making three standards in use during a calibration. These three standards must agree within established criteria before the calibration is acceptable.

Application of the Force

The force is generally applied in one of three ways.

1. Insert a hydraulic jack in line with the calibrated punches and use a hand pump to apply pressure to the piston. The punches are generally pre-aligned between the rolls. The load is applied gradually and many points can be obtained from zero to full scale. At SMI over 1000 points are obtained and a regression analysis is performed to obtain the stated sensitivity and errors.

FIGURE 20 Calibrated punch pocket design.

FIGURE 21 Cross section of pocket design.

2. Align the calibrated punches between the rolls as before and use the hydraulic system of the tablet press to produce a load in place of the in-line jack.
3. Position the calibrated punches before the compression rolls and rotate the turret manually through a compression cycle. This method is excellent for a quick check of the force measurement system at a limited number of force levels.

The calibration kit shown in Figure 22 shows some of the components used for method one above.

The instrument in the upper left is a transducer simulator and is used to apply a calibrated input to the balance of the data acquisition system.

The Balance of the System Requires Calibration Also!

The emphasis to date in this chapter has been on the actual force transducer installed within the machine. It is, however, only one link in the chain. Other components, collectively referred to as signal conditioning must be calibrated as well, such as power supplies, amplifiers, analog to digital converters.

The instrument in the upper left of Figure 22 is a transducer simulator and is used to apply a calibrated input to the balance of the data acquisition system. It is this instrument that is used to input a traceable ratio-metric mv/v signal into the signal conditioning. The transducer is temporarily disconnected from the signal conditioning and the transducer simulator installed in its place.

The transducer simulator inputs an ascending and descending signal to the system in 10% increments from 0% to 100% of full scale. All recorded data points are regressed to determine accuracy and linearity. Power supplies, amplifiers, and analog to digital convertors are so accurate today that a typical overall error is $<0.05\%$ of full scale with a rejection tolerance of 0.1% (4).

"It is much better to be approximately accurate than precisely wrong" (7).

Two terms that are frequently interchanged are accuracy and precision. They do not mean the same as illustrated in the example of the target below. Precision is the tight grouping of bullets (data) in a location not necessarily where desired. If you were a deer

FIGURE 22 Calibration kit view 2.

hunter every shot could be precisely in the same spot, all way over the top or short of the desired target. Making an adjustment in your rifle sights (instrumentation) could correct this problem. Accuracy is a random grouping within a specified tolerance of the target center. A tight accuracy tolerance would lead to precision at the target center (Fig. 23).

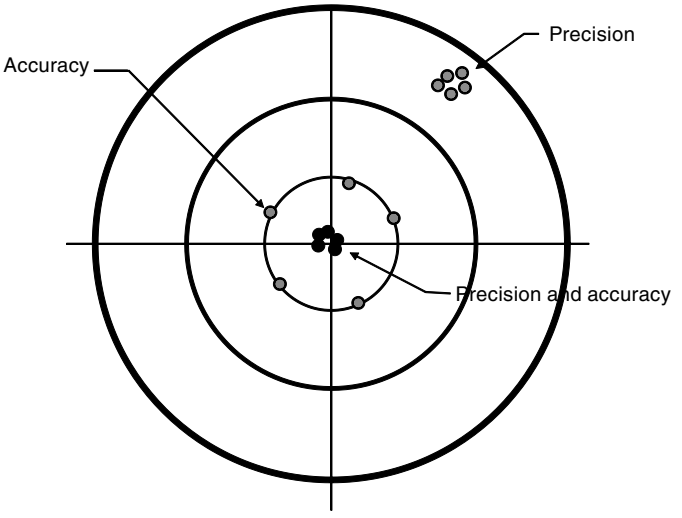


FIGURE 23 Accuracy vs. precision graphic.

ANALYSIS SOFTWARE

The software package is the means of presenting large amount of data into meaningful information, such as charts and graphs in engineering units. Because this front end interface is the only exposure the scientist has to the data acquisition system, it is often thought of as "*the data acquisition system*." This of course is not true; the software is the pretty front end of all the components of a data acquisition system and is perfectly willing to display incorrect results from a transducer in a very attractive format. Validation engineers often go to great lengths to ensure the compliance of the software only to neglect the balance of the data acquisition system. This may be true as most validation engineers have a computer, not an instrumentation background. Such an attitude will lead to a false sense of security if the entire system is not addressed in the validation.

A well designed software program will provide the press operator with real time force feedback, converting data streaming in at thousands of samples per second into useful information. Each manufacturer will have their own offering for displays and features; I will use the screens from the SMI Director Program to discuss the purpose and use of typical real time and post analysis data presentations.

Real Time Presentations

Peak Value Bar Charts

The graph in Figure 24 is displaying the peak forces for an eight station tablet press during the last turret revolution. Notice that in this example, all of the bars are the same length and the digital values are all 17.5 kN. In order to achieve this, tooling must be perfectly matched and the material flow into the die excellent, an unrealistic occurrence.

The information available with this type of presentation is of great value. A quick glance will verify not only the compression force levels, but the uniformity of the forces for each station. One station with a higher or lower force will stand out immediately and generally indicate a problem with the tooling in that station. A random distribution will speak to the flow ability of the material into the dies. The tabs at the top will allow the operator to display the available transducers.

Oscilloscope Display

The oscilloscope display displays the entire force time profile, not just the peak value. Figure 25 shows main compression for consistency, but the scope mode is most useful in trouble shooting ejection and take off forces because a punch that is showing a high ejection force or tablet removal from the lower punch tip is immediately obvious. This program displays the *x*-axis in degrees of turret revolution. Other programs may use time. The advantage of degrees is that a tooling station is always on the chart at the same location, independent of turret speed.

Figure 26 illustrates a potential ejection problem as the breakaway force, resulting from a higher static than dynamic coefficient of friction, is significant relative to the push out force. Although the actual ejection force levels are reasonable this situation is a red flag for much higher ejection forces to follow, as the data in Figure 27 shows. Ejection forces of this magnitude are excessive and will result in premature wear on both the ejection ramp and punch heads. The high ejection forces shown in Figure 27 occurred only a few turret revolutions later than that shown in Figure 26.

Looking at compression events with an oscilloscope function yields little additional information, perhaps even less, than with the use of a peak value bar chart. Looking at an ejection transducer such as Figure 26, on the other hand, is extremely useful in avoiding

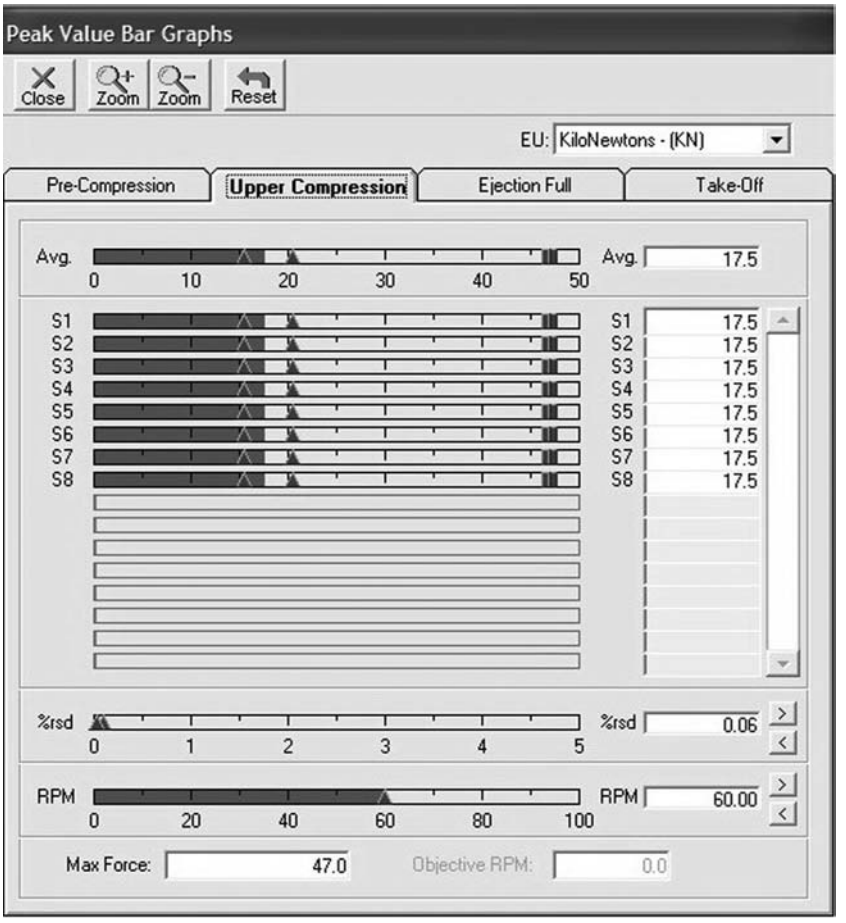


FIGURE 24 Oscilloscope display.

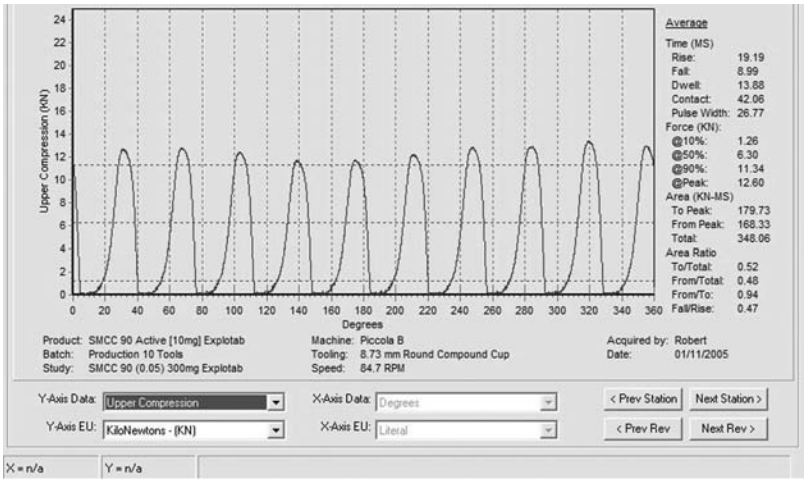


FIGURE 25 Compression scope traces.

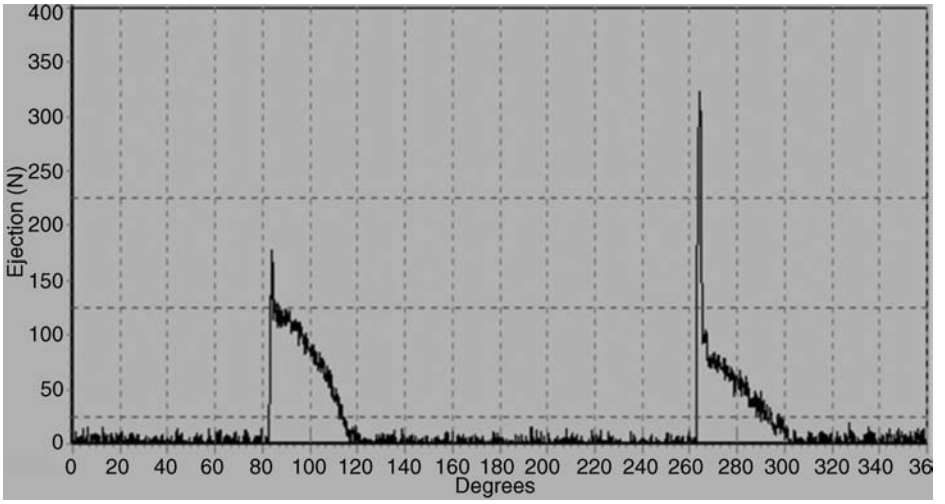


FIGURE 26 Ejection scope traces with high forces.

problems. Figure 28 demonstrates how two different excipients react to increasing compaction pressure, one increasing; the other remaining relatively constant. The upper trace is typical of lactose and mineral-based excipients, the lower, MCC.

Limits and Control Charts

Figure 29 is an example of a typical control chart where the dark dashed line in the middle represents the average compression force for 1000 turret revolution and the lightly dashed lines above and below are ± 1, 2, and 3 sigma standard deviation. Notice that a control chart does not display the target force or any limits, the intent of a control is merely to show that the process is in or out of control. The example shown in Figure 29 would be out of control as there are too many samples above the average between 450 and 600 revolutions.

A limits chart is the same data as shown in the control chart plotted against user defined limits and target. Generally, there are two upper limits, two lower limits, and a

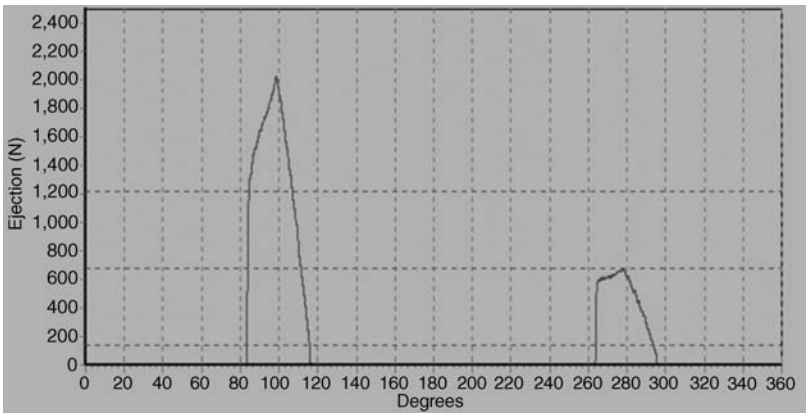


FIGURE 27 Excessively high ejection forces.

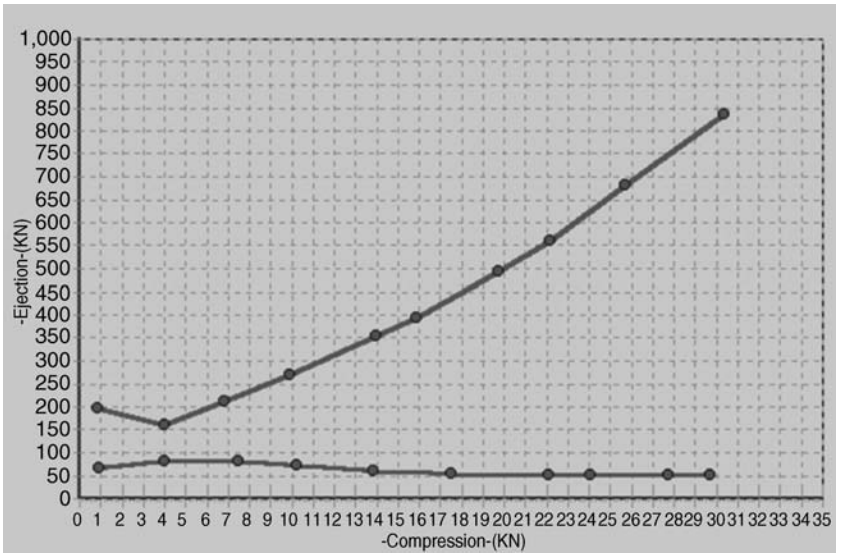


FIGURE 28 Ejection force versus compression force.

target specified. Figures 29–31 display the same data, the first a control chart, the second a limits chart and lastly a histogram. The tags at the top of the histogram bars represent the percentage of samples that fell within that bar.

Post-Acquisition Analysis

After the data are acquired and stored, additional analysis is generally possible beyond what was available in the real time displays.

Rotary tablet presses are frequently used to generate compaction and strain rate studies, detailed oscilloscope analysis of the compression or ejection events as well as several levels of summary reports.

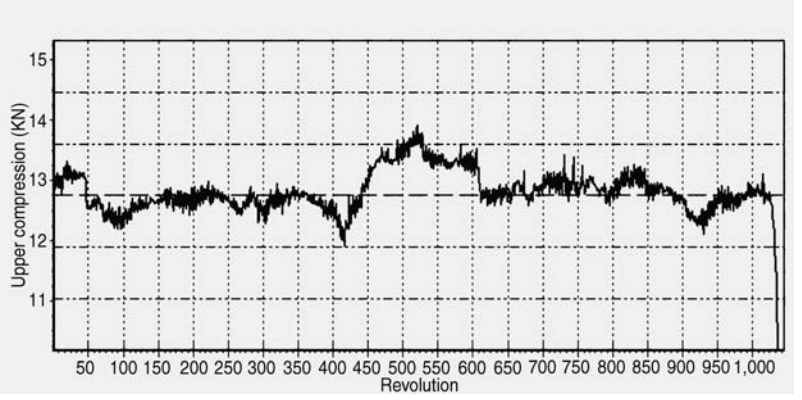


FIGURE 29 Control chart [SMCC 90 Active (10 mg) Explotab].

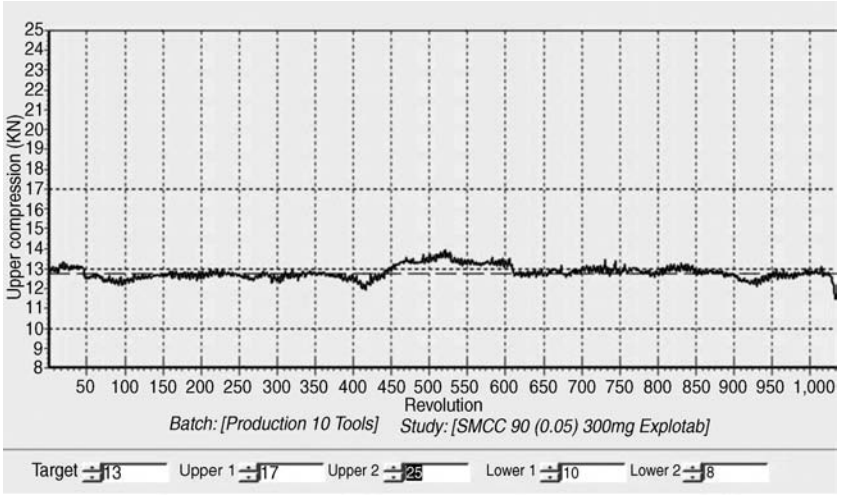


FIGURE 30 Limits chart [SMCC 90 Active (10 mg) Explotab].

Single station tablet presses can be used as a "cheap man's compaction simulator" to generate force displacement, work, heckel, porosity graphs, and radial die wall (8–10).

Detailed Oscilloscope Traces

A detailed analysis of a compression or ejection event is possible provided that the information is saved to a file. This detail can provide insight as to the compaction characteristics of a formulation, especially relating to the recovery process after main compression. Figure 32 shows a typical compression along with the details pertaining to the event. For the Director Analysis program the following definitions apply. Note that several ratios, such as fall time/rise time and area from peak/area to peak are calculated for the formulator to aid in characterizing the formulation. To aid in the visualization, the horizontal dashed lines represent 10%, 50%, and 90% of the peak force.

Rise time: The time from 10% of peak force to 90% of peak force.

Fall time: The time from 90% of peak force to 10% of peak force.

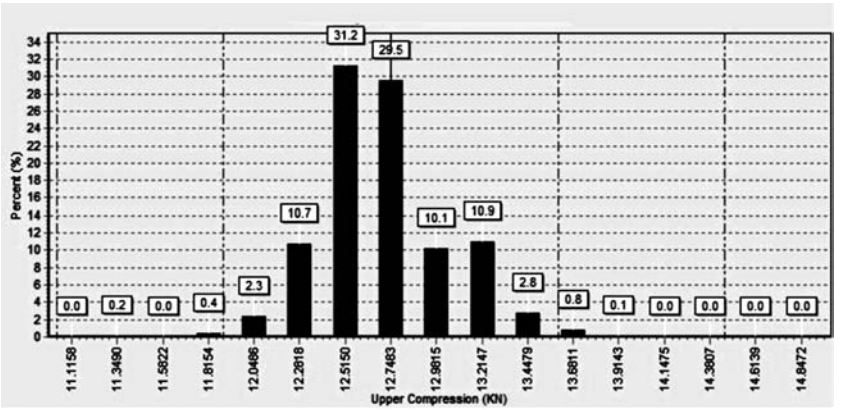


FIGURE 31 Histogram [SMCC 90 Active (10 mg) Explotab].

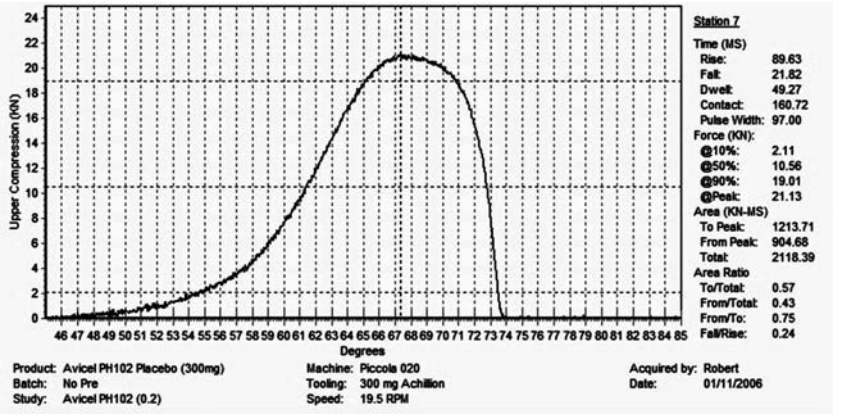


FIGURE 32 Detailed compression event.

Dwell time: The time from 90% of peak on the rise to 90% of peak on the fall.
Pulse width: The time from 50% of peak on the rise to 50% of peak on the fall.
Contact time: The time from 10% of peak on the rise to 10% of peak on the fall.

Compaction Profiles

During a compaction study the turret speed is kept constant and the compression force varied. Tablet breaking forces are measured for each compression force level and entered into the program. Based on the tablet geometry and breaking force, the program calculates the tablet tensile strengths for each compression force level and present the data in a graphical format. Overlays make for an easy comparison as shown in Figure 33.

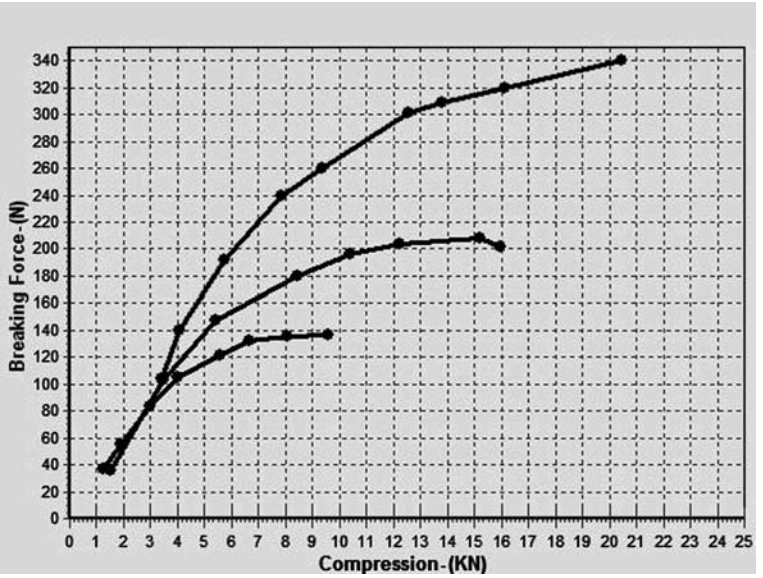


FIGURE 33 Breaking versus compression force.

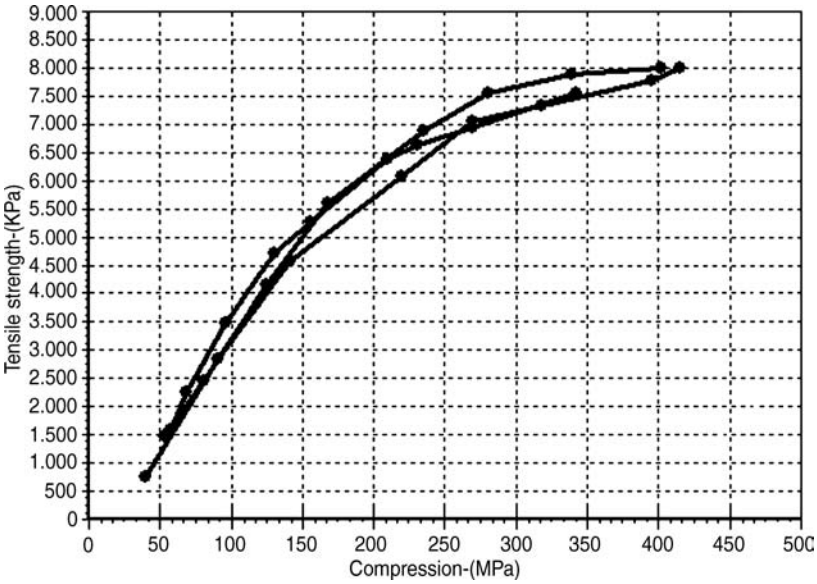


FIGURE 34 Tensile versus compression strength.

The curves shown in Figure 33 represent three different tablet sizes and weights from the same formulation. The lower graph is a 75 mg tablet, the middle a 150 mg, and the upper, a 300 mg tablet. It is clear and understandable that it takes more force to break a larger tablet than a smaller one of the same material and force level. Normalization of the compression force to compaction pressure and the breaking force to tensile strength yields almost identical results for the three sizes, as shown in Figure 34. All data, at least in the R&D environment should be presented in this manner. Basic understanding of tensile strengths that are required to withstand shipping and handling, coating, dissolution, etc. can easily obtained that are not obvious when the data are not normalized.

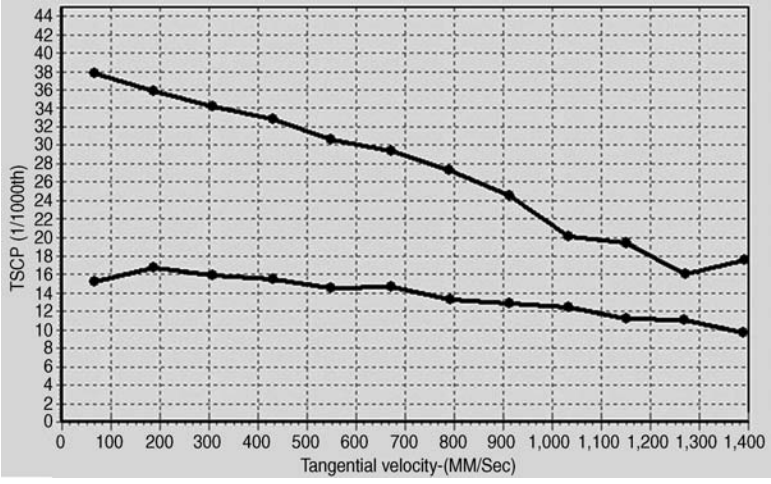


FIGURE 35 Strain rate study.

Strain Rate Studies

A strain rate study maintains a constant force and varies the turret speed from low to high. The intent is to evaluate how the material will perform when transitioned from a low-speed machine to a high-speed production model. The turret speed on different machines will result in different tangential velocities depending on the machine pitch circle diameter. The program should account for this in the analysis and graphical presentation. Figure 35 shows such a presentation for two different formulations, one of which is clearly more strain rate sensitive and might pose a problem in production.

SUMMARY

An instrumented tablet press in an R&D environment is not a luxury today; it is a necessity if one wishes to practice good science and have a deeper understanding of compaction principles. It is possible to design an in-house system and many have been built and put to good use. Today, there are several commercial options that should be considered first to see if they fit into the company needs as thousands of man-hours have been invested into their design by the manufactures. Whatever the path, do instrument or purchase an instrumented tablet press. It will shorten development time; enable easier transition from R&D machines into production models resulting in a quick return on the initial investment.

A properly designed data acquisition system needs to be based on sound mechanical and electrical principles. "You ask a measurements system for the truth, the whole truth, and nothing but the truth, not its opinion." Incorrect components are perfectly willing to moonlight providing more information than you wanted. Some force transducers produce a nice signal when exposed to a strong light source, others from temperature and still others due to improper mounting. This is not acceptable. There are many who purport to being "Instrumentation Experts," do not be duped into believing a fancy software program makes for a well-designed instrumentation system. The transducers must fit the application; power supplies must match the transducer requirements of either constant voltage or constant current, the resolution of the analog to digital conversion must be appropriate for the application and use ratio-metric measurements. Sample rates must be determined for the required frequency response and proper use of anti-aliasing filters employed. The entire system must be able to be calibrated, not just the transducers and finally there must be a software system that can condense all of the data into a meaningful and usable format.

BIBLIOGRAPHY

1. Celik M, Oktugen E. Dev Indust Pharmacy 1993; 19 (17&18):2309–34.
2. Cocolas HG, Lordi NG. Drug Dev Indust Pharmacy 1993; 19 (17&18):2473–97.
3. Hoag S. Tablet compaction issue. Eur Pharmceut Rev Issue 2005; 2:104–11.
4. Kistler Instrument Corp., Amherst, NY. (Accessed September, 2007, at, http://www.kistler.com/do.content.us.en-us?content=90_Support_Download)
5. Marshall K. KMA Associates, Brick, NJ, Conversation.
6. Microstrain, Williston, VT. (Accessed September, 2007, at http://www.microninstruments.com/support/help/index.htm)

7. PCB Piezotronics, Inc., Depew, NY. (Accessed September, 2007, at, http://www.pcb.com/techsupport)
8. RDP Electrosense, Pottstown, PA. (Accessed September, 2007, at http://www.rdpe.com/us/mendisp.htm)
9. Specialty Measurements Inc., Lebanon, NJ: Internal Publications.
10. Vishay Micro Measurements, Wendell, NC. (Accessed September, 2007, at, http://www.andrusspeskin.com/mg/mgnotes.html)

3
Pharmaceutical Manufacturing: Changes in Paradigms

Jean-Marie Geoffroy
TAP Pharmaceuticals, Inc., Lake Forest, Illinois, U.S.A.

Denise Rivkees
Pfizer, Inc., Morris Plains, New Jersey, U.S.A.

INTRODUCTION

Pharmaceutical science involves the study of dosage form design and physiologic disposition along with methods used to control and test the design and disposition. The ultimate goal of the dosage form design is to manufacture a dosage form that can be delivered through the market to the site of action in the patient. A thorough understanding of what manufacturing is, how it works, and the regulatory requirements that affect the manufacturing process can enable the delivery of a manufacturing process and product that meets the needs of the manufacturing organization, and the needs of the patient and the healthcare community.

The purpose of this chapter is threefold (*i*) to give an overview of pharmaceutical manufacturing in its current transitional state all the way from totally manual to completely automated, (*ii*) prepare the scientist for any working environment between the two, and (*iii*) help the scientist understand how the movement from manual systems to automated systems can improve production processes and the overall operation of the manufacturing facility.

There are two aspects of manufacturing for the pharmaceutical scientist: the verb manufacturing, for which the development scientist is involved with the design of a manufacturing process or making clinical supplies–and the noun, physical part of a company responsible for manufacturing marketed products (and in some companies, clinical supplies). This chapter will review the basic elements of a manufacturing organization and how these elements work together. The scientist who finds him/herself in a research and development organization will see elements of both manufacturing and the manufacturing organization within the research department to a greater or lesser degree depending on the company. Depending on the size of the company, the scientist may work completely with the commercial manufacturing organization. Most pharmaceutical scientists start in the preformulation or formulation area, and if there is interest in manufacturing, move closer to work on marketed products after some experience has been gained.

This chapter will also cover the way process automation is being integrated into manufacturing processes and operations. The first part will demonstrate how manufacturing

historically operates without information technology so that the scientist understands the basic operations being performed without the overlay of electronic controls. The transition from manual to automated systems has been enabled by the development of information technology. Through the use of computer systems and integration of sensors to those systems, we are now able to capture data, use mathematical modeling to make predictions, and document quality based on real time data. As discussed later in this chapter, Food and Drug Administration (FDA) and other governing and regulatory bodies has led the way in concert with industry to enable the use of technology to improve quality and business practices.

Keep in mind that companies differ for a variety of reasons. First, not all operations are exactly the same. Second, many companies in the pharmaceutical industry are decades old, and their processes have advanced with technical and scientific understanding. Third, there are thousands of dosage forms, active pharmaceutical ingredients, excipients and processes that are used to deliver a therapeutic effect to active physiologic sites. Fourth, although the functions performed by a company are the same, not all companies are organized the same way.

In this chapter, we have tried to summarize the general functions using the titles that most companies use, but the scientist should be prepared for differences in the way companies operate and how they label their departments and functions. Likewise, we will use solid dosage form examples in this chapter because they are the most common (these principles apply to any dosage form). The granulation or coating process for a solid dosage form may slightly vary between companies and/or products within the same company based on the available science and development philosophy of the company at the time the dosage form was developed.

The last part of the chapter transitions to the state of pharmaceutical manufacturing where the influence of technology in terms of its applicability to process monitoring and control with Quality by Design (QbD) will be discussed.

Manufacturing Goals

The goal of the manufacturing organization and technical operations is to make the same product(s) reproducibly over the lifecycle of the product. On the other hand, the goal of research is to define the parameters under which a new product can be consistently made, and to understand its disposition in the body. These two different paradigms lead to different cultures for the organizations. Manufacturing is a culture where rules must be followed and innovation must be introduced in the context of manufacturing where many activities occur simultaneously and the work of individuals overlaps. Manufacturing is a place where a predetermined set of systems control each step throughout production. These systems are not only required by regulations, but make good business sense as well. All functions in manufacturing are interrelated (similar to a mixture problem statistically), so when something happens to affect a single function, it has an impact on other functions.[a]

[a] When new products are introduced, when troubleshooting is required, or when changes are requested, communication must go through several departments before any change occurs, usually in the form of a written protocol. It is frequently the job of the pharmaceutical scientist to get "buy-in" from people in other departments before a study is started, even if it is analytical approval to analyze samples. As such, there may be a time delay before experimentation and/or implementation can occur. Upfront planning will minimize delays as much as possible.

There are three basic components that are used to make a product, raw materials, equipment, and a process. These basic elements evolve into all the parts of a company needed to manufacture a product. Although the exact mechanism of interdepartmental communication and organization depends on the culture of individual companies, all manufacturing organizations encompass the functions discussed below. We will start with organization from the view of materials entering the plant, how they are stored, processed, tested, documented, and regulated. We will then move to the broader context of an integrated manufacturing organization.

Supply Chain

The work of manufacturing is dictated by the Supply Chain. Orders for product originate with the Supply Chain, which is the shipping and inventory control part of manufacturing. The Supply Chain Department works with wholesalers and sometimes directly with pharmacists and physicians to deliver finished product to the market.

When the Supply Chain needs a product, orders usually go to some type of planning or Materials Department. The Materials Department keeps inventory control over raw materials and either issues the batch production record or directs that manufacturing or the quality department issue the batch production record. Production scheduling is governed by the generation of batch production records, usually called the batch record or production order.[b]

Materials

Upon ordering a raw material from a vendor, the raw material is shipped from the vendor, received by the manufacturer receiving department, stored in non-released raw materials warehouse (quarantine), sampled by the manufacturer, tested to meet certain specifications by the quality department (each test dictated by a standard operating procedure and carefully documented in a bound lab notebook or other document satisfactory for an audit), released for use by the quality department, moved to released material storage, requisitioned for use in a batch, dispensed by weight on an order from a batch production record, moved to a batch staging area, moved to the individual production module, then charged to the batch.[c]

At each movement of a raw material, it is stored in a preassigned area. Each movement of the material through the system generates documentation that must be signed by the person who completed the move, whether it is a material representative or a quality department release representative. At the same time, some organizations are able to allocate materials to batches while they are still in the same storage place for business planning, then when they are physically moved, the paperwork associated with the material is changed to reflect its physical status. These documentation requirements create an audit trail that can be traced at a later date when necessary and are subject to regulatory enforcement. In a facility that produces multiple products and hundreds of

[b] In Research, it is frequently the job of the formulator to generate the batch record for development batches. Your first goal as a formulator can be to study other batch records to see how they are constructed.

[c] When planning a study, be sure to keep all receipts for materials and leave plenty of time for them to arrive after an order has been placed.

batches a year, it is easy to see that materials handling is a major activity and not always something that occurs within a short time of initiation.[d]

Engineering and Information Technology

The physical facility, equipment, and software are the responsibility of a maintenance and/or engineering and/or IT department, which may be separate or one depending on the size of the company. Facilities and equipment are important parts of pharmaceutical production. They must be maintained and documentation kept with the same amount of effort and control as the drug product and quality testing equipment.[e]

Production of Drug Product Dosage Form

The batch production record consists of several parts that are controlled by the company's quality system and required by regulations. It usually has a section demonstrating the cleanliness of the manufacturing module and equipment followed by documentation of release by the quality department. It has a section for documentation of the dispensed materials, sometimes called a Bill of Materials. It has step-by-step directions on exactly how the raw materials are to be processed and stored. Each step must be accomplished by the operator, who must sign and date for each step, and somehow verified by a second person, whether it is another operator, a quality representative who works with the operator, or a supervisor. Some plants use automated processing for some or all steps as described later in this chapter. Individual steps along with other manufacturing procedures such as cleaning and equipment operation can also be dictated by standard operating procedures that are separate from the batch record.

At the completion of the batch production, the supervisor must review the batch and certify that all steps are complete. The supervisor, or sometimes someone from the quality department, must calculate the yield of product from the batch. If the yield is below a certain preset level, usually 90%, a quality investigation must be generated to identify the source of loss. This is because consistent yield is a leading indicator of reproducibility and for business reasons, low yields are costly to the company.[f]

Packaging

Finished product is then sent to a finished product warehouse to wait in line for packaging. The packaging order can either be part of the product production order or separate. A packaging Bill of Materials, packaging instructions, and yield calculations are also required. In order to avoid mislabeling, the room must be scrupulously inspected by the quality department, usually while the equipment is disassembled. The equipment is then assembled in time for the arrival of the product and packaging materials. Strict count of bottles and labels is kept in order to avoid mislabeling. The labels are numbered on the

[d] Maintenance of all documentation in an orderly manner will create the audit trail as study progresses. Do not wait until the end of a 2-year study to get your records in order.

[e] Good relationships with engineering, maintenance, and IT colleagues is paramount to your success as a pharmaceutical scientist, whether you work in a laboratory or in process.

[f] When documentation is incomplete, it can hold up progress of the batch to release and interfere with the supply chain or timeliness of regulatory submissions. As such, an important part of a scientist's job is to make sure batch records are complete.

back of the label for audit trail purposes. Because of its intricate mechanical and computing intensive nature, packaging equipment may require frequent adjustments to maintain high throughput, and engineers are frequently on stand by.

Once packaging is complete, the material is stored in a non-release finished product warehouse (quarantine) and samples dictated in advance are sent to the quality department for testing. Once the testing is complete and passes, the material can be moved to a released product warehouse and staged for shipping. Product cannot be moved from the quarantine area until the quality release is finished and found to be acceptable.[g]

Validation

Phase III of the New Drug Application (NDA) process is the start of large-scale clinical trials for efficacy. At this point, the pharmaceutical organization begins to verify that the process will produce the same product and quality every time it is repeated. Equipment, software, and facilities verification are also part of this responsibility. Depending on the size of the pharmaceutical company, the department that developed the formulation and/or process may perform what is called the Technology Transfer to manufacturing, or there may be a separate department. After the NDA is filed, the process must be fully validated in the manufacturing facility. The Manufacturing operation will have a team that accomplishes Validation (usually working in unison with the research group), whether it is part of the Technical Services or Quality Department, or a stand-alone Validation group. Validation is the collection of data to provide a degree of certainty that a particular set of raw materials, equipment, and processes will produce the same product time after time. What type and how much data is required to attain what degree of certainty is a matter of scientific, theoretical, and experiential (historical) perspectives. When validation became a regulatory requirement, the production of three batches meeting specifications was considered to satisfactory. With the introduction of electronic data collection, analysis, and control, the field of validation will further evolve, as discussed later in this chapter.[h]

Quality

The Quality Department is involved in all aspects of manufacturing, from the installation of facilities and equipment, to the ordering and receipt, and use of raw materials, to production, packaging, testing, and shipping. The Quality department is responsible for Quality Systems throughout the entire organization. In earlier times, the Quality function was a matter of "Quality Control," which meant testing to specifications and release of the product. In the past 10–15 years, the Quality Assurance role has evolved to one that is over and above the Quality Control role.

With respect to improvements and changes, all changes, even change of a small part on a piece of equipment, must be assessed. Changes that are considered significant are made through a process called change control. In change control, formal notification is issued to inform affected parties of the impending change, a study is performed to

[g] As packaging is frequently accomplished by another department, be sure to leave enough time for packaging. Be sure to plan the start date for a stability study after packaging is complete.

[h] Facilities, equipment, and software validation include three phases: installation qualification, operation qualification, and production qualification. If a new piece of equipment is ordered, you will need to qualify the data produced by the machine before you start a study using it. You will need to leave enough time for the qualification stage(s) necessary to be completed.

document the suitability of the change, and some type of documentation such as a report that may be on a preprinted form must be issued to document the change. Change controls are recorded in some type of change control log open to regulatory inspection. Currently, major changes to a product involving excipients and processing steps frequently require regulatory review and approval prior to implementing the change.

When, for some reason, a manufacturing step does not go as planned or a laboratory test does not give the expected answer, an investigation is required. Investigations are conducted by the Quality Department and include the participation of any department that was involved with the unexpected result. All investigations are documented in an investigation log that is open to regulatory inspection.[i]

Regulatory Affairs

The Regulatory Department is responsible for filing and maintaining required documents with regulatory agencies. Along with Quality and Manufacturing Management, the Regulatory Affairs Department is responsible for insuring regulatory compliance. In the United States, the FDA is responsible for providing public safety with respect to drugs. Other major regulatory agencies include The European Agency for Evaluation of Medical Products, The Japanese Ministry of Health Labor and Welfare, and The Australian Drug Evaluation committee. Smaller countries have their own regulatory agencies as well. International organizations that coordinate the efforts of the individual agencies include, but are not limited to, the International Conference on Harmonization (ICH), the World Health Organization, and the European Union (EU).

Regulatory agencies have traditionally used two main ways of enforcing compliance to standards. One is through the use approvals to manufacture, whether it is for a clinical trial or marketed product and in the form of the New Drug Application or an Annual Review, or Facilities Inspection. The other is through the use of standardized test methods listed in compendia such as the United States Pharmacopeia (USP), National Formulary, the Japanese Pharmacopeia, and the European Pharmacopeias. Methods listed in these references are referred to as compendial methods.

Regulatory inspections can either be to examine the site for compliance to regulatory requirements (Good Manufacturing Practices), to inspect a site prior to approval of a new product, or to investigate product failures. When a routine Good Manufacturing Practices (GMP) inspection occurs or product failure inspection occurs, the Change Control and Investigation logs are of central importance.[j]

TECHNOLOGICAL INTEGRATION OF MANUFACTURING FUNCTIONS

From the sequence of events discussed above, one can readily see the main departments that carry out the production part of a manufacturing organization. Usually, they are: Materials, Shipping and Receiving, Production Planning, Production, Engineering,

[i] The Quality Department controls parts and materials in the company through the issuance of part numbers. It controls standard operating procedures through SOP numbers. Be sure to work up front with the Quality Department to get batch records, part numbers, and SOPs issued in advance of when you will need them.

[j] Sometimes particular companies have a sensitivity about the way studies are conducted because of a past regulatory action. Be sure to find out ahead of time if there will be any preferences with the way a study is planned at your company.

Maintenance, Packaging, Quality Control (for testing), Quality Assurance (for systems), Validation, Regulatory, Supply Chain, and Technical Services. Information Technology is now an integral part of manufacturing organizations, although the degree of daily involvement in batch production is dependent upon the company. The corporate functions of Human Resources, Accounting, Safety, Finance, Security, and General Management oversee the structural, money, and people management of the company.

The workflow in pharmaceutical manufacturing is driven by the issuance of a batch record. As the product progresses through the various manufacturing stages, all of the manufacturing and testing records as well as material movements, room clearances, equipment clearances, and management reviews are kept in the batch record. The batch record then contains a complete history of the drug product by the time it is released for shipment. Almost all of the departments listed above enter data and have a signature on the batch record.

All of the manufacturing functions, documentation, and interactions between departments can quickly lead to complex relationships. Once a step is taken by one department, several other departments are automatically staged to perform their part. All of these functions occur simultaneously, 24 hours a day in some organizations.

The size of the organization adds to the complexity. A small manufacturing organization might have one manufacturing site with just 10 products with three dosage strengths, each with three different package configurations, which equates to 90 individual stock keeping units (SKUs), for only 10 products, and all of these SKUs are in different stages of manufacturing on any 1 day. Large manufacturing organizations may be global, have 10–50 plants worldwide, and must meet regulatory requirements of multiple government agencies.

In considering all the files and documents that go into making an audit trail for every single batch of all the different packaging configurations along with all of the stability records, it is easy to see that processing and retrieving all that information in an efficient and timely manner is a large task.

In the past, all of this paperwork was manual with the exception of some generation of electronic batch records by a few companies and secondary storage of laboratory data in laboratory information management systems (LIMS). Even with the use of electronic batch records and LIMS systems, it is necessary to retrieve the records one at a time so that putting concurrent and retrospective data together for trend analysis requires a large undertaking to perform in the absence of sophisticated data management, analysis, and reporting systems.

Process Understanding

As such, the industry has transitioned to a state where the goal is to understand processes well enough to (*i*) write a mathematical model (usually a polynomial) relating the critical process parameters (CPP) to the critical quality attributes (CQA), (*ii*) collect data throughout the process, and (*iii*) feed the data into intelligent computer systems that constantly monitor the CPP and CQA in real time.

Data collected on CQA at low and high values of the CPP during research or process improvement is used to create the multivariate mathematical equations, or models, which describe processes. Development of a product in this way, with a range of equipment settings along with raw and in-process material variances, allows the product and process to be more fully understood. We call the multivatiate mathematical "space" determined by this process the design space.

Many companies have already made considerable progress in moving their new and or older products to technology-based systems. Most companies are transitioning in some

way, with monitoring, process understanding, risk analysis, QbD, and statistical processing, as discussed later in this chapter. Many new sensors and software programs continue to be developed for in-process monitoring that can be interfaced to intelligent computer systems that analyze the data and compare it to historical data.

Scientists can prepare themselves by understanding how the CPP and CQA of products and processes that they want to create can be monitored, and how the collected data can be used in multivariate modeling to understand the entire design space around the process.

Monitoring during production, adjustment of the process to attain the desired outcome using process understanding, and storage of the data in a way that allows all of the test results to be trended over time is a way to create more efficiency in the pharmaceutical manufacturing industry. Not only does a computer screen show progress on results from a single manufacturing line, but intelligent systems can be set up to monitor business cycles as well, bringing the information to corporate level functions of accounting, finance, supply chain, and general management.

From these considerations, the pharmaceutical scientist can see that his/her interaction with manufacturing is not the simple matter that it can appear to be. Careful thought about the way formulations and processes are designed is required to support smooth operations in manufacturing.

The remainder of this chapter will focus on the scientific aspects of pharmaceutical manufacturing and the role the pharmaceutical scientist can play to improve manufacturing.

Process Endpoints

Historically, manufacturing unit operations are typically concluded after predefined periods of time. For example, granulation processes are concluded after reaching a time endpoint. Compressing and milling unit operations are set to prespecified speeds. Along with many other industries, the pharmaceutical industry recognizes that not only are the endpoints of a manufacturing process important, but the path or trajectory taken to get to the endpoint can also be important in controlling product quality (1–3). Recent changes embrace the ability to stop a process when a certain quality is attained rather than after a preset time. This ability is based on the use of in-line sensors, intelligent interfaces, and information technology.

Regulatory Support

A number of positive changes in the regulatory environment are supporting the use of technology. From a U.S. perspective, the release of FDA's Process Analytical Technology (PAT) Guidance (Guidance for Industry: PAT-A Framework for Innovative Pharmaceutical Development, Manufacturing, and quality Assurance at www. Fda.gov/cder/guidance/index) was instrumental. From the ICH's perspective, the release of ICH Q8, Q9, and Q10 (www.ich.org) documents which cover product QbD, risk management and quality systems, respectively, was also instrumental. The U.S.FDA, the EU, and the Japanese Ministry of Health, Labour and Welfare along with regulatory bodies from many smaller governments and organizations have encouraged the use of risk- and science-based methodologies. The industry itself is utilizing the potential of more efficient processes by:

1. further characterizing raw materials for functional attributes,
2. QbD,
3. utilizing advanced analytics,

4. data management and acquisition,
5. modernization of the manufacturing process through real-time process control,
6. modernization of the manufacturing process through continuous manufacturing processes,
7. risk management,
8. systems to support these concepts.

In the remainder of this chapter, we will discuss the evolution of the industry and present a future state which helps to ensure the industry achieves its key objectives.

THE USEFULNESS OF MANUFACTURING DATA

Pharmaceutical manufacturing in a traditional batch mode begins with fixed quantities of materials charged into equipment. Each slug of material is pushed through each unit operation within hours; however, the time between each unit operation can range from minutes to weeks or months depending upon the schedule of the manufacturing operation. Traditionally, a few in-process tests were used for unit operations to ensure that the step was completed adequately. For example, loss on drying measurements are usually taken after each drying step to ensure that the moisture content is within an acceptable range. Most specification testing was completed at the end of the stage or on the finished product.

With its most modern technology, a manufacturing facility can test at the end of a stage as well as while the process is running at a variety of locations. Data can be collected at any place that a sensor can be installed with electronic archiving of the data over time. This data can be used in combination with statistics software packages to transform the data into multidimensional descriptions of the process. Frequently, multidimensional graphs can be generated that enable scientists to visualize the design space.

Commercial Product Manufacturing

In order to demonstrate the utility of online data collection, QbD, and process understanding, consider a typical solid oral drug product manufacturing process, in which powders are blended then granulated, compressed, and coated.

Raw materials arrive to the manufacturing facility and are minimally tested for identity since the organization will, whenever possible, test only a few lots per year. The vendors' analytical methods have been validated or verified to the pharmaceutical company's satisfaction and the pharmaceutical manufacturer accepts the material based on the vendors' certificate of analysis. Once any required testing has been completed (typically in weeks), the material is sent to the warehouse until dispensing requires its use for product. The materials are stored in the warehouse until needed for manufacture.[k]

The required amount of material is weighed according to the manufacturing directions. The materials are staged in a suitable warehouse pending readiness of the manufacturing area responsible for starting the process. The material may be in a staged

[k] Compendial tests were originally designed for chemical, not physical properties and the specifications of the compendial test could be quite large. As such, companies began adding additional tests that affect functionality, such as particle size analysis. In addition, the specification range for a compendial test could be larger than the acceptable range for the product. Without process understanding, this could lead to unexpected movement of product properties within the compendial range (see the PAT example).

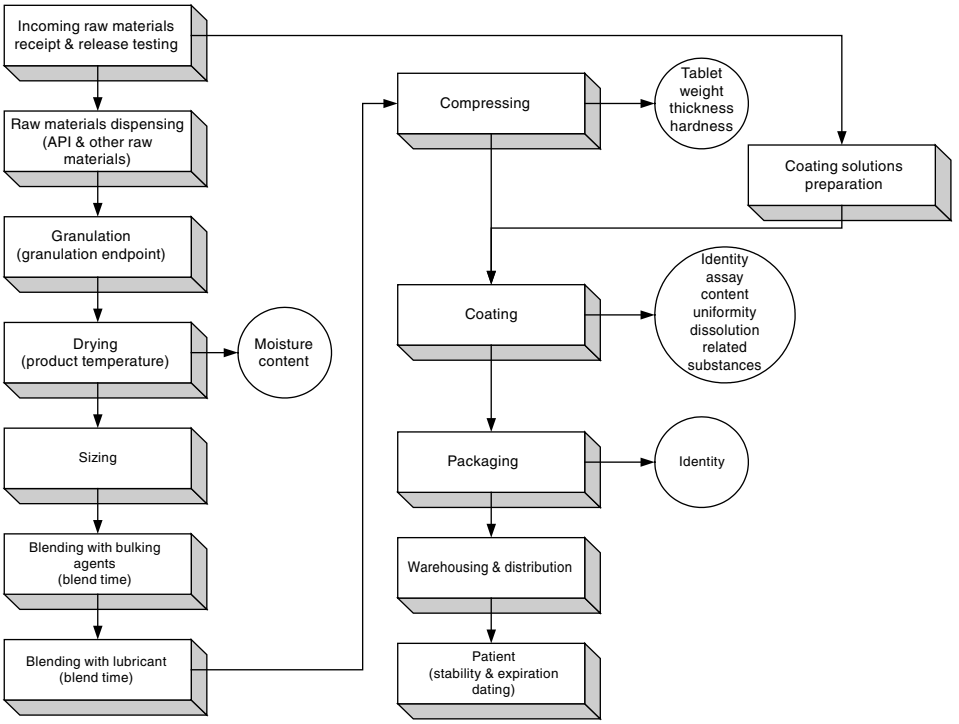


FIGURE 1 Typical manufacturing process flow diagram for a solid oral dosage form.

area for a few hours to many days before consumption by the next manufacturing step. In the example to be used, the granulation area is responsible for initiating the process. Please note that the expiration date for this product is typically defined as the first day that the drug substance is consumed or modified in any way.[l]

In our example, the drug is placed into a high shear mixer and other raw materials added in order to impart the required qualities for the dosage form (diluents, glidants, binders, and compressing aids). The high shear mixer is started and allowed to run for a predefined period of time at a suitable speed. After the materials are mixed for the required time, the granulating solution (typically water) is added either through a pipe or spray system. The impellers continue to turn until all the water is added. Once water addition is complete, the impellers are turned to high speed and allowed to run for a predefined period of time. The wet granulation is sent immediately to be dried in a fluid-bed dryer as wet material may compact on itself and make it impossible to fluidize, and even develop microbial growth. The quality of the granulation could change if allowed to sit wet for prolonged periods of time.[m]

---

[l] When designing development studies, be sure to take this date into consideration when planning stability studies.

[m] Granulation is frequently used for several reasons, the main one being that granulated material flows through compression and encapsulation equipment better. Although it is not always possible, creation of a direct compression process eliminates a step from the process. Eliminating steps from the process creates a more efficient process. The more steps in a process the more it costs the company to make the product because every step requires more space, more people, and more documentation.

Once granulation is complete, the product is placed into a fluid-bed dryer. Enough dry air is introduced into the dryer in order to fluidize the material in the dryer and achieve rapid drying. After a predefined temperature is reached, the dryer is stopped and one or more samples are retrieved for moisture testing. If the test results are acceptable, the material proceeds to the next unit operation. If the test results are too high, the material is sent back to the dryer where the process is repeated. The test method is typically a Loss on Drying method which is not specific to water itself. The dried granulation is then milled through a high speed mill equipped with a fixed screen size in order to reduce the particle size to the desired range.

The milled granulation is introduced into a diffusion mixer and additional materials added as appropriate. In this example, a glidant is added and mixed in first with a time endpoint, and then a lubricant, typically magnesium stearate, is added in order to impart the right lubricity to the granulation. At this point, material could remain staged for the next unit operation for hours to months before compressing.

The lubricated granulation is then sent to the compressing area so that tablets can be made. The tablet press is setup and the tablets made from the granulation while using a constant press speed. Adjustments are made in order to ensure that tablet weight, thickness, and hardness are acceptable, and are confirmed by the Quality Assurance (QA) department. At this point, material could remain staged for the next unit operation for hours to months before coating.

The compressed tablets are sent to the coating area. Solutions or suspensions are prepared and sprayed onto the tablets. Airflow rates, temperatures, spray rate, and atomization air pressure are kept within predetermined ranges to ensure that the coating quality is adequate. Inspection of the final, coated tablets by the QA department assures acceptability of the appearance of the tablets. At this point, material could remain staged for the next unit operation for days to months before packaging.

The packaging operation can usually be executed within hours; however, the setup of the equipment itself can be quite complicated and time consuming. Whenever possible, it is highly desirable to keep this equipment running by packaging many lots of the same product at the same time, thereby reducing the number of changeovers for other products. In this example, tablets are placed into High Density Polyenthylene (HDPE) bottles, a label with an appropriate expiration date applied, a suitable cap added and closed to the correct torque to ensure that it is properly closed, and bottles then sent to a cartoner where a package insert is also added. Finally, several bottles are placed into a larger corrugate box which is then sealed for shipment. These boxes are then placed on a crate, shrink-wrapped in plastic, and stored in a warehouse until final product testing is complete.

Once final product testing is complete and found to be acceptable, the product is released by the quality function and the product shipped to a distribution warehouse that is strategically located around the world to meet the companies supply chain distribution needs, or to a customer directly.

The QA department ensures through inspection that the process was properly executed per the manufacturing directions and that all in-process and final product release results are acceptable before proceeding.

Measures taken during any unit operation are typically not utilized to make adjustments to the downstream manufacturing steps. For example, moisture determinations are not used for adjusting blending or tableting operations.

Granularity of Data

Modern pharmaceutical plants are equipped with a significant amount of electronics and measuring devices. Computers are installed for nearly every piece of operating equipment.

The amount of real-time data which can be captured by this equipment is significant and poses a challenge to the pharmaceutical organization. What is the right data to capture, through instrumentation or otherwise? How much data should be collected and stored? How frequently should we capture this data? What shall the pharmaceutical organization do with this data? and Can an organization make quality and product release decisions with real-time data, in real-time?

The granularity or detail contained within the data depends on its source (Fig. 2). Time-series data can be analyzed for each unit operation. Key results and findings from the real-time data can be analyzed along with batch information, which in turn can give an overall view of the product quality. The data can then be further analyzed for trends across product lines within the plant, for trends between plants manufacturing the same product in different plants, and other analysis (4).

This data can be utilized for generating process understanding. This understanding can be obtained through simple methods such as trending, graphing, or process capability analyses to sophisticated methodologies such multivariate data analysis and neural networks (5–7,103).

The source of the data is not limited to equipment data but can also reflect incoming raw material property information contained with LIMS, electronic batch data, any other potentially useful data stores including financial and plant management systems (7).

If appropriate, raw material properties should be used to not only predict downstream operations but also to make adjustments to the manufacturing process as a result of those properties. This is known as feed-forward control. Data generated during manufacturing should be utilized to make adjustments to the process for the next batch which is about to be processed. For example, the amount of granulating water could be adjusted so that the process trajectory for granulation (rate of power/torque produced over time and final power/torque) is constant for each granulation run. In this way, product variability from within and between lots is minimized (104).

Similarly, conditions for the drying process are adjusted to reflect the changes in moisture of the granulation. Blending conditions are calculated to achieve a uniform product, understanding that the blending process itself is influenced by not only moisture but also other factors such as granulation particle size and shape (96).

Environmental Conditions

Most industries have a great appreciation of the impact of environmental conditions on a process. Published data suggest that in fact environmental conditions do play a significant role. Though the authors did not determine the reason for the impact, Stryczek et al. (8) found that outside temperature and/or humidity have direct impact on processing and dissolution. The authors investigated approximately 140 process parameters including raw materials properties, processing conditions, and environmental conditions, and their impact on dissolution. The authors concluded that ambient humidity/temperature is critically important. As the temperature and humidity in the tropics track very well with each other, one can choose either variable for process monitoring. Keeping in mind that some of the key raw materials are shipped to the tropics via boat, and that many of these raw materials are stored in polyethylene and/or paper bags and subsequently stored on the islands in vendor warehouses which are not temperature or humidity controlled, it would be expected that storage time, environmental temperature and humidity would change the moisture levels of the raw materials before they reach the humidity and temperature controlled manufacturing facility. Upon receipt, these materials will acclimate and change to their new lower temperature and lower humidity environment. This may make

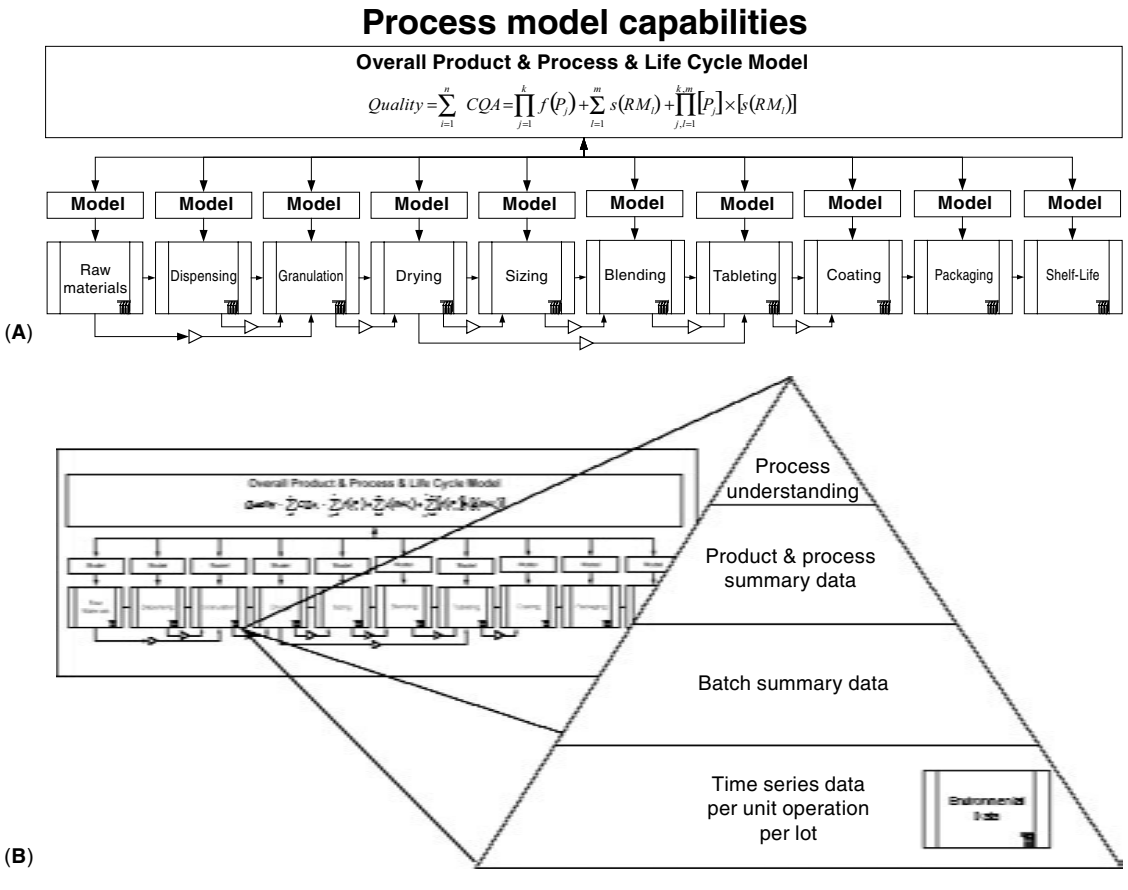


FIGURE 2 Process flow diagram for a product utilizing modern process control and measurement scheme: (**A**) manufacturing operations flow; (**B**) granularity of data available at the dispensing stage.

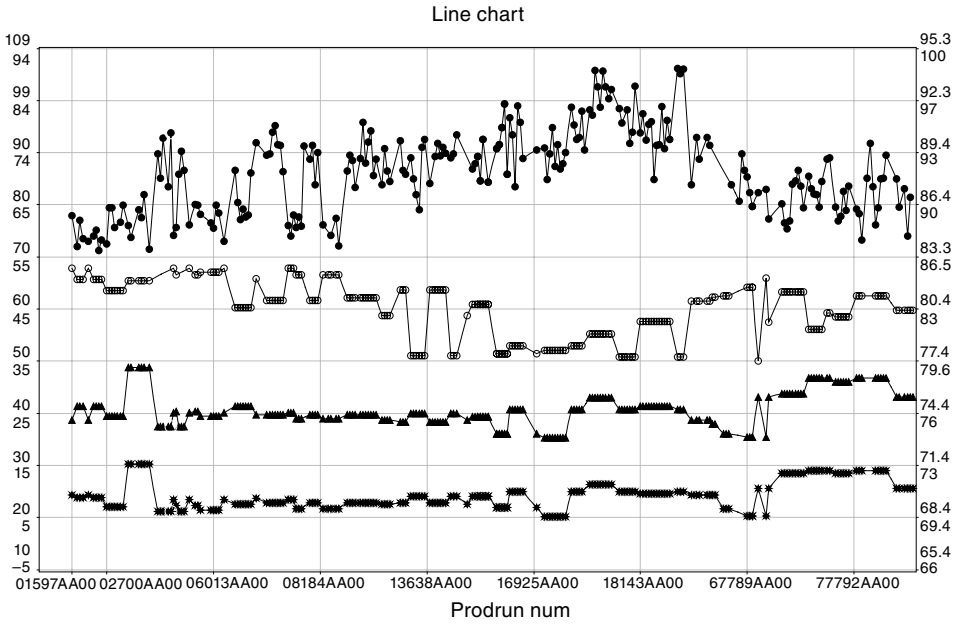


FIGURE 3 Changes in raw material properties over several hundred lots of product. Red—product dissolution rate. Blue, black, and green—raw material properties.

it challenging for the pharmaceutical firm to control the manufacturing process (and final product quality) if it does not have an adequate understanding off these effects.[n]

Troubleshooting Manufacturing Operations

Supporting manufacturing operations can be quite challenging in terms of constraints (i.e., time, capital, money, and manpower). There are many methodologies which can be used for designing and optimizing manufacturing operations (9–18). One of the keys to becoming proficient at troubleshooting manufacturing operations is to be able to quickly diagnose the source of the problem. Where the issue manifests itself is not necessarily where the source of the problem occurred. For example, the data in Figure 3 shows how the results of dissolution testing for a modified release product. One can readily see that the dissolution rate increased during a period of time. The dissolution results for this particular product were not available until after all manufacturing had been completed. This is unfortunate as three raw material properties for one excipient shifted. A test, such as near infrared (NIR), which explores multiple aspects of this raw material may have assisted in detecting an issue with the raw material before it was consumed. In this way, the organization would have minimally known that there is something fundamentally different with the raw materials before they were consumed in the product. Had this test been in place, the company would have saved several million dollars in lost inventory as

[n] Cold, dry weather can also affect the product. Always take into account the weather effects from outside the facility on the temperature and humidity inside the manufacturing module. Make sure the temperature and humidity monitors are in place, adequately maintained, and documented prior to proceeding to a manufacturing experiment. Be sure you also understand how your vendors are storing their materials. Their methods could affect your product as well.

the raw material and product no longer met specifications. When corrections were made, a return to near baseline for drug release was achieved.

As scalable models are typically not available for many pharmaceutical operations, it is often challenging to troubleshoot a manufacturing process in the pilot plant, and to then return to full-scale manufacturing. One cannot expect to be successful without some additional exploratory trials to further define optimum manufacturing conditions at full scale.

Stryczek et al. (8) published data for multiple manufacturing sites. In this example, commercial operations had been in effect for one plant for several years. The company wanted to transfer the process to a distant facility. As is common, a direct transfer to similar (but not identical) equipment is not necessarily straightforward. For example, the granulators, tablet presses and tablet coaters at each facility were made by different vendors. The authors executed experimental designs to define the granulating and compression conditions to achieve optimal dissolution rates at the new commercial site. The scientists then used the results obtained from these studies (i.e., mathematical algorithms) and applied them to the original manufacturing facility. They were able to improve the level of understanding at the original commercial facility with the new understanding obtained from these experiments.

It is not unusual for manufacturing operations to experience some sort of difficulty. For example, during coating of one product, scientists were notified of manufacturing defects for coated tablets. Previous coating runs ran rather well, however, the subsequent coating run yielded an unacceptable physical appearance. As two coaters were contained within the same manufacturing suite, scientists could compare directly to the two coaters and noted no issues with that unit. When the scientists retrieved the raw data from that coating operation, they discovered that the coating temperature was fluctuating in a sinus wave (Fig. 4). Availability of the raw data from each coating run was key in determining what happened. If this equipment were not fit with these sensors, an investigation into the matter would likely have yielded no conclusive results and subsequent batches would have also had the same difficulties. In this way, future coating runs were spared and dollars were saved. Manufacturing operations quickly resumed after the accurate diagnosis.

Quality by Design

Quality by design (19) is defined as "Designing and developing a product and a manufacturing process that ensures that the product consistently achieves the pre-determined quality characteristics." It is holistic in scope in that it encompasses the entire life-cycle of a product including its initial concept phases through development, commercialization, and eventual removal from the market. Table 1 is one adaptation to the pharmaceutical industry of "Juran on QbD"(20) where activities and outputs for QbD were identified. In this environment, marketing identifies key opportunities for development through rigorous market research. From this information, key patient populations are identified for each potential indication. For each of these indications, the needs of the patient are clearly defined (i.e., reduce risk of heart disease caused by high cholesterol). These needs are then transformed into product quality attributes from which a dosage form can be designed (e.g., reduce Low Density Lipoproteins (LDL) cholesterol with compound A by 30%; with an immediate release, solid, oral dosage form). From these criteria, a dosage form can be produced using a process which has been optimized using advanced analytical techniques. Since appropriate product attributes are correlated to performance (e.g., dissolution to LDL levels), in-process controls can be put in place throughout the process, guaranteeing that an optimal process has been used. This will ensure that the

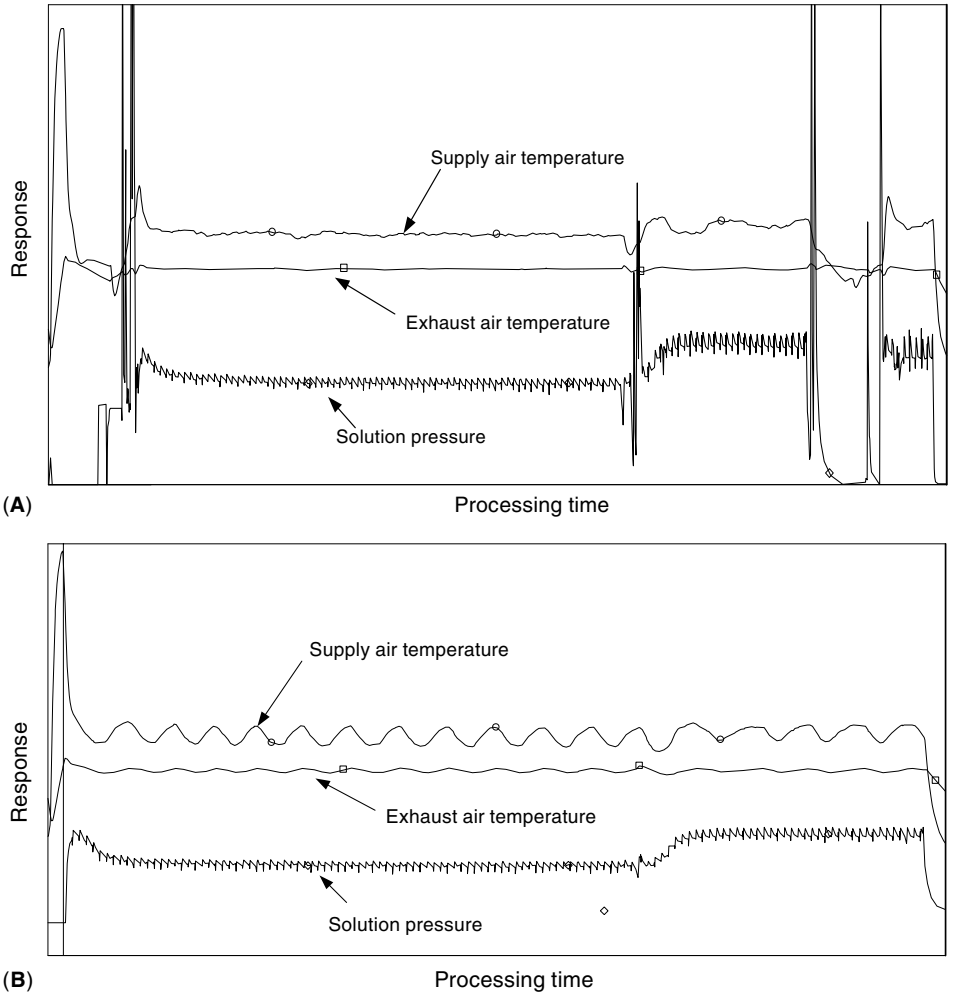


FIGURE 4 Real-time coating conditions: (**A**) typical coating run; (**B**) problematic coating run showing fluctuation in the temperature control.

desired final quality attributes of the dosage form (i.e., dissolution) are achieved as they were controlled throughout the process. The process is then scaled-up, and commercialized. Validation in this environment is actually continuous verification of key quality attributes on each batch that are meaningful to the performance of the dosage form in the patient. Continuous improvement throughout the life-cycle occurs, thereby constantly updating and reducing the risk profile of the product.

Process Development and Monitoring Using Quality by Design

Traditional process development was considerably more limited in its ability to properly define processes. This is due to many reasons including the following:

1. continuing evolution of the understanding for first principles;
2. limitations in sensor technologies in terms of new measurement devices;
3. limitations in sensor technologies relative to data collection rates;

TABLE 1 Example Activities, Outputs and Responsibilities for Quality by Design from Inception through Commercialization

Activities	Outputs	Responsibility
Identification of potential patient populations	List of patient populations by indication	Marketing
Determine patients' needs	List of patients' needs	Marketing and therapeutic area
Develop pharmaceutical quality attributes	Dosage form size, shape, etc.	Pharmaceutical R&D
Develop process features	Identification of appropriate unit operations for process	Pharmaceutical R&D
In-process controls, PAT, product specifications	Commercial manufacture start-up	Pharmaceutical R&D, quality and operations
Process validation and process capability	Routine commercial manufacture	Pharmaceutical R&D, quality & operations
Continuous improvement and risk management	Improved processes & products	Pharmaceutical R&D[a], quality & operations

[a]R&D involvement in continuous improvement depends on the company.

4. development of statistical methods including but not limited to design of experiments, Taguchi methods, and robust engineering;
5. limitations in computing processor speed.

Pharmaceutical scientists typically would create a formulation based on their education and experience. Though often close to their ideal, minor tweaks in the formulations and supporting processes were made using one factor at a time methods. As is well documented, though improvements were made, the optimum was often not reached.

The tool set available to today's scientist is quite varied and powerful. A scientist will first identify which process parameters could have an impact on final product quality. Figures 5 and 6 provide examples of a Fishbone diagram for a wet granulation process and design space for a dry granulation process, respectively. Key in product development and troubleshooting efforts is the design of the experiments up front. Not only is it important to understand what the experimental design will deliver, it is equally important to understand what is does not deliver. Planning ahead and anticipating processing events is the key to rapid development. Under the best of circumstances, the first few times a product is manufactured by R&D, it is not always clear if the process conditions and raw materials are close enough to be process capable. For example, can a granulation actually be produced under these conditions? A smart scientist will not only plan for the number of experimental runs detailed by his statistical design, he will also allow for additional runs, if possible, in order to further explore things he learns as he executes the experiments. He may confirm a previous trial which appears to produce enhanced quality or processability, or he may choose to investigate an area which is slightly beyond the experimental design space because the data he has collected in the first few trials point him in that direction. This is especially useful on intermediate and large scale where the time to get into the facilities is typically quite long between experimental campaigns. The time savings are not only substantial, but it may also allow the scientist to salvage an entire campaign because of an inadequate initial design.

In other situations, the scientist may plan for a certain experimental design but may leave the actual conditions unspecified up front. That is, he will attempt to manufacture

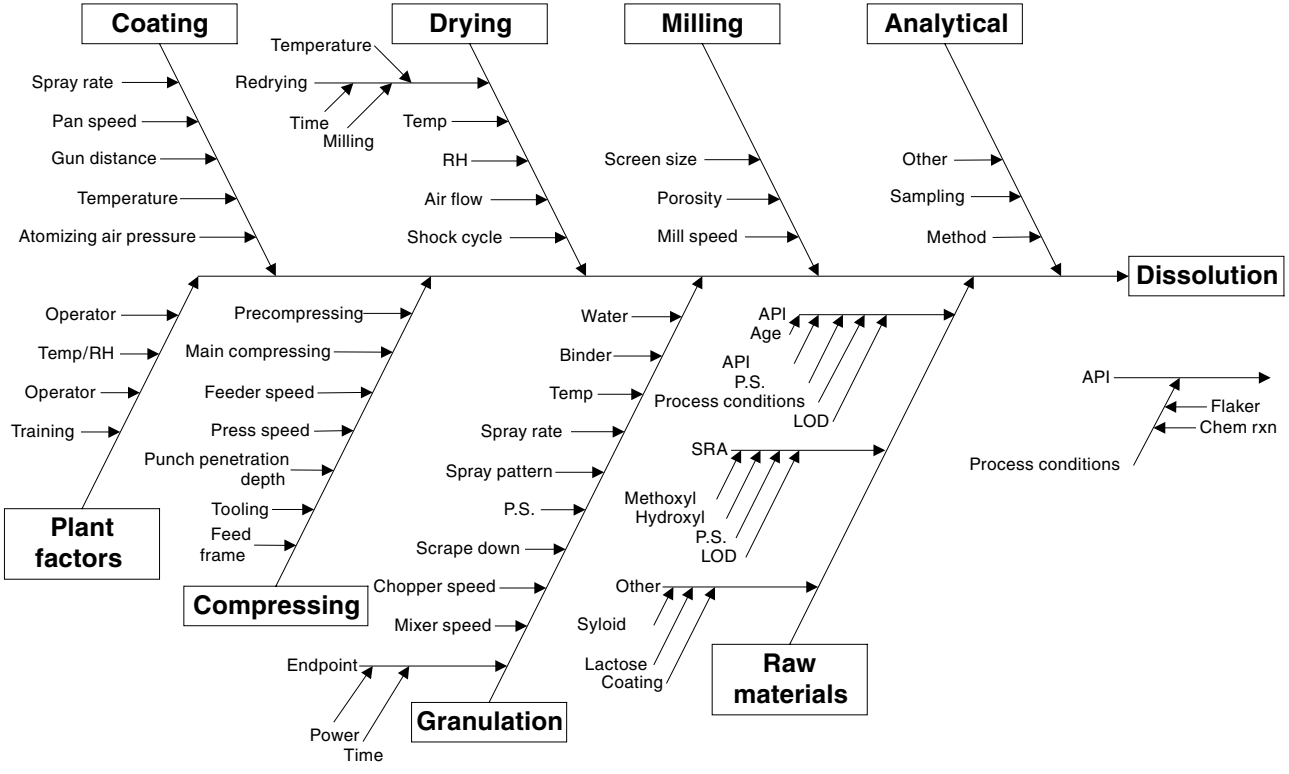


FIGURE 5 Fishbone (Ishikawa) diagram for dissolution.

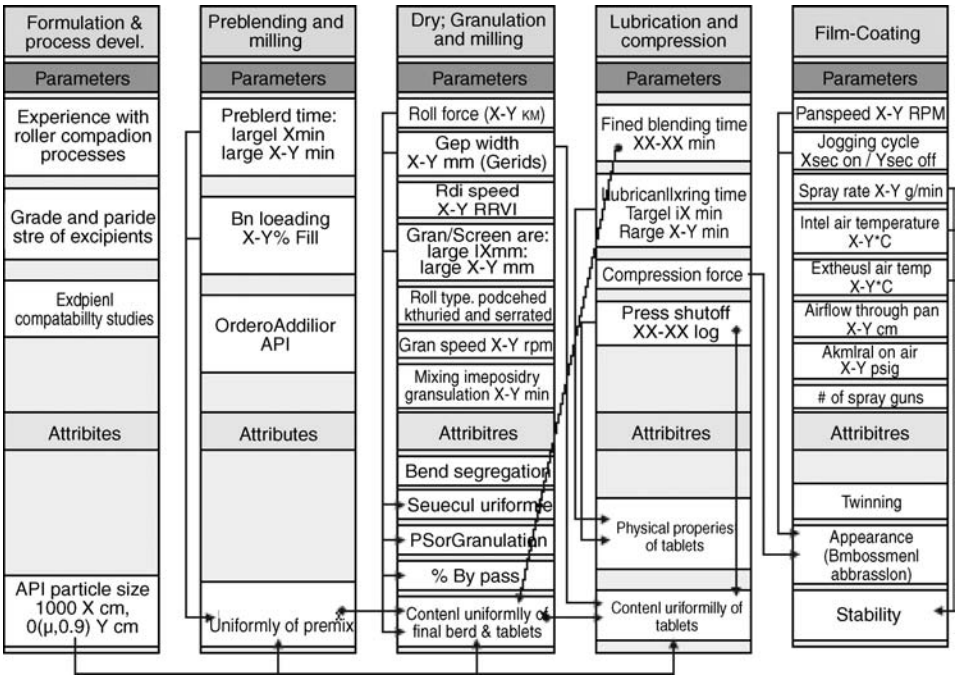


FIGURE 6 Dry granulation design space. *Source*: From Ref. 21.

the initial batch under target conditions. If the batch fails to process properly, we will then redesign his experiment and compensate for the additional understanding obtained from the first trial.

Process Development and Monitoring with Quality by Design and Process Analytical Technology

Today's scientist also has a significant amount of on- or at-line analytical support. NIR, Raman, acoustic, and other measurement methods are now commercially available (22–36) and can be mounted to manufacturing equipment so that the sensor beam goes through a window to the sample without coming in direct contact with the materials (37–55). Wireless transmission to databases provides real-time data collection and process monitoring. Figure 7 shows a corona NIR attached to a Patterson Kelley V-Blender. Figure 8 shows and overlay of NIR spectra for individual raw material, demonstrating unique peaks of interest for each raw material. Figure 9 shows the spectrum of the blend collected during a single time point. Figure 10 shows the change in concentration of individual raw materials over time. These graphs were generated by plotting the magnitude of the signal at unique wavelengths in the individual blend spectra over time for each of the raw materials. The data from these spectra can be used in combination with variables collected from other methods, including materials characterization, to develop a process model. Analysis of data from wavelengths unique to each excipient allows exploration of disposition of each excipient.

In another example using the same study, off-line NIR chemical imaging (56–65) was used to further understand blend qualities on CQA of the finished product. Figure 11 shows an example of chemical images obtained for the blend.

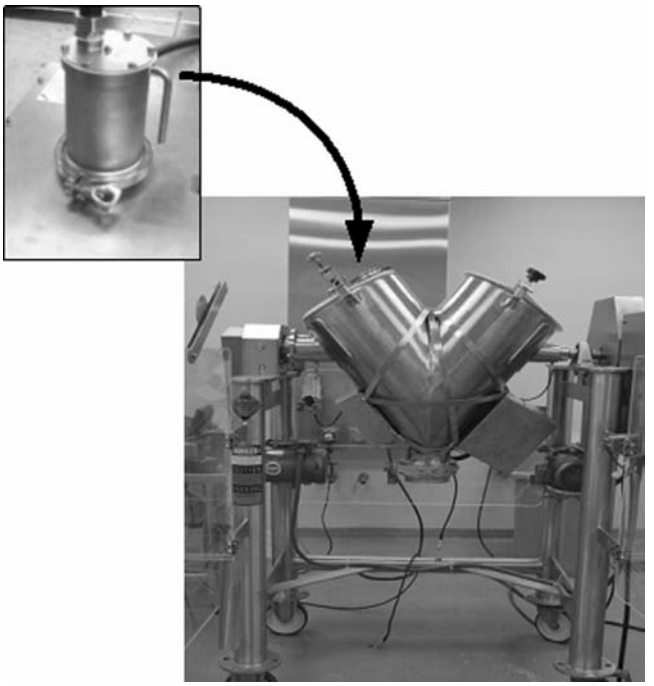


FIGURE 7 Corona NIR and wireless data collector attached to Patterson Kelley V-Blender. The detector at the left is mounted to the sapphire window on the hatch. Data is collected when the hatch is down and powder is on the window. A trigger stops data collection when the hatch is up. *Abbreviation*: NIR, near infrared.

Software can be used to analyze the size and number of the colored domains in the images (66). The number or size of the domains for a particular ingredient can then be plotted against a CQA of interest. In the current example, the design variables were particle size (unmilled, milled, and milled twice) and blend time (15, 45, or 75 minutes). Figure 12 shows a scatter plot that reveals clusters of data. The spacing of these clusters was a reflection of the design space with the blue and purple cluster representing data from the milled API at a shorter blend time, the middle of the cluster being the 45-minute milled material and the red and green cluster representing the unmilled API.

The data from the domain analysis can also be plotted as a function of the study inputs. Figure 13 shows a response surface analysis where the blend time and the API particle size were used as inputs (*X*- and *Z*-axes) with the resultant API domain size on the the *Y*-axis.

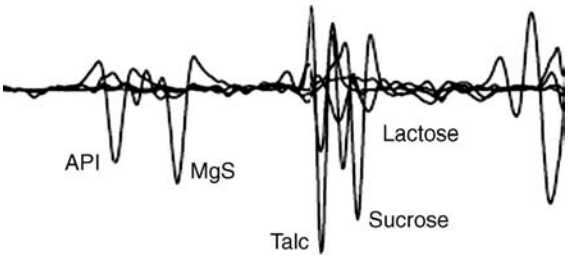


FIGURE 8 Overlay of individual raw material spectra demonstrating unique peaks of interest for each raw material.

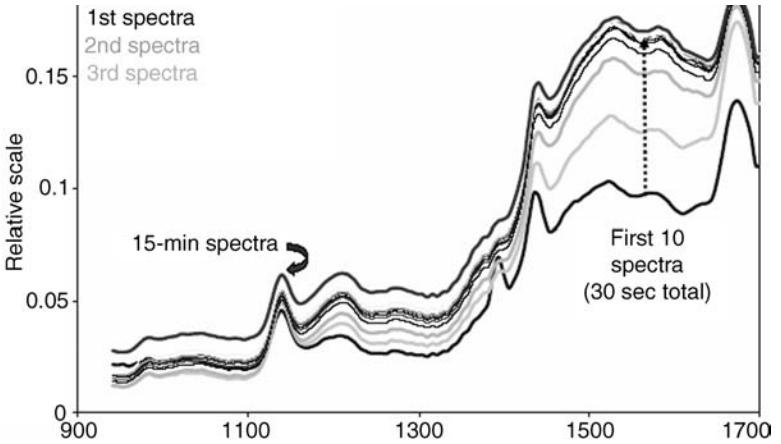


FIGURE 9 Online NIR spectrum (nm) for the blend in the PK blender. Bottom scale = time in minutes. *Abbreviations*: NIR, near infrared; PK, Patterson Kelley.

Once understood, process models can be developed for process control at the R&D and commercial scale. Many recent articles have been published in the area of process monitoring. There are few pharmaceutical papers which discuss the control aspects themselves. The pharmaceutical scientist can also look to other industries for useful process control information (67–94). In addition, recent American Society for Testing and Materials (ASTM) (94), ICH (121), and FDA publications can point the pharmaceutical scientist in the right direction.

Raw Materials Characterization

Raw materials characterization is an area where the pharmaceutical industry has a great opportunity to gain efficiencies. From the authors' experiences in commercial operations support, raw materials contribute to a significant portion of manufacturing investigations related to the drug product. In the past 10–20 years, much greater emphasis is being placed on additional functional characterization for performance in the dosage form (95) and its link to bioavailability. As previously discussed, pharmacopeial specifications are typically geared towards identity and chemical integrity testing, along with some basic physical characterization, but a stream of new on- or at-line methods continues to become available These methods can be used at-line during development to more fully understand the entire design space.

Figure 14 shows what may happen for a typical product from initial R&D development through commercialization. Usually, the development scientist is not successful in securing multiple lots of key raw materials that have a range of properties. The pharmaceutical scientist first develops a dosage form and to the best of his ability attempts to obtain raw materials with varying properties. Though, the compendial range is quite wide, his actual experience is quite narrow. For practical reasons, the pharmaceutical organization files their drug application with the compendial limits as this is an acceptable practice. During product launch, he may experience a little more raw material variability than during development but the process is still relatively well behaved. However, over time, the process continues to drift and issues start to occur. Perhaps dissolution is no longer acceptable, or tablet hardness has unexpectedly fallen off, or even the granulation endpoint can no longer be reached, etc.

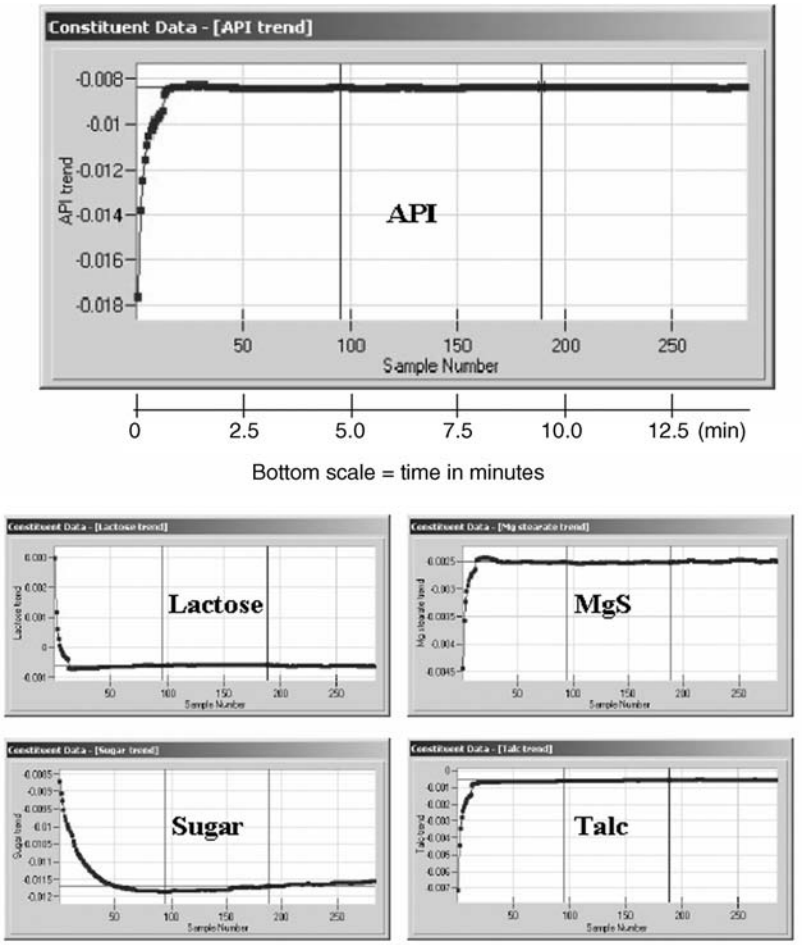


FIGURE 10 Online NIR data for a single wavelength over time. The top graph shows data for the active ingredient (API). The API was "sandwiched" between the other excipients when the blender was charged. The next four graphs show the excipients. Materials closer to the outside of the blender when charged decreased in concentration seen by the sensor while materials not close to the outside upon charging increased. Blend homogeneity was attained within 5 minutes. *Abbreviation*: NIR, near infrared.

This is a significant area of concern. Most pharmaceutical companies are not large enough to demand "special treatment" from their raw material vendors. In order to properly characterize a raw material for long-term, commercial-scale production, the pharmaceutical scientist should identify those properties which could have significant impact on product quality and manufacturability. That said, obtaining various lots of raw materials with different quality attributes requires commitment not only on the part of the pharmaceutical manufacturers, but also the raw material vendors themselves. In many if not most cases, the financial return to a raw material vendor is not justified for making "special lots of raw material" in order to allow the pharmaceutical organization the opportunity to investigate an acceptable design space. Therefore, the pharmaceutical scientist is typically limited in his attempts to properly define the design space for the raw material. He must launch the product with a narrowly defined window for potentially critical attributes of the raw material. The burden of further refining the raw material

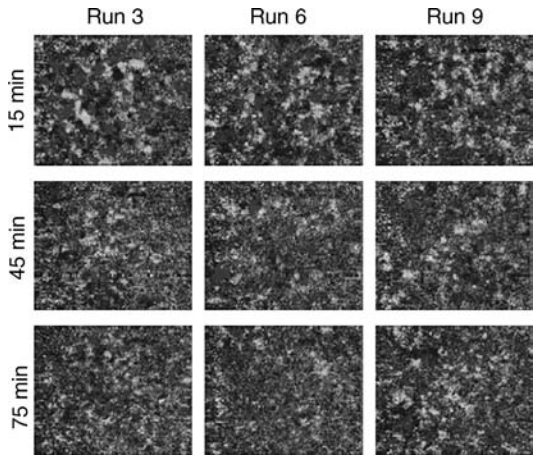


FIGURE 11 NIR chemical images of experimental blend. *Abbreviation*: NIR, near infrared.

specifications then falls on the manufacturing organization itself. Perry's Chemical Engineering Handbook list several potentially useful examples such as shear indices, compressing indices, etc. (96).

Utilizing Advanced Analytics

The advancement of sensor technology is facilitating the deployment of PAT. Equipment vendors are becoming aware of the needs of the R&D and commercial organizations, and how to properly deploy these technologies. Since the introduction of the FDA PAT Guidance (FDA Guidance for Industry, PAT: a Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance, September 2004) the industry has shifted its focus from trying to understand the implications of the guidance to implementation of its concepts.

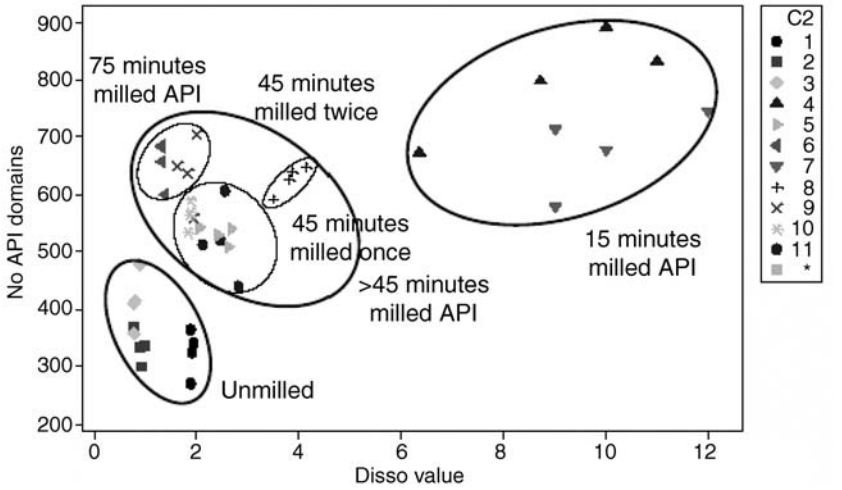


FIGURE 12 Offline chemical imaging analysis of API domain size *versus* capsule dissolution. Dissolution rate value was determined from USP dissolution testing at 30, 60, and 120 minutes and calculated from the ratio of the difference between 60 and 30 minutes, and divided by the difference between 120 and 60 minutes. Fiber optic UV dissolution data were not available for this analysis. *Abbreviations*: NIR, near infrared; API, active ingredient; USP, United States Pharmacopeia.

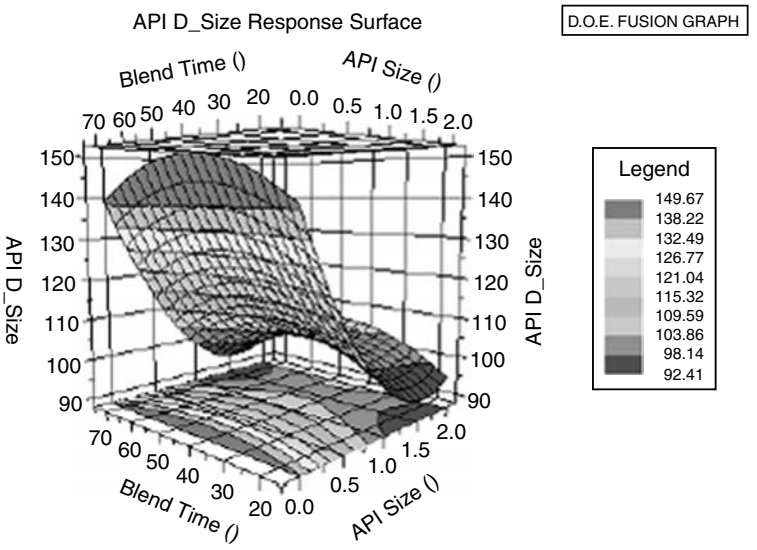


FIGURE 13 Response surface plot of active ingredient (API) particle size (z), blend time (x), and powder blend API domain size (y).

Figure 15 shows an example output of a NIR sensor mounted to a tablet coater used for a product in which the coating controls dissolution (97). The NIR senses material in front of its laser optic probe. As the coating process proceeds, the NIR senses the change in materials in front of its laser optic probe. The initial NIR scans represent the core tablets themselves and as coating proceeds, the NIR scan changes to one representing the coating materials. From this methodology, it is possible to quickly determine when an adequate amount of coating was applied. As shown in Figure 16, the NIR results can then be further correlated to dissolution results.

Contrast this to prior art which required that the coating process be stopped after a prespecified period of time. It was not always clear if the coating was properly applied or if enough coating was applied during the coating process. Coating, if not properly performed, can yield different film properties if the conditions of the process are not adequately controlled. This type of analysis can yield further insight into coating quality.

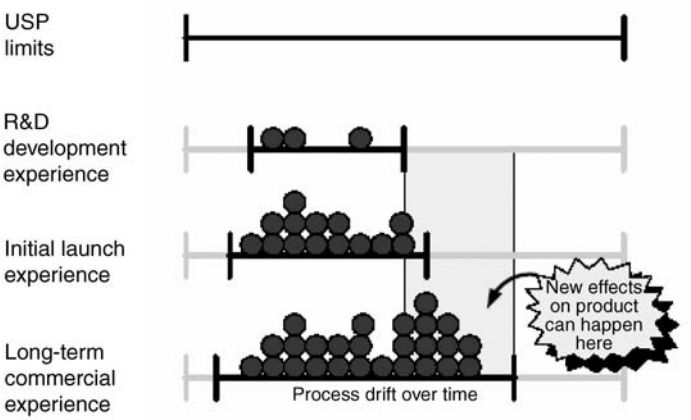


FIGURE 14 Typical drift in raw material attributes over the life of a product.

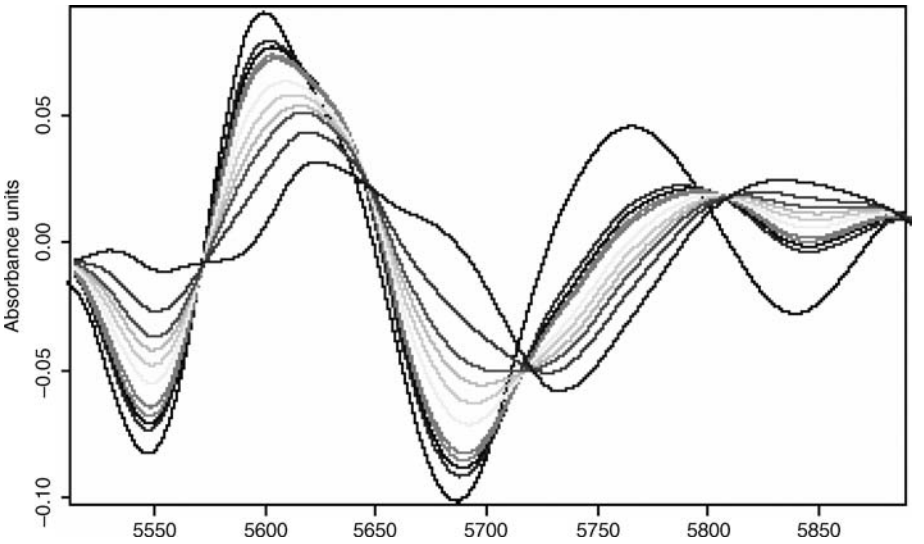


FIGURE 15 NIR results from tablet coating monitoring coating thickness. *Abbreviation*: NIR, near infrared.

The coating example is but one example. For nearly every unit operation, examples of real-time data monitoring and analysis exist. For drug substance, reaction endpoint monitoring, crystallization, and impurity monitoring have been reported (98–101). For drug product, blending, compressing, coating, milling, and roller compaction, have all been demonstrated in the literature (102–121).

Data Management and Acquisition

With the advancements in analytical data and sensor technologies, methods for collecting and managing these data sets are required. Geoffroy (7) reported an example where data collection systems were developed for batch production and laboratory results.

The authors reported how batch data, laboratory data, and user-defined data (notebook information, results, and data from outside sources) were combined into a data warehouse or data repository. As the data contained within each of these systems are linked through the lot number and other descriptive information, it is possible to analyze data for a product in a very holistic manner. That is, it is possible to analyze batch-to-batch information for multiple types of data including but not limited to processing information and operator specified information. It is also possible to trend that data against quality measures, either in-process or at release. One can analyze trends for raw materials as quantities and lot numbers are specified in the bill of materials, equipment used as equipment numbers are specified by the operator in the batch record, process parameters as either they are predefined in the batch record or specified by the operator during manufacture. In some cases, equipment usage and frequency can be monitored as the equipment can be used for multiple product lines. Equipment maintenance can be specified according to the number of times it has been used. This can also assist the QA organization in assessing the level of equipment qualification that should be performed.

Obtaining the data electronically is critical to success in that the efficiency obtained in designing such a system is huge. In the traditional manufacturing organization, batch and test data are stored on paper records. Collating the information from hundreds of lots

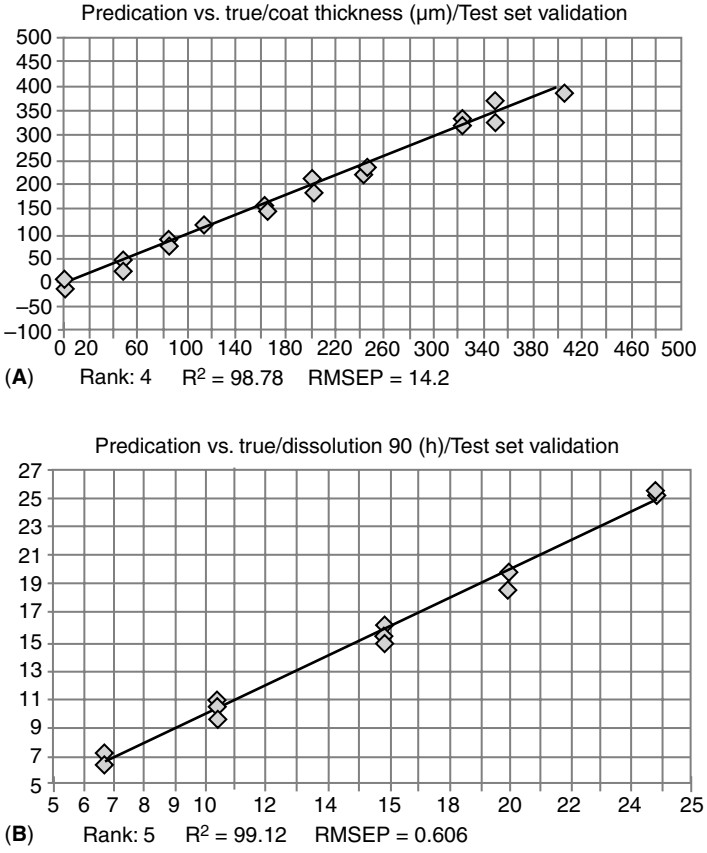


FIGURE 16 Correlation and predictability of NIR data to: (**A**) coating thickness; (**B**) dissolution. *Abbreviation*: NIR, near infrared.

with a paper-based system is incredibly time consuming. When one considers that the average batch record is 100–500 pages in length depending upon the complexity of the manufacturing process, the time to go through a batch record, re-enter data into an electronic repository (spreadsheet or database), verify the accuracy of the data, and then to begin the analysis process, this methodology is just not efficient or cost effective.

Risk Management

The ASTM E55 (Standard Terminology Relating to PAT in the Pharmaceutical Industry E 2363–06a) defines risk as "a combination of the probability of occurrence of harm and the severity of that harm." It is a structured evaluation of the impact or severity if something went wrong (e.g., patient death, dosage form rendered ineffective), and the occurrence (frequency) that the event will occur. Oftentimes, risk also takes into account whether it is possible to detect whether the issue will occur. This is done by evaluating the severity, probability, and detectability using a predefined ranking system. Several risk management processes have been developed, one common process being Failure Modes and Effects Analysis (FMEA).

An example of an FMEA evaluation is provided below in Figure 17. As can be seen, the impact of a failure for each step in a process is evaluated against the severity this risk may have on the patient. The severity of the impact should be performed by a

Frequency Scale		Severity Scale		Detectability Scale	
10	Frequent: Happens several times a year	10	Failure that can result in serious harm	10	Less than 50% of the time
7	Occasional: May happen a few times a year	7	Failure that can cause non-serious harm and/or significant dissatisfaction	7	50% of the time
4	Uncommon: May happen 2-5 times a year	4	Minor event causing delays	4	70% of the time
1	Remote: May happen sometimes in 5 to 30 years	1	Failure not noticeable or would not effect the delivery of the therapeutic effect	1	90% of the time before it reaches the patient

Step or Link in Process	Potential Failure Mode	Potential Cause or Mechanism	Frequency Likeliness Scale:1-10	Potential Effect of Failure Mode	Severity Potential for harm Scale: 1-10	Design Controls	System Control or Test Detect-ability Scale: 1-10	Risk Priority Number (RPN)*	Rank
Prep of Coating Suspen-sion	Solid powder for suspend-sion does not suspend properly	Mixing time too short	5	-Inconsistent coating layer with potential of decreased elegance or therapeutic effect -Clogged line causing room throughput delays -Clumps in bottom of tank leading to decreased suspension concentration	10 if coating is control release	Visual	10	500	13
						Viscosity	4	200	7
						Spectro-scopy	1	50	2
		Mixing speed too slow	5		10 if coating is control release	Visual	10	500	14
						Viscosity	4	200	8
						Spectro-scopy	1	50	3
		Charge to mixing tank too fast	8		10 if coating is control release	Visual	10	800	16
						Viscosity	4	320	11
						Spectro-scopy	1	80	5
		Charge to mixing tank inconsistent	4		10 if coating is control release	Visual	10	400	12
						Viscosity	4	160	6
						Spectro-scopy	1	40	1
		Change in raw material	3		10 if coating is control release	Release Test Gel Chroma-tography	8	240	9
		Suspending medium too cold	7		10 if coating is control release	Visual	10	700	15
						Viscosity	4	280	10
						Spectro-scopy	1	70	4

*(RPN)=Product of Freq times Severity times Detectability. This chart generated for purposes of this discussion.

FIGURE 17 Failure mode effects analysis for wet granulation process.

medical professional who can properly assess the impact of a dissolution failure on a patient. In addition, the occurrence and detectability of a failure mode occurring should be determined by the pharmaceutical scientist.

The initial assessment of the risk occurrence can be made during development through scientific judgment and experience and/or R&D batch data. Process capability can be estimated from R&D trials for each unit operation and formulation. Continued data collection during manufacturing will improve the accuracy of the overall risk assessment. Similarly, detectability can be further understood from the data obtained during method development and method validation. Sampling and acceptance criteria are also very important in this assessment.

Equally important to the risk evaluation is the process for mitigating product risks. An organization must decide how much risk it is willing to assume. A measure of that risk can be estimated by multiplying the ranking for severity and occurrence ($S \times O$ method) or severity, occurrence and delectability ($S \times O \times D$ method). The $S \times O \times D$ calculation gives a risk priority number (RPN). The higher the RPN, the higher the risk the organization is assuming.

Once the RPN number has been determined, mitigating the risk is typically accomplished through process improvements, lean manufacturing, six sigma programs, etc. These projects will generate new operating conditions that are more optimal for the product in terms of quality and risk. Once the process, measurement systems have been updated, the risk analysis should be performed again to determine if in fact the risk has been reduced to an acceptable level. If it has not, the cycle is repeated. If it has, the organization can move on to a higher priority project.

SUMMARY

A brief review of manufacturing operations is discussed. The interrelationships between each functional area requires that each area communicate effectively to ensure product quality and future success. The complexities of running a manufacturing organization are numerous.

The pharmaceutical industry is rapidly changing, using more advanced methods of developing and maintaining products on the market. Learnings from other industries are playing a key role in this evolution including how quality is viewed and evaluated, and advanced analytics, both in terms of instrumentation and mathematical methods. These advances will lead to even higher product quality at lower overall costs.

ACKNOWLEDGEMENTS

The authors would like to thanks specially to Alton Johnson, Tom Garcia, and Steve Hammond of Pfizer for Editorial review.

REFERENCES

1. Kourti T. The process analytical technology initiative and multivariate process analysis, monitoring and control. Anal Bioanal Chem 2006; 384:1043–8.
2. Garcia-Munoz S, Kourti T, MacGregor JF, Apruzzese F, Champagne M. Optimization of batch operating policies. Part I. Handling multiple solutions. Ind Eng Chem Res 2006; 45: 7856–66.
3. Garcia-Munoz S, Kourti T, MacGregor JF. Model predictive monitoring for batch processes. Ind Eng Chem Res 2004; 43:5929–41.
4. Geoffroy J. Resolving Manufacturing Problems and Continuous Improvement. New York: Arden House Conference, Harriman, January 29, 2004.
5. Montgomery DC. Design and Analysis of Experiments. New York: John Wiley & Sons, 2001.
6. Myers RH, Montgomery DC. Response Surface Methodology: Process and Product Optimization Using Designed Experiments. 2nd ed. New York: John Wiley & Sons, 2002.
7. Geoffroy J. Leverage Process Information to Advance Pharmaceutical Manufacture: Obtaining, Analyzing & Summarizing Drug Product Data to Improve Process Efficiency, Cost and Quality. ISPE Philadelphia, PA2005:14–6.
8. Stryczek K, Horacek P, Klema J, Castells X, Stewart B, Geoffroy J-M. Capitalizing on aggregate data for gaining process understanding: effect of raw material, environmental and process conditions on the dissolution rate of a sustained release product. J Pharm Innov 2007; 2:6–17.
9. Cox DR (1992). Planning for experiments. In: Hunter WG, Hunter JS, eds. Statistics for Experimenters: an Introduction to Design, Analysis and Model Building. New York: John Wiley & Sons, 1978.
10. Box GEP, Draper NR. Evolutionary Operation: a Statistical Method for Process Improvement. New York: John-Wiley & Sons, 1969.
11. Taguchi G, Wu Y, Wu A. Taguchi Methods for Robust Design: American Society of Mechanical Engineers, New York, NY 2000.
12. Ross PJ. Taguchi Techniques for Quality Engineering. New York: The McGraw-Hill Companies, Inc, 1996.
13. Montgomery DC. Introduction to Statistical Quality Control. 4th edn. New York: John Wiley & Sons, 2001.
14. Juran JM, Godfrey AB. Juran's Quality Handbook. 5th edn.: McGraw-Hill, New York, NY 1999.

15. Duncan AJ. Quality Control and Industrial Statistics: Richard D. Irwin, Inc., New York, NY 1974.
16. Wheeler DJ. Beyond Capability Confusion: the Average Cost-of-Use: SPC Press, Chicago, USA 1999.
17. Taguchi G, Chowdhury S, Taguchi S. Robust Engineering: Learn how to Boost Quality while Reducing Costs & Time to Market McGraw-Hill, New York, 1999.
18. Box G, Draper NR. Evolutionary Operation: a Statistical Method for Process Improvement. New York: John Wiley & Sons, Inc., 1998.
19. Minsk A, Hoffman D. Minimizing pharmaceutical company's risk, 2007.
20. Juran JM, ed. Juran on Quality by Design: the New Steps for Planning Quality into Goods and Services. New York: Free Press, 1992.
21. Nosal R, Garcia T. Capitalizing on Quality by Design Principles to Define Design Space for Drug Product Content Uniformity. AAPS Workshop: Pharmaceutical Quality Assessment—a Science and Risk-Based CMC Approach in the 21st Century, North Bethesda, MD, October 5, 2005.
22. Blanco M, Beneyto R, et al. Analytical control of an esterification batch reaction between glycerine and fatty acids by near-infra-red spectroscopy. Anal Chim Acta 2004; 521:143–8.
23. Shah RB, Tawakkul MA, Khan M. Process analytical technology: Chemometric analysis of Raman and near-infrared spectroscopic data for predicting physical properties of extended release matrix tablets. J Pharm Sci 2007; 96:1356–65.
24. Bo LH, Chen-Yunqing, Zhang XC. Characterization of anhydrous and hydrates pharmaceutical materials with THz time-domain spectroscopy. J Pharm Sci 2007; 96:927–34.
25. Minjunk K, Hoeil C, Youngah W, Kemper M. A new non-invasive, quantitative Raman technique for the determination of an active ingredient in pharmaceutical liquids by direct measurement through a plastic bottle. Anal Chim Acta 2007; 587:200–7.
26. Blanco M, Alcala M, Gonzalez JM, Torras E. A process analytical technology approach based on near infrared spectroscopy: tablet hardness, content uniformity, and dissolution test measurements of intact tablets. J Pharm Sci 2006; 95:2137–44.
27. Otsuka M, Yamane I. Prediction of tablet hardness based on near infrared spectra of raw mixed powders by chemometrics. J Pharm Sci 2006; 95:1425–33.
28. Medendorp J, Lodder R. Acoustic-resonance spectronomy as a process analytical technology for rapid and accurate tablet identification. AAPS PharmSciTech (electronic publication) 2006; 7:E25.
29. Pollanen K, et al. DRIFT-IR for quantitative characterization of polymorphic composition of sulfathiazole. Conference information: presented at the 9th International Conference on Chemometrics in Analytical Chemistry, Lisbon, Portugal, September 20–23, May 10–12. Anal Chim Acta 2004; 544:108–17.
30. Pollanen K, et al. ATR–FTIR in monitoring of crystallization processes: comparison of indirect and direct OSC methods. Chemometrics Intell Lab Syst 76:25–35.
31. Rantanen J, et al. Use of in-line near-infrared spectroscopy in combination with chemometrics for improved understanding of pharmaceutical processes. Anal Chem 77:556–63.
32. Workman J, Jr, et al. Process analytical chemistry. Anal Chem 2005; 75:2859–76.
33. Ciurczak EW. Near-infrared spectroscopy: Why it is still the number one technique in PAT. PAT 2006; 2003; 3:19–21.
34. Ciurczak EW. The process analytical technologies initiative – what is it, and where does spectroscopy come in? Spectroscopy 2003; 18:20–1.
35. Tatavarti AS, et al. Assessment of NIR spectroscopy for nondestructive analysis of physical and chemical attributes of sulfamethazine bolus dosage forms. AAPS PharmSci Tech (electronic resource) 2005; 6:E91–9.
36. Ciurczak E, Drennan J. Pharmaceutical applications of near-infrared spectroscopy. Near-Infrared Appl Biotech 2001; 25:349–66.
37. Hammond SV, Lyon R. Pfizer FDA Collaborative Research Agreement Study: process monitoring of pilot-scale pharmaceutical blends by near-infrared chemical imaging and spectroscopy. Proceedings of the International federation of Process Analytical Chemistry, Grayslake, IL 2006.

38. Watts DC, Afnan AM, Hussain Ajaz S. Process analytical technology and the ASTM Committee E55. Stand News 2004; 32:25–7.
39. Yu LX, Lionberger RA, Raw AS, D'Costa R, Wui H, Ajaz AH. Applications of process analytical chemistry to the crystallization process. Adv Drug Deliv Rev 2004; 56:349–69.
40. Hammond SV, et al. NIR analysis of antibiotic fermentations. In: Proceedings of the 4th International Conference on NIR Spectroscopy, 1991.
41. Hammond SV, et al. The cost avoidance role of NIR in pharmaceutical production. In: Proceedings of the 6th International Conference on NIR Spectroscopy, 1994.
42. Hammond SV, et al. Sample presentation in NIR. In: Proceedings of the 7th International conference on NIR Spectroscopy, 1995.
43. Hammond SV, et al. Focusing NIR spectroscopy on the business objectives of modern pharmaceutical production. In: Proceedings of the 8th International Conference on NIR Spectroscopy, 1997.
44. Aldridge PK. Apparatus for mixing and detecting on-line homogeniety. US Patent No. 5,946,088.
45. Axon TG, Hammond SV. Powder analysis. US Patent No. 6,362,891.
46. Aldridge PK, Hammond SV, Axon TG. Spectrophotometric Analysis. US Patent No. 5,859,703.
47. Hammond SV, Axon TG. Vial autosamples. US Patent No. 5,969,813.
48. Hammond SV, Axon TG, Brown R. Method and apparatus for spectrophotometrically analyzing characteristics of a tablet. US Patent No. 6,014,212.
49. Afnan AM, Chisholm RS. Mixing apparatus and method. US Patent No. 6,874,928.
50. Gehrlein L, Ciurczak E, Ritchie G, Bynum K. Method and apparatus for the determining the homogeniety of granulation during tabletting. US Patent No. 7,057,722.
51. Ufret C, Morris K. Modeling of powder blending using on-line near-infrared measurements. Drug Dev Ind Pharm 2001; 27:719–29.
52. El-Hagrasy AS, D'Amico F, Drennan JK. A process analytical technology approach to near-infrared process control of pharmaceutical powder blending. Part I: D-optimal design for characterization of powder mixing and preliminary spectral evaluation. J Pharm Sci 2005; 95:392–406.
53. El-Hagrasy AS, Delgado-Lopez M, Jrennan JK. A process analytical technology approach to near-infrared process control of pharmaceutical powder blending: Part II: Qualitative near-infrared models for prediction of blend homogeneity. J Pharm Sci 2005; 95:407–21.
54. El-Hagrasy AS, Drennan JK. A process analytical technology approach to near-infrared process control of pharmaceutical powder blending, Part III: Quantitative near-infrared calibration for prediction of blend homogeneity and characterization of powder mixing kinetics. J Pharm Sci 2005; 95:422–34.
55. Almeida-Prieto SA, Blanco-Mendez J, Otero-Espinar FJ. Microscopio image análisis techniques for the morphological characterization of pharmaceutical powders: influence of process variables. J Pharm Sci 2006; 95:348–57.
56. Hammond SV, et al. NIR microspectroscopy and the control of quality in pharmaceutical production. Eur Pharm Rev 1998; 2(1):6–17.
57. Clarke F, Clark D, Hammond SV, Moffat A. Chemical image fusion—the synergy of FT NIR and FT Raman spectroscopy to enable a more complete visualization of pharmaceutical formulations. Anal Chem 2000; 73:2213.
58. Clarke F, Lewis EN, Carroll J. A near infrered view of pharmaceutical formualtions. NIR News 2001; 12:16–8.
59. Clarke F, Hammond SV, Mattison CA. The development of NIR microscopy for process control in pharmaceutical manufacturing. J Process Anal Chem 2001; 7:115–8.
60. Clarke FC, Hammond SV, Jee RD, Moffat AC. Determination of information depth and sample size for the analysis of pharmaceutical materials using reflectance NIR microscopy. Appl Spectrosc 2002; 56:1475–83.
61. Clarke FC, Hammond SV. Near-infrared microscopy. In: Chalmers JM, Griffiths PR, eds. Handbook of Vibrational Microscopy, Vol. 2. Chichester: Wiley and Sons, 2002: 1405–18.

62. Clarke FC, Hammond SV. NIR microscopy of pharmaceutical dosage forms. Eur Pharm Rev 2003; 8:41–50.
63. Clarke F. Extracting process-related information from pharmaceutical dosage forms using near infrared microscopy. Vib Spectrosc 2003; 34:25–35.
64. Lewis NE, Schoppelrei, Lee E. Near-infrared chemical imaging and the PAT initiative: NIR-CI adds a completely new dimension to conventional NIR spectroscopy. The Role of Spectroscopy in Process analytical Technologies, 28–34. (Accessed January 2005, at www.spectroscopyonline.com) Spectroscopy 19(4) April 2004, 24–36
65. Lyon RC, Lester DS, Lewis NE, et al. Near-infrared spectral imaging for quality assurance of pharmaceutical products: analysis of tablets to assess powder blend homogeniety. AAPS PharmSciTech (electronic resource) 2002; 3:E17.
66. Hamad ML, Ellison CD, Khan M, Lyon R. Drug product characterization by macropixel analysis of chemical images. J Pharm Sci (electronic publication) 2007; 13:96, No. 12, (2007).
67. Barnes SE, Brown EC, et al. Vibrational spectroscopic and ultrasound analysis for the in-process monitoring of poly(ethylene vinyl acetate) copolymer composition during melt extrusion. Analyst (Cambridge, U.K.) 2005; 130:286–92.
68. Blanco GD, Sanchez NN, et al. High speed liquid chromatography for in-process control of rifabutin. Anal Chim Acta 2005; 531:105–10.
69. Blanco GD, Sanchez NN, et al. Fast high-performance liquid chromatography method for in-process control of sulbactam. Anal Chim Acta 2003; 498:1–8.
70. Blanco M, Beneyto R, et al. Analytical control of an esterification batch reaction between glycerine and fatty acids by near-infra-red spectroscopy. Anal Chim Acta 2005; 521:143–8.
71. Braatz RD, Hasebe S. Particle size and shape control in crystallization processes. AIChE Symposium Series 326 (Chemical Process Control-VI) 2002; 307–27.
72. Clegg IM, Everall NJ, et al. Online analysis using Raman spectroscopy for process control during the manufacture of titanium dioxide. Appl Spectrosc 2001; 55:1138–50.
73. de Diego A, Usobiaga A, et al. Application of the electrical conductivity of concentrated electrolyte solutions to industrial process control and design: from experimental measurement towards prediction through modelling. Trends Anal Chem 2001; 20:65–78.
74. Dorfner R, Ferge T, et al. Laser mass spectrometry as online sensor for industrial process analysis: process control of coffee roasting. Anal Chem 2004; 76:1386–402.
75. Doyle MJ, Newton BJ. Chromatography with online HPLC and ion chromatography for process control. CAST, Chromatography and Separation Technology. Cast, January/February, 2002; 4–7:9–12.
76. Epshtein NA. Structure of chemical compounds, methods of analysis and process control – validation of HPLC techniques for pharmaceutical analysis. Khim 2004; 38:212–28.
77. Faure A, Grimsey IM, et al. Process control in a high shear mixer-granulator using wet mass consistency: effect of formulation variables. J Pharm Sci 1999; 88:191–5.
78. Fransson M, Sparen A, et al. On-line process control of liquid chromatography. Anal Chem 73:1502–8.
79. Joergensen P, Pedersen JG, et al. On-line batch fermentation process monitoring (NIR) introducing "biological process time". J Chemometrics 2004; 18:81–91.
80. Kohori F, Yokoyama M, et al. Process design for efficient and controlled drug incorporation into polymeric micelle carrier systems. J Control Release 2002; 78:155–63.
81. Kristensen J, Schaefer T, et al. Direct pelletization in a rotary processor controlled by torque measurements. Part 1. Influence of process variables. Pharm Develop Technol 2000; 5: 247–56.
82. Larsen CC, Sonnergaard JM, et al. A new process control strategy for aqueous film coating of pellets in fluidised bed. Eur J Pharm Sci 2003; 20:273–83.
83. Laviana L, Fernandez MF, et al. HPLC for in-process control in the production of sultamicillin. J Pharm Biomed Anal 2003; 31:321–8.
84. Laviana L, Mangas C, et al. Determination and in-process control of zolpidem synthesis by high-performance liquid chromatography. J Pharm Biomed Anal 2004; 36:925–8.

85. Pollanen K, Hakkinen A, et al. DRIFT-IR for quantitative characterization of polymorphic composition of sulfathiazole. Presented at the 9th International Conference on Chemometrics in Analytical Chemistry, Lisbon, Portugal, September 20–23, 2004, May 10–12, 2004. Anal Chim Acta 2004; 544:108–17.
86. Pollanen K, Hakkinen A, et al. ATR–FTIR in monitoring of crystallization processes: comparison of indirect and direct OSC methods. Chemometrics Intell Lab Syst 2005; 76: 25–35.
87. Radtke G, Knop K, et al. In-process control of direct pelletization in the rotary fluidized bed using NIR spectroscopy. NIR News 1999; 10:4–5.
88. Rantanen J, Wikstroem H, et al. Use of in-line near-infrared spectroscopy in combination with chemometrics for improved understanding of pharmaceutical processes. Anal Chem 2004; 77:556–63.
89. Rantanen JT, Laine SJ, et al. Visualization of fluid-bed granulation with self-organizing maps. J Pharm Biomed Anal 2004; 24:343–52.
90. Rasenack N, Steckel H, et al. Micronization of anti-inflammatory drugs for pulmonary delivery by a controlled crystallization process. J Pharm Sci; 92:35–44.
91. Rogers TL, Gillespie IB, et al. Development and characterization of a scalable controlled precipitation process to enhance the dissolution of poorly water- soluble drugs. Pharm Res 2004; 21:2048–57.
92. Wille C, Pfirmann R. Pharmaceutical synthesis via microreactor technology increasing options for safety, scale-up and process control. Chimicaoggi 2004; 22:20–3.
93. Workman J, Jr, Koch M, et al. Process analytical chemistry. Anal Chem; 75:2859–76.
94. ASTM E2537-08 Standard Guide for Application of Continuous Quality Verification to Pharamaceutical and Biopharmaceutical Manufacturing, ASTM International, Conshohocken, PA, 2008.
95. Perry RH, Green DW, eds. Perry's Chemical Engineers' Handbook, 7th edn. New York: McGraw-Hill, 1997.
96. Clegg I, Pysik A, Brody R. Meeting the Regulatory Challenges Associated with Process Analytical Technologies. AAPS Annual Meeting, 2006. AAPS Phar Sci Tech 2005; 6(3) Artical 62 (http://www.aapspharmscitech.org).
97. Blanco M, Beneyto R, et al. Analytical control of an esterification batch reaction between glycerine and fatty acids by near-infra-red spectroscopy. Anal Chim Acta; 521:143–8.
98. Braatz RD, Hasebe S. Particle size and shape control in crystallization processes. AIChE Symposium Series 326 (Chemical Process Control-VI) 2002:307–27.
99. Rasenack N, Steckel H, et al. Micronization of anti-inflammatory drugs for pulmonary delivery by a controlled crystallization process. J Pharm Sci 2003; 92:35–44.
100. Wille C, Pfirmann R. Pharmaceutical synthesis via microreactor technology increasing options for safety, scale-up and process control. Chimicaoggi 2004; 22:20–3.
101. Gupta A, Peck G, Miller R, Morris K. Real-time near-infrared monitoring of content uniformity, moisture content, compact density, tensile strength, and Young's modulus of roller compacted blends. J Pharm Sci 2005; 94:1589–97.
102. El-Hagrasy AS, D'Amico F, et al. A Process Analytical Technology approach to near-infrared process control of pharmaceutical powder blending. Part I: D-optimal design for characterization of powder mixing and preliminary spectral data evaluation. J Pharm Sci 2006; 95:392–406.
103. Schweitz L, Fransson M, et al. On-line process control of gradient elution liquid chromatography. Anal Chem 2004; 76:4875–80.
104. El-Hagrasy AS, Drennen JK. A Process Analytical Technology approach to near-infrared process control of pharmaceutical powder blending. Part III: Quantitative near-infrared calibration for prediction of blend homogeneity and characterization of powder mixing kinetics. J Pharm Sci 2006; 95:422–34.
105. Faure A, Grimsey IM, et al. Process control in a high shear mixer-granulator using wet mass consistency: effect of formulation variables. J Pharm Sci 1999; 88:191–5.

106. Joergensen P, Pedersen JG, et al. On-line batch fermentation process monitoring (NIR) – introducing "biological process time". J Chemometrics 2004; 18:81–91.
107. Joergensen SB, Bonne D, et al. Monitoring and Control of Batch Processes. Boca Raton, FL: Chemical Industries, 106 (Batch Processes), 2006:419–62.
108. Kourti T. The Process Analytical Technology initiative and multivariate process analysis, monitoring and control. Anal Bioanal Chem 2006; 384:1043–8.
109. Kristensen J. Direct pelletization in a rotary processor controlled by torque measurements. III. Investigation of microcrystalline cellulose and lactose grade. AAPS PharmSciTech 2005; 6(3) Article 2 (http://www.aapspharmscitech.org)..
110. Kristensen J, Schaefer T, et al. Direct pelletization in a rotary processor controlled by torque measurements. Part 1. Influence of process variables. Pharm Dev Technol 2000; 5:247–56.
111. Kueppers S, Haider M. Process analytical chemistry – future trends in industry. Anal Bioanal Chem 2003; 376:313–5.
112. Larsen CC, Sonnergaard JM, et al. A new process control strategy for aqueous film coating of pellets in fluidised bed. Eur J Pharm Sci 2003; 20:273–83.
113. Liang J, Qian J. Multivariate statistical process monitoring and control: recent developments and applications to chemical industry. Chin J Chem Eng 2003; 11:191–203.
114. McCarthy MJ, Walton JH, et al. Development of microscale NMR sensors for control of food processing. Food Sci Biotechnol 2004; 13:848–51.
115. Meurens M, Kadji E, et al. Process control by NIR spectroscopy in the brewery. Cerevisia 2005; 30:195–8.
116. Olinga A, Siesler HW. Quality control and process monitoring by vibrational spectroscopy. NIR News 1998; 11:9–11.
117. Radtke G, Knop K, et al. In-process control of direct pelletization in the rotary fluidized bed using NIR spectroscopy. NIR News 1999; 10:4–5.
118. Rantanen J, Wikstroem H, et al. Use of in-line near-infrared spectroscopy in combination with chemometrics for improved understanding of pharmaceutical processes. Anal Chem 2004; 77:556–63.
119. Rantanen JT, Laine SJ, et al. Visualization of fluid-bed granulation with self-organizing maps. J Pharm Biomed Anal vol.24 Issue 8, January 2001;24:343–352
120. Siren H, Luomanpera K, et al. Process control and drug analysis with an on-line capillary electrophoresis system. J Biochem Biophys Methods 2004; 60:295–307.
121. ICH Q9 Quality Risk Management, 2006.

4

A Forward-Looking Approach to Process Scale-Up for Solid Dose Manufacturing

Fernando J. Muzzio, Marianthi Ierapetritou, Patricia Portillo and Marcos Llusa
Department of Chemical and Biochemical Engineering, Rutgers University, Piscataway, New Jersey, U.S.A.

Michael Levin
Metropolitan Computing Corporation (MCC), East Hanover, New Jersey, U.S.A.

Kenneth R. Morris, Josephine L.P. Soh, and Ryan J. McCann
Department of Industrial and Physical Pharmacy, Purdue University, West Lafayette, Indiana, U.S.A.

Albert Alexander
AstraZeneca, Wilmington, Delaware, U.S.A.

INTRODUCTION

The purpose of this chapter is to provide a realistic discussion of both current practices and emerging issues in process scale up for pharmaceutical oral solid products. At the time when this chapter is being written (late Summer, 2007), the pharmaceutical manufacturing community is actively engaged in a broad dialogue regarding modernization of methods used for pharmaceutical product and process design. In the preceding five years, under the banners of process analytical technology (PAT) and quality by design (QbD, also known in other fields as "model-based design and optimization"), the pharmaceutical industry has focused substantial efforts on improving its understanding of key unit operations, and on developing statistical, instrumental, and fundamental methods for characterizing and controlling sources of variability in product performance.

In recent discussion forums, it has became increasingly clear that application of QbD methods is not a discrete activity to be "done and done with" at an early stage of product/process development, but rather a longitudinal component of the product life cycle, to be used initially as a formulation design/screening methodology, later on as a product/process optimization approach, and finally as a continuous improvement method during commercial manufacturing.

However, while the conceptual use of statistical QbD methodologies is straightforward and the necessary toolbox is well developed and has been used in other industries for decades, actual implementation is a very large task, for several reasons:

1. There is incomplete knowledge regarding which "product performance parameters" are actually relevant to in vivo product performance. As a result, "quality improvement" efforts typically involve meeting standard values in performance parameters (such as RSD in drug content, or F1&F2 "indexes" in in vitro dissolution) that are regarded by many as somewhat arbitrary
2. In spite of much recent progress by regulatory bodies, the current global regulatory framework does not facilitate implementation of continuous process improvement approaches
3. Mechanical and physicochemical properties of many active pharmaceutical ingredients (APIs) and excipients are at best only partially understood, limiting identification of critical material variables
4. For many process components there is an incomplete knowledge of critical process variables
5. Because the theoretical, all-encompassing parametric space of all conceivably relevant variables is very large, and because of the incomplete knowledge of what is critical and what is not, many current attempts at application of QbD methodologies are likely to be sub-optimal.

This chapter is organized as follows: First, we discuss in general terms the current state of pharmaceutical product and process development, and we identify some roadblocks that emerge frequently during process scale up. Subsequently, we briefly review QbD methodologies. The next several sections discuss essential issues that are important in the scale up of the most common process components used to manufacture oral solid dosage forms (blending, lubrication, wet and dry granulation, and compaction). We then shift our attention to an emerging issue. In recent years, substantial interest has emerged on the implementation of continuous methods for solid dose manufacturing. While some of the actual process components used in continuous manufacturing approaches are quite similar (and sometimes identical) to those used in batch processing, operation of a continuous process provides substantial opportunities for improved performance, increased controllability, and reduced cost. However, effective implementation of continuous approaches capable of realizing such gains also requires some evolution in the regulatory perspective. This topic is addressed in the closing comments of this chapter.

GENERAL ISSUES IN SCALE-UP OF SOLID DOSE MANUFACTURING PROCESSES

Traditional pharmaceutical product and process development, illustrated in Figure 1, largely follows a sequential task structure (1). Typically, the first stage (drug synthesis) yields a drug substance in powder form. At this stage, material properties needed to achieve desired product performance are largely unknown. In the formulation stage the material is turned into a preliminary product employing small-scale experiments following a recipe that is expected to achieve the desired release profile. However, at this stage it is not generally known how processing choices will affect manufacturability. In the next stage, the process is scaled up to a pilot plant, and later, to the manufacturing scale, by successively testing and adapting the tasks of the recipe to larger scale equipment. Rigorous scale-up methods are seldom available (2). Processing parameters

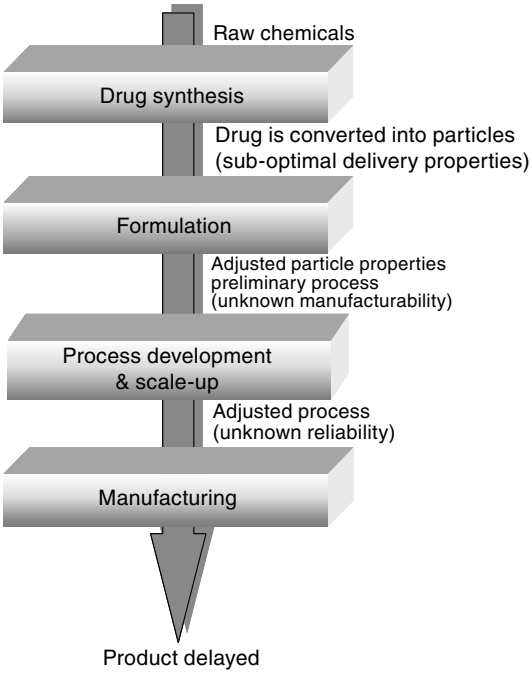


FIGURE 1 The current product development process, its major stages, and their outcomes.

are empirically adjusted until performance likely to satisfy regulatory compliance is achieved. Once this is accomplished, the manufacturing process becomes much harder to improve because the rigorous science base does not exist for reliably predicting the impact of further material or process changes on the final product. This knowledge gap, current regulatory practice and the business pressure to speed the product to market significantly hinder product and process optimization and adoption of new technologies (http://www.fda.gov/cder/pike/July2004.htm).

Throughout this process, lack of predictive methods for identifying and controlling critical material and process variables hinders implementation of development and optimization methods, and is the main reason for the lack of flexibility in the regulatory framework. For example, an often serious gap in our ability to predict scale-up from early solid oral dosage form (SODF) product development through the pilot plant/clinical supplies and manufacturing is the uncertainty in the API characteristics as the parallel API development and scale-up proceed. In the pursuit of efficient commercial synthetic pathways, engineers will often make logical changes that may change the physical properties of the final API. The changes may or may not negatively impact on the use of the API in product production; however, the impact is typically only retrospectively addressed. It would of course make the most sense to coordinate the API and product development efforts; however, this is made more difficult because many of the variables that determine the limits of the physical properties needed for production are not firmly known early in the product development process. Some of these variables include:

1. *The final process*. It is often the case that during early product development, sometimes even through clinical supply manufacture, the final manufacturing site and equipment have not been selected. This may be due to uncertainties in the volume of the product to be produced and/or the type of equipment available that is appropriate for the process select. As the type of processing equipment may change either

in principle of operation or manufacturer, the impact on the product produced will be necessarily less certain. One approach to obviating the differences is to use material monitoring such as described in the PAT guidance to ensure that the product quality is maintained even in the face of needed adjustments to remain within a design space.

2. *The final dose*. In early development the final dose required of the dosage form may still be undetermined. This may be of particular importance for directly compressed or roller compacted dosage forms if the dose is higher than anticipated. Such changes may impact the ability to blend and/or compact sufficiently for manufacture. This requires that the micromeritic and mechanical properties of the API be well understood in order to alter either the formulation or the processing variables to try to achieve the required product properties.
3. *The quantities of API required*. Another often missed issue is underestimating the demand for the product and therefore the need for higher volumes of API. As the volumes of API required increase, the throughput may be enhanced by *crashing out* or rapid crystallization of the API while still remaining within specifications. However, if these changes result in the production of small needlelike crystals where more regular and/or larger crystals had been formed in the past, the process may be negatively impacted. This is why understanding the process sufficiently to set meaningful specifications on the API is so important.
4. *Full characterization of the solid-state of the API*. As has often been said by Professor Stephen Byrn of Purdue University, "the best polymorph screen is a scale up." This means that unanticipated crystal form or solid form changes may occur as the API process is scaled up which may make material and production different than that which was tested in the clinic. Again, full understanding of thermodynamics of the materials is essential to anticipate, avoid, or troubleshoot such changes.
5. *Flow properties of the powder stream under actual conditions*. Another potentially major gap in the SODF product scale up procedure lies in the methods of material transfer between unit operations on the small scale versus full scale. At the small-scale material transfer is typically done manually, i.e., scooping powders into hoppers or tablets into coding pans, etc. However at full scale it is more typical to have dense phase transfer via pneumatic systems or to accomplish transfer by moving pieces of equipment adjacent to other pieces of equipment and directly discharging, e.g., the contents of a bin blender into the feed of a tablet press. Essentially, material transferred full scale represents a new unit operation not modeled or even considered at the small-scale.

REGULATORY ISSUES AND THE QBD INITIATIVE

For the past decade, Scale-up and process improvement has been largely ruled by FDA regulations broadly known as the Scale-Up and Post-Approval Changes (SUPAC) framework (3–11). The main issue and challenge of scale-up is that R&D, clinical studies and production are using equipment of a different scale. Pre-approval changes caused by dimensional dissimilarities of equipment may require repetition of expensive clinical studies. On the other hand, once approved, a process is very difficult to change or transfer due to the SUPAC regulations, except for a well-defined list of changes that are regarded to have relatively small impact. Such "annual report" changes can be implemented without requiring prior approval and only require a post-implementation report to the regulatory agency.

Current practices in pharmaceutical process development involve univariate (OVAT, "one variable at a time") efforts. One variable is examined for a few conditions, which in practice, are selected within a "safe" subset of the permissible range of variation. A value of this parameter is selected and kept subsequently constant. Another variable is then examined, a value is chosen, and the process continues sequentially. Intuitively, unless the target function is essentially a plane, if the end result is anywhere near the global optimum, it is only by chance. A historical reason for this dated practice is that the regulatory framework greatly discouraged implementation of the virtuous cycle mentioned above, which is the heart of the optimization process. Once a process was approved, the cost of implementing improvements (and the risk of examining process performance outside approved sets of parameters) were simply too high. As a result, while the rest of the industrial world embarked in wave after wave of quality revolutions, pharmaceutical process development practices stayed frozen in decades-old paradigms from a time before computer models.

The Process Analytical Technology Guidance (12), introduced four years ago, represented a significant attempt to evolve from this situation. The scientific approach to scale-up is referred to as one of the primary sources of data and information needed to understand the "multi-factorial relationships among various critical formulation and process factors and for developing effective risk mitigation strategies (e.g., product specifications, process controls)". One of the declared PAT goals is "to design and develop processes that can consistently ensure a predefined quality at the end of the manufacturing process". Since each operation along the scale-up path can be intimately understood and controlled through PAT, a concept of "Make Your Own SUPAC" was developed (alternatively called PAT-SUPAC, or SUPAC-C) by Ajaz Hussain the former deputy director of the Office of Pharmaceutical Sciences at FDA.

Discussions concerning the use of QbD methods, which started around 2004, have intensified in the last two years, and have captured the attention and interest of both agencies and industry. The fundamental assumption underlying QbD is that if critical sources of variability can be understood, then product performance can be controlled by using the manufacturing process to mitigate variability in material properties. The ultimate goal of QbD is "real-time release" of finished product. As mentioned above, this is a conceptually clear proposition, but in practice it involves a substantial amount of effort. Even more importantly, implementation of QbD-based processes requires deep transformation of the regulatory mentality: in a post-QbD era, the process is no longer fixed; far from it, it is a dynamic exercise that continuously mutates to accommodate variations in raw material properties.

An appropriate starting point for a discussion of model-based design and optimization requires clarification of some terminology. Certain engineering terms are often used in pharmaceutical manufacturing but not necessarily with the same meaning, generating significant confusion. Consider, for example, the term "optimization." In pharmaceutical process development "optimization" often refers to the practice of examining process performance empirically for a small set of parameter values, often chosen based on experience (such as three different blending times), and then selecting the value that gives the results that are deemed most adequate (usually without sufficient replication of results and often without use of statistical methods to determine significance). "Scale up" refers to a process development stage (Fig. 1) where the process recipe is carried out in larger equipment, and scale equivalence is "established" by demonstrating the ability to manufacture "acceptable product." A manufacturing process is said to be "in control" when it is possible to make a large number of batches of product within specification.

To an engineer in most other industries, these terms have radically different meanings. Optimization is the use of a predictive model to determine the best possible design of a product, or the best possible operating condition for a process. To find "the best," the *design space* (the permissible region of parameters given technical, regulatory, or economic constraints) is identified. A quantitative target function describing the property (or properties) to be optimized is developed. The target function can be a single performance attribute (quality, technical performance, profit), or a combination of multiple parameters after they are assigned a given weight. Once the design space and the target function are known, the absolute minimum (or maximum) of this function is found.

In contrast, in many other industries, the optimization process is multivariate (multiple variables *and their interactions* are simultaneously examined) and the design effort is conducted in iterative fashion (Fig. 2), beginning with the development of a model of the process. The model can be statistical (13), fundamental (based on conservation laws for momentum, mass, and energy, thermodynamics, constitutive models, etc) or some hybrid combination therewith. In early stages of product or process design, relatively little is known, and only a preliminary version of the model can be developed. A "first pass" optimization exercise is conducted. Model predictions are compared with actual performance, and results are used to improve the model itself. Results are also used to refine knowledge about design space boundaries. The more refined model is used to generate higher quality performance predictions, which are again used to predict an optimum operating regime. Comparison of prediction and practical observations are used to further improve the model, the target function, and the design space. The process continues *ad infinitum* following a virtuous cycle that leads to ever better predictive power.

Since economic conditions, process capabilities, and regulatory requirements change over time, both the so-called design space and the target function are dynamic structures, and the optimum product or process design is, in fact, a moving target, although the underlying physics is the same. Model-based optimization is ideally suited to respond to these dynamics. Once a high quality model is available, the change in conditions can be incorporated into the process, and a new iteration along the virtuous cycle is performed to generate the new selection of optimum processing conditions.

True process optimization can be challenging. The design space can be a complex, irregularly shaped region (or set of disconnected regions) in an n-dimensional space. The target function can have local minima that can "trap" the trajectory of the

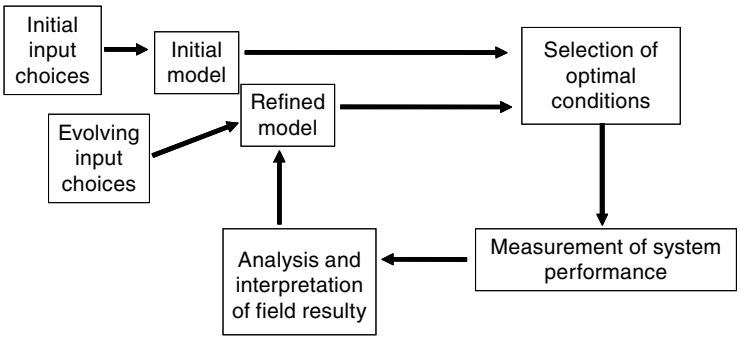


FIGURE 2 The iterative optimization process. An initial model is developed, used to predict process performance, tested by comparison with experiment, refined, and used to improve prediction. The process naturally accommodates changes in economic or regulatory constraints.

solution-seeking algorithm. To avoid such "non-convex" situations, searching algorithms have been developed that incorporate a certain measure of randomization in the sequential selection of process conditions to be examined. Ample literature exists on the topic and is not reviewed here in the interest of brevity, for an introduction see (14,15).

Two other important issues deserve mention here. A common misconception is to assume that the optimization effort is a discrete activity to be completed prior to product approval. In reality, any such attempt to front-loading the development method is unlikely to succeed for several reasons. First, as mentioned before, both materials and processes exhibit dynamic change, and the optimum process is a moving target. Second, the amount of work needed to identify, characterize, and control all variables affecting product performance is quite large, so at best only a first-pass design can be achieved within the short time frames associated with product development in the current pharmaceutical business cycle. Third and most important, extremely valuable information is generated by the manufacturing operation, which can be used to further refine models and improve performance. The second issue, which is a logical consequence of this reality, is that in an enlightened post-QbD regulatory framework, it is understood, accepted, and even encouraged, to use dynamic control specifications that allow for more flexibility at the beginning of the manufacturing life cycle (when knowledge is sparser) but benefit from greatly improved quality once the process reaches maturity.

CURRENT PRACTICES IN SCALE-UP OF BATCH PROCESS COMPONENTS—SCALE UP BY SIZE ENLARGEMENT

Blending and Lubrication

General Issues

The quality of a final product is a direct measure of the success of any manufacturing operation. Processes that incorporate powder or granular blending steps are often highly dependent on the degree of homogeneity of the final mixture. In the pharmaceutical context, inefficient blending can lead to increased variability of the active component in the final dosage form, threatening the health of patients. Content Uniformity issues have four main root causes: (*i*) weight variability in the finished dose, which is often related to flow properties of the powder stream, (*ii*) poor equipment design or inadequate operation, (*iii*) particle segregation (driven by differences in particle properties), and (*iv*) particle agglomeration, driven by electrostatics, moisture, softening of low melting point components, etc.

Additional problems may occur when a lubricant is added to the mixture (as in the case of most pharmaceutical formulations). Lubricants such as magnesium stearate (MgSt), work by interposing a film of low shear strength material at the interface between the tablet mass and the die wall. The addition of dry lubricants allows compression at lower pressure and reduces the generation of heat during tablet compression. The effect of the lubricant depends on the amount and intensity of shear energy that is applied to the lubricated mixture. Although small amounts of MgSt are used (around 1%), it is known that the insolubility of this material poses a problem to the penetration of the solid dosage form by the gastrointestinal fluids intended to dissolve it. It can also impart other undesirable characteristics to tablets. The interactions between the lubricant and excipient or between the lubricant and the active ingredient may cause insufficient mechanical strength of tablets and capsules. Poor lubrication also leads to variability in the compaction step (i.e., the tablet will stick to the press) and it may hinder powder flowability.

Over-lubrication is also a situation that must be avoided. Overlubrication occurs whenever the addition of dry lubricant tends to coat the particles of the formulation, thus decreasing the binding between particles, decreasing the strength of the tablets, and resulting in decreased tablet solubility, increasing the disintegration and dissolution time.

Tumbling blenders remain the most common means for mixing granular constituents in the pharmaceutical industry. Tumbling blenders are hollow containers attached to a rotating shaft; the vessel is partially loaded with the materials to be mixed and rotated for some number of revolutions. The major advantages of tumbling blenders are large capacities, low shear stresses, and ease of cleaning. These blenders come in a wide variety of geometries and sizes, from laboratory scale (<16 qt.) to full-size production models (>500 ft^3). A sampling of common tumbling blender geometries include the v-blender (also called the twin-shell blender and the PK blender), the double cone, the bin blender (also known as the IBC blender, and the tote blender), and the rotating drum. Surprisingly little is known about flow patterns, mixing dynamics, and segregation in these devices [for a review on solid mixing devices, see (16–19) and references therein]. Flow patterns are believed to consist of a combination of thin, rapid flow regions characterized by high shear and density gradients in areas where the yield strength of the powder is exceeded, and nearly non-deforming regions everywhere else (20–21). The main transport mechanisms, nevertheless, are yet to be well characterized in realistic blenders. To date, the design and control of three-dimensional blenders have been based more on trial and error than on quantitative or analytic methods. Even quantitative characterizations of mixing performance as a function of the most basic parameters, such as vessel speed or filling level, are scarce in the literature (22–26).

The other most common type of mixer is the convective blender, where flow is created by one or more impellers rotating within a fixed shell. Their main advantages are ability to impart high shear when needed, reduced ingredient segregation, and the ability to use them for wet granulation. While they are also available in a wide range of sizes, the largest available capacity is often an inverse function of the maximum shear rate they can apply. Examples of convective mixers include ribbon blenders, high-shear granulators, and plow-mixers.

There are currently no rigorous techniques to predict blending scale-up criteria in either type of blender without prior experimental work. Typically, blending studies performed in industry start with a small-scale, try-it-and-see approach. The following questions usually arise:

1. What rotation rates should be used?
2. Should filling level be the same?
3. How long should the blender be operated?
4. Are variations to the blender geometry between scales acceptable?

Further complicating the issue is that rotation rates for typical commercially available equipment are often fixed, obviating question (1) and suggesting that, under such conditions, true dynamic or kinematic scale-up may not be possible.

Defining Mixedness

The final objective of any granular mixing process is to produce a homogenous blend. Determining mixture composition throughout the blend is a difficulty for granular systems. As yet, few reliable techniques for on-line measuring of composition have been developed and granular mixtures are almost always quantified by removing samples from the mixture. To determine blending behavior over time, the blender is stopped at fixed

intervals for repeated sampling; a process that may change the state of the blend. Once samples have been collected, the mean value and sample variance is determined and then often used in a mixing index (16,27). In general, the pharmaceutical industry has relied on the relative standard deviation [(RSD) aka coefficient of variability], and the usual specification is that the measured RSD should be smaller than a given value (6% and 5% are the two most commonly cited values). This approach contains the intrinsic assumptions that the blend is a random structure with a Gaussian (normal) distribution of compositions, and that a small number of samples can sufficiently characterize variability throughout the blend. Unfortunately, in many instances where blends exhibit segregation, agglomeration, and/or incomplete mixing, *distributions deviate substantially from normality*, and a simple measure of breath such as the RSD does not predict the frequency of extreme values.

Furthermore, sample size can have a large impact on apparent variability. Samples that are too small can show exaggerated variation and magnify sampling error, while too large a sample can blur concentration gradients. Hence it is paramount that a sufficient number of samples are taken representing a large cross-section of the blender volume. Another concern is whether standard sampling techniques retrieve samples that are truly representative of local concentration at a given location. Thief probes remain the most commonly employed instrument for data gathering. These instruments have been demonstrated to sometimes induce large sampling errors coming from poor flow into the thief cavity or sample contamination (carry over from other zones of the blender) during thief insertion (16) (a method to assess blend uniformity and blend sampling error is given in PDA Technical Report #25 (17)).

Finally, the degree of mixedness at the end of a blending step is not always a good indicator of the homogeneity to be expected in the final product. Many granular mixtures can spontaneously segregate into regions of unlike composition when perturbed by flow, vibration, shear, etc. Once a good blend is achieved, the mixture still must be handled carefully to avoid any "de-mixing" that might occur.

Mixing Mechanisms

Current thinking describes the blending process as taking place by three essentially independent mechanisms: convection, dispersion, and shear. Convection causes large groups of particles to move in the direction of flow (orthogonal to the axis of rotation), the result of vessel rotation or impeller motion. Dispersion is the random motion of particles as a result of collisions or inter-particle motion, usually orthogonal to the direction of flow. Shear separates particles that have joined due to agglomeration or cohesion and requires high forces. While these definitions are helpful from a conceptual standpoint, blending does not take place as merely three independent, scaleable mechanisms. Rather, the mechanisms act simultaneously, and exhibit different scale dependence, making scale up a difficult task at best.

Let us now describe the main phenomena in each of the two types of blenders. Powder mixing in tumbling blenders takes place as the result of particle motions in a thin cascading layer at the surface of the material, while the remainder of the material below rotates with the vessel as a rigid body. All the mixing (and all the segregation) in a tumbling blender occurs in the cascading region. Tumbling blenders impart very little shear, unless an intensifier bar (I-bar) or chopper blade is used (in some cases, high shear is detrimental to the active ingredient, and is avoided). Without an intensifier bar, the little shear that is present occurs at the powder cascade, concurrently with tensile normal stresses, which tend to separate adjacent particles. Compressive normal stresses are static and are due entirely to the weight of the powder loaded to the vessel.

In convective mixers, homogenization is driven by the flow field created by the motion of the impeller. Typically, the entire powder mass experiences a certain amount of shear at all times. Shear levels are controlled entirely by the speed of the impeller that drives the flow. Shear always results in tensile stresses. However, differently from tumbling mixers, convective mixers also apply compressive normal stresses that can be much larger than those due to the powder weight (hence their use as granulators).

In general, regarding scale-up requirements, mixing processes can be classified into two fundamentally different groups, free-flowing and cohesive materials, having different mixing requirements.

Free-Flowing Materials

Free-flowing materials are powders and granulations where inter-particle cohesive forces are small enough to allow particles to move individually. Typically, this situation is descriptive of materials where particles are larger than ~100 μm and where attractive forces between particles are similar or smaller than the particle weight. These materials do not require substantial shear to be mixed, and tumbling blenders are often the preferred route. The main process risks, beside those emanating from incorrect operation (discussed below), are due to segregation either within the blender or after blender discharge.

To understand scale-up requirements, one must first recognize that most tumbling blenders are symmetric in design; this symmetry can be the greatest impediment to achieving a homogeneous mixture. The mixing rate often becomes limited by the amount of material that can cross from one side of the symmetry plane to the other (18–22). Some blender types have been built asymmetrically (e.g., the slant cone, the cross-flow v-blender), and show greater mixing proficiency. Furthermore, by rocking the vessel as it rotates, the mixing rate can also be dramatically increased (23). Asymmetry can be "induced" through intelligent placement of baffles, and this approach has been successfully tested on small scale equipment (21,24–26) and used in the design of some commercial equipment. But, when equipment is symmetric and baffles unavailable, careful attention should be paid to the loading procedure as this can have an enormous impact on mixing rate.

Non-systematic loading of multiple ingredients will have a dramatic effect on mixing rate if dispersion is the critical blending mechanism. For instance, in a v-blender, it is preferable to load the vessel either through the exit valve or equally into each shell. This ensures that there are near equal amounts of all constituents in each shell of the blender. Care must be taken when loading a minor (~1%) component into the blender—adding a small amount early in the loading process could accidentally send most of the material into one shell of the blender, and substantially slow the mixing process. Smaller blenders entail shorter dispersal distances necessary for complete homogeneity, and thus, may not be as affected by highly asymmetric loading. As a final caution, the order of constituent addition can also have significant effects on the degree of final homogeneity, especially if ordered mixing (bonding of one component to another) can occur within the blend (28).

Inter-shell flow is the slowest step in a v-blender because it is dispersive in nature while intra-shell flow is convective. Both processes can be described by similar mathematics, typically using an equation such as

$$\sigma^2 = Ae^{-kN} \tag{1}$$

where σ^2 is mixture variance, N the number of revolutions, A an unspecified constant, and k the rate constant (20,29). The rate constants for convective mixing, however, are orders

of magnitude greater than for dispersive mixing. Thus, unequal loading across the symmetry plane places emphasis on dispersive mixing and is comparatively slow compared to top-to-bottom loading which favors convective mixing.

When discussing tumbling blender scale-up, one parameter consideration that arises is whether rotation rate should change with variations in size. Previous studies on laboratory scale v-blenders and double cones have shown that, when far from the critical speed of the blender, the rotation rate does not have strong effects on the mixing rate (20,21) (the critical speed is the speed at which tangential acceleration due to rotation matches the acceleration due to gravity). These same studies showed that the number of revolutions was the most important parameter governing the mixing rate. Equation (1) was derived by assuming that the mixture went through a specific incremental increase in mixedness with each revolution (either by dispersion or convection). While this approach has been shown to be successful at modeling increasing in mixture homogeneity, no scaling rules have been determined for the rate constants that govern this equation, and this remains an open question for further inquiry.

Given a geometrically similar blender and the same mixture composition, it would seem obvious that the fill level should also be kept constant with changes in scale. However, an increase in vessel size at the same fill level may correspond to a significant decrease in the relative volume of particles in the cascading layer compared to the bulk—this could be accompanied by a large decrease in mixing rate. It has been shown in 1 pint v-blenders that running at a 40% fill brings about a mixing rate that is nearly 3 times faster than at 60% fill (20). Thus, although fill level should be kept constant for geometric similarity, it may be impossible to match mixing rate per revolution across changes in scale if the depth of the flowing layer is a critical parameter.

In the literature, the Froude number ($\mathrm{Fr} \equiv \Omega^2 R/g$; where Ω is the rotation rate, R is the vessel radius, and g is the acceleration from gravity) is often suggested for tumbling blender scale-up (30–33). This relationship balances gravitational and inertial forces and it can be derived from the general equations of motion for a general fluid. Unfortunately, no experimental data has been offered to support the validity of this approach. Continuum mechanics may offer other dimensionless groups, if a relationship between powder flow and powder stress can be determined. However, Fr is derived from equations based on continuum mechanics, but the scale of the physical system for blending of granular materials is on the order of the mean free path of individual particles, which may invalidate the continuum hypothesis. A less commonly recommended scaling strategy is to match the tangential speed (wall speed) of the blender; however, this hypothesis also remains untested.

As an example, consider the general problem of scaling a 5- to 25-ft^3 blender using *Fr* as the scaling parameter: The requisites are to ensure geometric similarity (i.e., all angles and ratios of lengths are kept constant), and keep the total number of revolutions constant. With geometric similarity, the 25-ft^3 blender must look like a photocopy enlargement of the 5-ft^3 blender. In this case, the linear increase is ($5^{1/3}$) or a 71% increase. Also for geometric similarity, the fill level must remain the same. To maintain the same Froude number, since R has increased by 71%, the rpm (Ω) must be reduced by a factor of $(1.71)^{-1/2} = 0.76$, corresponding to 11.5 rpm. In practice, since most blends are not particularly sensitive to blend speed, and blenders available are often fixed speed, the speed closest to 11.5 rpm would be selected. If the initial blend time were 15 minutes at 15 rpm, the total revolutions of 225 must be maintained with the 25 ft^3 scale. Assuming 11.5 rpm were selected, this would amount to a 19.5-minute blend time. Though this approach is convenient and used often, it remains empirical.

Common violations of this approach that can immediately cause problems include the attempt to scale from one geometry to another (e.g., v-blender to in-bin blender),

changing fill level without concern to its effect, and keeping blending time constant while changing blender speed.

Cohesive Powders

A substantially different scenario arises for cohesive powders. The effect of cohesion of powder flow and scale-up, in particular for mixing operations, remains an open problem, and only a brief discussion is provided here. In simple terms, a cohesive powder can be defined as a material where the adhesive forces between particles exceed the particle weight by at least an order of magnitude. In such systems, particles no longer flow independently; rather, they move in "chunks" whose characteristic size depends on the intensity of the cohesive stresses. Two main effects are often observed for cohesive blends: (*i*) the overall mixture is sufficiently cohesive to affect the flow of the material in the blender, and (*ii*) a specific ingredient (often the active) is cohesive enough to display formation of agglomerates. Let us discuss the separately:

The effective magnitude of cohesive flow effects depends primarily on two factors: the intensity and nature of the cohesive forces (e.g., electrostatic, van der Waals, capillary moisture) and the packing density of the material (which determines the number of interparticle contacts per unit area). This dependence on density is the source of great complexity: cohesive materials often display highly variable densities that depend strongly on the immediate processing history of the material. In spite of this complexity, a few "guidelines" can be asserted within a fixed operational scale:

1. Slightly cohesive powders mix faster than free flowing materials.
2. Strongly cohesive powders mix much more slowly.
3. Strongly cohesive powders often require externally applied shear (in the form of an impeller, and intensifier bar, or a chopper.
4. Baffles attached to vessels do not increase shear substantially.

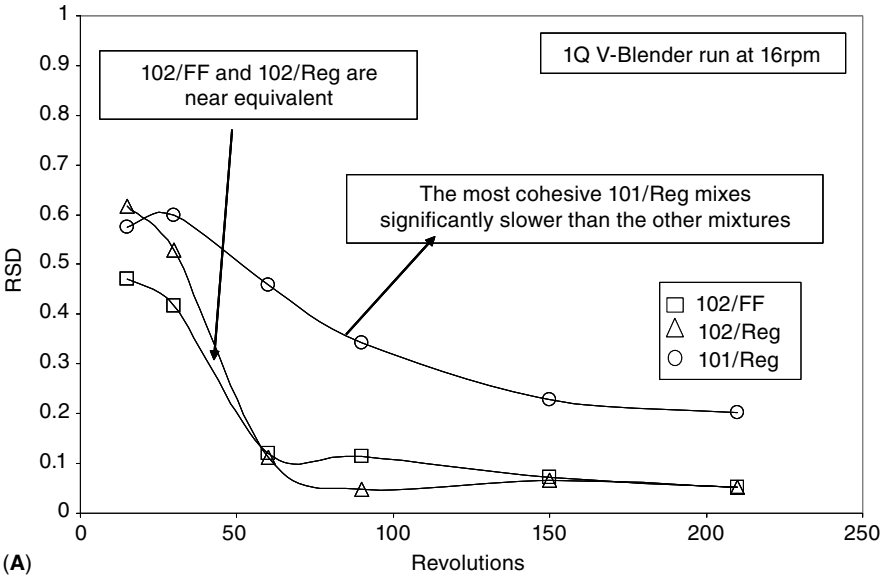
Lacking a systematic means to measure cohesive forces under practical conditions, the effects of cohesion on scale-up have been studied rarely. The most important observation is that cohesive effects are much stronger in smaller vessels, and their impact tends to disappear in larger vessels. The reason is simple: while cohesive forces are surface effects, the (gravitational and convective) forces that drive flow in powder blenders grow proportionally to the vessel volume. Thus, as we increase the scale of the blender, gravitational and convective forces grow faster, overwhelming cohesive forces. This can also be explained by remarking that the characteristic "chunk" size of a cohesive powder flow is a property of the material, and thus to a first approximation it is independent of the blender size. As the blender grows larger, the ratio of the "chunk" size to the blender size becomes smaller.

Both arguments can be mathematically expressed in terms of a dimensionless "cohesion" number Π_c

$$\Pi_c = \sigma/\rho g R = (\sigma/\rho g)/R = S/R \tag{2}$$

where σ is the effective (surface averaged) cohesive stress (under actual flow conditions), ρ is the powder density under flow conditions, g is the acceleration of gravity, and R is the vessel size. The group $S = (\sigma/\rho g)$ is the above mentioned "chunk" size, which can be more rigorously defined as the internal length scale of the flow driven by material properties.

Thus, as R increases, Π_c decreases. This is illustrated in Figure 3, which shows the evolution of the RSD of a blending experiment in a small V-blender for three mixtures of



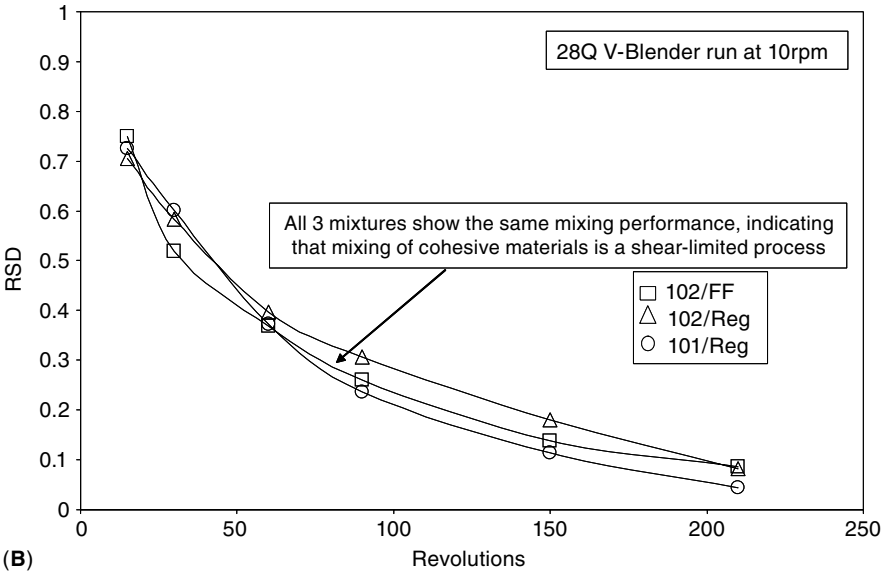


FIGURE 3 (**A**) RSD measured for axially segregated blends of different cohesion in a 1-qt V-blender. As cohesion increases, blending becomes slower. (**B**) RSD measured for axially segregated blends of different cohesion in a 28-qt V-blender. For a large vessel, the effects of cohesion become unimportant. *Abbreviation*: RSD, relative standard deviation.

different cohesion. Three systems were studied: a low-cohesion system composed of 50% Fast-Flo Lactose and 50% Avicel 102; a medium cohesion system composed of 50% Regular Lactose and 50% Avicel 102, and a high cohesion system composed of 50% Regular Lactose and 50% Avicel 101. In all cases, an aliquot of the system was laced with 6% micronized Acetaminophen, which was used as a tracer to determine the axial mixing rate in V-blenders of different capacities (1Q, 8Q, and 28Q).

Core-sampling was used to gather 35–70 samples per experimental time-point from 3 cores across each half of the blender. Samples were quantified using NIR spectroscopy, which was shown to be an accurate and efficient method for quantifying mixture quality. A simple model was used to determine mixing rates for both top/bottom and left/right loaded experiments. Variance measurements were split into axial and radial components to give more insight into mixing mechanisms and the separate effects of cohesion and vessel size on these mechanisms.

Convective mixing rates for radially segregated (top/bottom) loading were nearly constant regardless of changes in vessel size or mixture cohesion. Measured variances at short mixing times (i.e., 5 revolutions) were highly variable. These variations were attributed to unpredictable cohesive flow patterns during the first few rotations of the blender. An important conclusion was that scale-up of radial mixing processes could be obtained by simply allowing for a few (fewer than 10) "extra" revolutions to cancel this variability. As long as the shear limit was reached, the mixing rates was the same for all mixtures and vessel sizes, indicating that required mixing times (in terms of revolutions) needed to insure process outcome could be kept constant regardless of mixture cohesion or mixer size.

However, for axially segregated (left/right) loading, the scale-up factors depended on cohesion, indicating that scale-up is a mixture-dependent problem. As shown in Figure 3A, the most cohesive system mixed much more slowly in the smaller (1Q) blender. However, all three systems mixed at nearly the same rate in the larger (28 Q) vessel (Fig. 3B).

The conclusion from these results is that lab-scale experiments for cohesive powders are of questionable validity for predicting full-scale behavior. Behavior at small scales is likely to be strongly affected by cohesive effects that are of much less intensity in the large scale. Moreover, the density of the powder, and therefore the intensity of cohesive effects, might also depend on vessel size and speed.

An additional important comment is that the discussion presented in this section does not address another important cohesion effect: API agglomeration. As particles become smaller, cohesive effects grow larger. At some point, agglomeration tendencies become very significant. The critical factor in achieving homogeneity becomes the shear rate, which is both scale- and speed-dependent. This effect, which is familiar to the experienced formulator, occurs when a specific ingredient, typically the API, shows a tendency to agglomerate. In the authors' opinion, this problem is very common in direct compression applications, but has been rarely identified primarily due to the small number of samples typically used to characterize blends. Two situations should be distinguished: (*i*) agglomerates that do not reform once destroyed can be eliminated simply by implementing adequate "delumping" methods, preferably when loading ingredients to the blender, and (*ii*) agglomerates that form within the blender, and therefore pose a much more significant challenge. Here we only discuss the second case.

Several mechanisms drive the dynamic formation of agglomerates in a blender: (*i*) electrostatic charging, where polar materials can develop surface charges leading to aggregation, (*ii*) moisture transfer, where hygroscopic materials can sequester moisture from other ingredients and develop solid or capillary bridges with each other, and (*iii*) softening of MgSt or other low melting point ingredients, which can act as a glue to create "lumps" of non-polar ingredients. A full discussion of these effects would be beyond the scope of this chapter. Here, we limit our comments to three main observations:

1. In every instance known to the authors, this type of problem can be managed by judicious application of shear within the blender (i.e., use of an intensifier bar) or at the

discharge (passage through a mill) immediately prior to compression or encapsulation.
2. The most common scale-up criterion for the application of shear via impellers and I-bars is to match the linear speed of the moving element. It needs to be clearly understood, however, that while intuitively appealing, this criterion is scientifically untested.
3. Even when shear is used, dynamic agglomeration might re-surface. Thus, diagnostic of dynamic agglomeration is an exceedingly important issue. Combination of stratified sampling and multi-batch statistical analysis seeking to identify the presence of non-Gaussian super-potent tails in the composition distribution are a powerful method for monitoring the presence of agglomerates.

Summary

A systematic, generalized approach for the scale-up of granular mixing devices is still far from attainable. Clearly, more research is required both to test current hypotheses and to generate new approaches to the problem. Still, we can offer some simple guidelines that can help the practitioner wade through the scale-up process.

1. Make sure that changes in scale have not changed the dominant mixing mechanism in the blender (i.e., convective to dispersive). This can often happen by introducing asymmetry in the loading conditions.
2. For free-flowing powders, number of revolutions is a key parameter, but rotation rates are largely unimportant.
3. For cohesive powders, mixing depends on shear rate, and rotation rates are very important.
4. When performing scale-up tests, be sure to take enough samples to give an "accurate" description of the mixture state in the vessel. Furthermore, be wary of how you interpret your samples; know what the mixing index means and what your confidence levels are.
5. One simple way to increase mixing rate is to decrease the fill level—while this may be undesirable from a throughput point of view, decreased fill level also reduces that probability that dead-zones will form.
6. Addition of asymmetry into the vessel, either by design or the addition of baffles, can have a tremendous impact on mixing rate.

Until rigorous scale-up rules are determined, these cautionary rules are the "state of the art." The best advice is to be cautious—understand the physics behind the problem and the statistics of the data collected. Remember that a fundamental understanding of the issues is still limited and luck is unlikely to be on your side, hence frustrating trial-and-error is still likely (and unfortunately) to be employed.

Wet Granulation

Even more than blending, pharmaceutical granulation processes are still very much based on a batch concept despite efforts to switch to continuous manufacturing. The difficulty to fully embrace and implement continuous granulation throughout the pharmaceutical industry is often due to the challenging task of scaling up particulate processes. With the paradigm shift of moving towards "engineered particulate systems" in designing granular products, there is an increasing need for granules to possess certain physico-mechanical characteristics so that they can achieve the goal of enhancing product performance.

However, the sensitivity of particulate systems to scale and processing history makes them difficult to quantify, understand, model and control. Furthermore, characterization and identification of critical attributes must be achieved across several scales of scrutiny: micro- to meso- (bulk) to pilot- and finally full production scale. Consequently, modeling and simulation tools take on more integral and important roles in establishing the product–performance correlations across multiple scales.

Issues involved in the scale up of wet granulation processes were comprehensively addressed in a review by Mort (34,35). Some of the key points can be summarized as follows:

Concepts of dimensional similarity are often employed for scaling on the macroscale where the requisite operating conditions are determined over a range of dimensionally similar unit operations using dimensionless terms such as Froude, Stokes, and Reynolds numbers (36–40). Other commonly used dimensioned terms that can affect particulate growth processes include tip speed, swept volume and specific energy input.

However, the concepts of dimensional similarity are not without limitations. In fact, a classic example is one where the Froude number and tip speed cannot be kept constant as the impeller diameter increases. As the need to simultaneously maintain similarities in equipment shape and velocities, power input is not always possible and the choice of important factors to control becomes critical.

1. *Torque of the impeller blade* (41–45) *and power consumption* (46–48): Often used as parameters to determine the end point of wet granulation processes. Empirical adjustments are still required to achieve the desired granular product characteristics such as particle size and density.
2. *Specific energy*: This relates to the work done on the particulate system to bring it through the stages of granule formation. The net energy required in the agglomeration process is determined by integrating the net power draw over the residence time. When the net energy is expressed as a function of product mass, the specific energy is calculated. While this is an appealing approach, it is limited by the difficulty in determining the net powder draw which is used to bring about the agglomeration/coalescence process. It can, however, be estimated from the difference between the gross power draw and the baseline power consumption.
3. *Relative swept volume*: Defined as the volume of product swept away by the impeller blade in a given time, having considered the effects of product fill level, impeller speed and design. This idea is often combined with a modeling approach such as discrete element method (DEM) to measure the probability, frequency, and distribution of interactions between active mixing elements and product (34). A tight distribution of interaction frequency is desired to ensure that the amount of shear (energy) imparted to the product is uniform. The impact velocity and frequency can be used as a means to scale up coalescence and densification.

Modeling Techniques

Modeling techniques such as population balance, discrete element (DEM) and computational fluid dynamics are increasingly being applied to process simulations and control of continuous systems. It is common to have models with 20 variables, up to 200 variables can also be identified. Evidently, each model has its limitations and has yet to achieve complete validation. For instance, DEM requires mechanical properties of individual particles which can be difficult to determine. This, in turn, requires extrapolation from bulk calculations which can differ significantly between research groups. Moving forward, the continual refinement of modeling techniques and a combination of a

few of these still holds great promise in the accurate prediction of particle flow pattern, shear distribution, impact frequency and velocity for granulators of different scale.

Dry Granulation—Roller Compaction

As discussed, pharmaceutical scale-up is commonly thought of as the process by which batch size is increased. This can be accomplished by enlarging the physical dimensions from lab to pilot to plant scale or by increasing the output from a certain piece of equipment (2). Roller compaction is a unit operation that readily lends itself for scaled-up by either method. Through the use of continuous processing, larger batches of powders can be compacted using the same piece of equipment used for smaller scale batches by running for a longer period of time. The two main advantages of continuous processes are that ease of scale up for larger batches and a 24-hour automatic production line is possible (49). For example, a roller compaction process could be scaled-up using the WP 120 V Pharma roller compactor (50) from a 40- to 400-kg batch by running the compactor for 10 hours.

Ideally, when scaling up by enlarging the physical dimensions of the roller compactor from one production scale to another, the equipment should be similar geometrically, dynamically, and kinematically (49). The geometric condition is fulfilled when the ratio of physical dimensions between the small scale and the scaled-up version are constant. Dynamic similarity is seen when the ratio of forces exerted between matching points in the two roller compactors are equal. Finally, kinematic similarity is met when the ratio of velocities between matching points in both systems are equal (49). In reality, the scale-up process is more complicated because the equipment ratios between different scales may not match exactly. For instance, the WP 120 V Pharma roller compactor (50) is capable of running from 1g batches up to 40 kg/h, whereas the WP 200 C1 is capable of handling 100-kg batches up to 400 kg/h. These two roller compactors operate on the same operating principles and have the same design, thus making this scale-up a "level 1 equipment change" according to the Food and Drug Administration's (FDA) Scale-Up and Post Approval Changes guidance document for immediate-release solid oral dosage forms (SUPAC-IR) (3,9,51). Also, the increase in batch size from the WP120 V to the WP 200 C1 can be considered a "level 2 batch size change" due to the 100,000 fold increase in the batch capabilities and a "level 1 batch size change" with regards to the continuous manufacturing capabilities. Level 1 batch changes occur when the production batch is up to ten times larger than the pilot or bio batch size while a level 2 change occurs when the batch is greater than 10-fold for equipments operating on the same operating principles and design (3,9).

Apart from considering the physical dimensions, ratios of velocities, and ratios of pressures between two pieces of equipment of different scales, the design of roller compactors and their rolls are also important factors to consider in scale-up. According to the SUPAC-IR/MR-Manufacturing Equipment Addendum guidance (FDA), a level two equipment change only occurs when there is a change from one equipment class to another equipment class (9). One such example is the change from a dry granulator to a wet granulator even though this addendum classifies slugging and roller compaction together despite differences in their mechanism of powder densification. Although physics and finite element models have been investigated to describe the compaction process, none have yet been demonstrated to facilitate equipment or scale changes for practical purposes.

Even within the class of dry granulators, specifically roller compactors in this context, the direction of powder feed (vertical, horizontal, or angled) to the nip region varies among different equipment manufacturers with claimed advantages for each. Depending on the formulation, certain designs may be more suitable. A change of roller

compactor from one manufacturer to another requires a level 1 equipment change where application/compendial release requirements must be documented. Additionally, new batch records and long term stability results on the batches must be submitted to the FDA (3). Apart from the regulatory requirements, it is important to understand the effects of this change on the compacted ribbon and subsequently, the final dosage form. For example, horizontal feed roller compactors require formulations with higher levels of lubricant than vertical feed roller compactors to facilitate the compaction process. This change can, in turn, alter the hardness of the ribbons and resulting tablets.

Common rolls used in pharmaceutical roller compaction processes can be smooth, knurled fluted, knurled grooved and pocket design (52). Powders that are compacted using a smooth roll at lab scale may need to be compacted with knurled rolls on the pilot or manufacturing scale so that the powder can be gripped better, pulled through the nip region, and compacted by the rolls.

Compression

A typical problem of tableting scale-up is the loss of mechanical strength with increased speed. The strain rate sensitivity of viscoelastic and plastic materials is well documented (53–63). The resulting failure of tablets (Fig. 4) can classified as:

1. *Capping*: Due to release of elastic energy compared to a lesser increase of plastic energy and slow process of stress relaxation. It is often associated with air entrapment but this has been disputed in literature. Capping tendency is increasing with tableting speed (64,65), compression force, precompression force (66), punch penetration depth and tablet thickness (67).
2. *Lamination (tablet splits apart in single or multiple layers)*: Due to elastic recovery during decompression and ejection. Lamination is often blamed on over compressing—too much compression force flattens out the granules, and they no longer lock together. Lamination can also occur when groups of fine and light particles do not form enough interparticulate bonds during plastic deformation. Lamination tendency is increasing with speed, compression force and precompression force (68,69):
   a. Stress cracking—due to elastic recovery during ejection.
   b. Picking/sticking to punch faces—formulation, tooling and speed dependent.
   c. Chipping—may be caused by inadequate (brittle) formulation, take-off misalignment, and sticking.

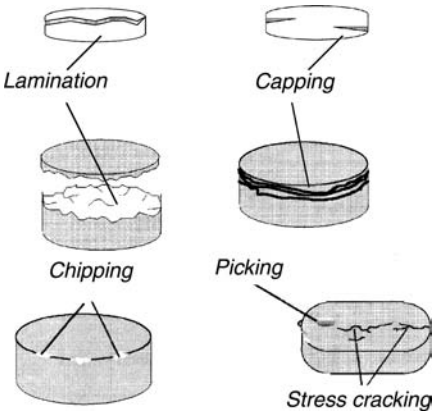


FIGURE 4 Tablet failure types.

Compression Factors

Apart from force and tooling that can be matched during scale-up or process transfer, the most important compaction factors are press speed and geometry. As the punch speed increases, so does the in-die temperature, friability, and porosity of tablets and their propensity to capping and lamination. The tensile strength of compacts tends to decrease with faster speeds, especially for plastic and viscoelastic materials, such as starch, lactose, avicel, ibuprofen, or paracetamol, as the rate at which the strain is applied and the duration both change. With the increase in porosity, one should expect a drop in disintegration and dissolution rates, but the interplay of the force-speed relationship may confound the effect. Although the energy absorbed by the tablet may not change, the power expended in the compaction process may decrease greatly with speed, and this, in turn, may have an effect on tablet properties. For the same linear speed of the press, tablets may be stronger if compression roll diameter is larger because this factor contributes to increase in consolidation and contact time.

Compression Time Events

Compression scale-up is generally governed by modeling principles that require geometrical, kinematic, and dynamic similarity of the physical process at different scales. Dimensional analysis of compaction process may lead to unified formulation-dependent theoretical equations that predict tablet properties on the basis of various processing factors (70). However, unlike all other unit operations in solid dosage development and production, scale-up of compression on a tablet press takes place in the same volume (die) using the same process geometry (tooling) and dynamic factors (compression force). The only practical differences between development and production conditions are press speed and the diameters of compression roll and die table (Table 1). In practical terms, compaction velocity and press geometry can be expressed and matched through characteristic process time components. The following times (Fig. 5) can be calculated on the basis of press speed and mechanical (geometric) parameters (71):

- Consolidation (solidification) time, T_s, is the time when punches are changing their vertical position in reference to the rolls, decreasing the distance between the punch tips.
- Dwell time, T_d, is the portion of the time when punches are not changing their vertical position in reference to the rolls.
- Decompression (relaxation) time, T_r, is the time when punches are changing their vertical position in reference to the rolls, increasing the distance between the punch tips before losing the contact with the rolls.
- Contact time, T_c, is the time when both punches are moving having their tips in contact with the material that is being compacted, and their heads are in contact with the compression rolls: $T_c = T_s + T_d + T_r$.
- Ejection time, T_e, is the time when the tablet is being ejected from the die.
- Total time, T_t, is the time required to produce one tablet on a press (including time between tablets).

It may be noted here that peak of compression force precedes the mid-point of dwell time because of the stress relaxation due to plastic flow for plastically deforming materials (the so-called peak offset time). It is this time during "quasi-constant" strain conditions that makes dwell time such an important factor in compaction process. Other scale-up considerations include feeding time, instrumentation grade, measurement of speed and mechanical strength, and variations in tooling, powder flow, raw materials,

TABLE 1 Similarity Factors in Tableting Scale-Up

Similarity	Production press vs. R&D press
Geometric similarity	
Die	Same
Upper punch	Same
Lower punch	Same
Turret	Different
Upper compression roll	Different
Lower compression roll	Different
Kinematic similarity	
Punch velocity Linear (horizontal, tangential), V_h Average vertical, V_v Maximum vertical	Can be matched in a limited range, depending on press speed and geometry
Punch acceleration Average vertical A_v	Can be matched in a limited range
Critical compaction times Consolidation time T_s Dwell time T_d Relaxation time T_r Contact time $T_c = T_s + T_d + T_r$	Can be matched in a limited range, depending on V_h and diameter of turret and compression rolls
Dynamic similarity	
Applied force	Can be matched

variation and tablet weight. Critical compaction times reflect differences in press speed and geometry. Consolidation and dwell time parts of the compaction cycle (during the "rise-time" of the force–time profile) is 6–15 times more important than the decompression part as a factor contributing to capping and lamination (69,72–74). It stands to

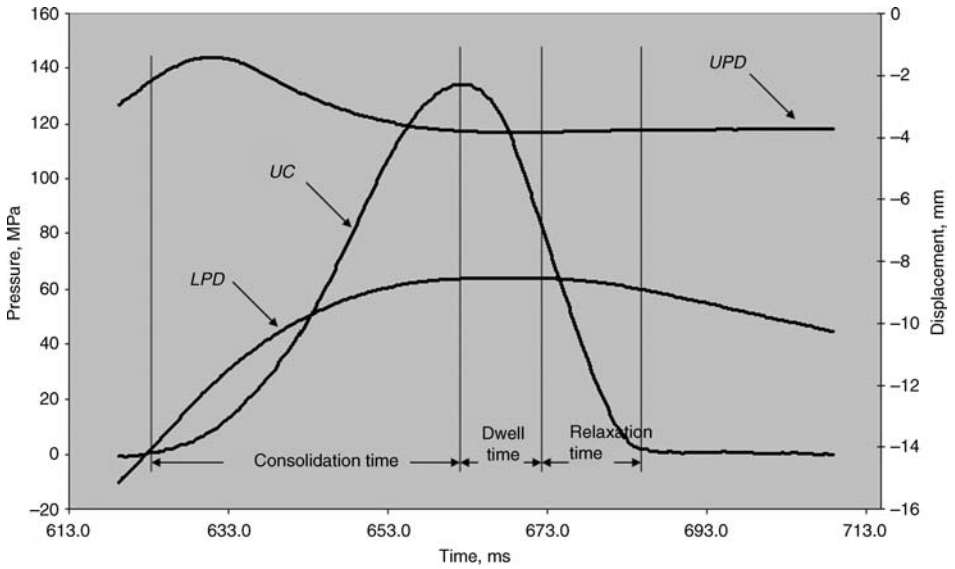


FIGURE 5 Time events in compaction. *Abbreviations*: UC, upper compression; UPD, upper punch displacement; LPD, lower punch displacement.

reason, therefore, to attempt to match $T_s + T_d$ as the most significant factors in compaction scale-up.

Compression Scale-Up: A Practical Example

As a practical example, consider a problem of scaling-up a perfect formulation from 16-station Manesty Betapress to 36-station Korsch PH336, or 36-station Kikusui Pegasus 1036, or 37-station Fette P3000. Let us say that the formulation was based on a wet granulation of brittle API, Avicel PH102, and 0.5% MgSt. The ideal tablet was made at 10 kN compression force, Betapress speed of 50 RPM, with TSM B 3/8 in. round flat tooling, 10 mm depth of fill, and the resulting out of die tablet thickness was 5 mm. Under these conditions, one may attempt to match $T_s + T_d$ on the target presses as seen in Table 2.

It turns out that both the Korsch and Kikusui presses have to operate at the lowest end of their speed range, while the Fette is not slow enough to reach the required (slow) speed. If the Fette is preferred, the Betapress speed should be increased up to at least 60 RPM (Table 3).

A maximum speed of an R&D press can barely reach half the range of production press speed in terms of $T_s + T_d$ (e.g., $T_s + T_d = 24$ ms for maximum Betapress time at 104.2 RPM, which corresponds to 51.3 RPM on Fette 2090 or 41.4 RPM on Fette 3000). Therefore, the best way to eliminate scale-up problems without limiting the production outputs would be to develop your formulation using a high-speed compaction simulator. Such devices attempt to mimic compaction profiles of any press with the obvious benefit of forecasting formulation behavior under the production conditions.

Effect of Shear and Strain on Material and Product Properties

Important variables seldom taken into account during scale up are the shear rate and the total strain experienced by the material during processing (75). It has been known that excessive shear applied to a pharmaceutical blend for a significant amount of time decreases hardness, increases capping and decreases dissolution of subsequently compressed tablets. For direct compression cohesive blends, intensity of applied shear also

TABLE 2 Matching $T_s + T_d$ for Manesty Betapress at 50 RPM

Tablet Press	Stations	RPM	TPH	T_s	T_d	$T_s + T_d$
Manesty Betapress	16	50.0	48,000	42.1	15.5	57.6
Korsch PH336	36	33.4	72,169	44.6	13.0	57.6
Kikusui Pegasus 1036	36	34.8	75,230	42.6	15.0	57.6
Fette P3000	37	30.0	133,200	36.7	11.7	48.4

TABLE 3 Matching $T_s + T_d$ for Manesty Betapress at 60 RPM

Tablet Press	Stations	RPM	TPH	T_s	T_d	$T_s + T_d$
Manesty Betapress	16	60.0	57,600	35.1	13.0	48.1
Korsch PH336	36	40.1	86,603	37.2	10.8	48.0
Kikusui Pegasus 1036	36	41.8	90,277	35.5	12.5	48.0
Fette P3000	37	30.2	134,112	36.4	11.6	48.0

affects particle size and shape, the density, flowability, and content uniformity of powder, and weight variation of the resulting tablet. Finally, total applied shear correlates directly to electrostatic charging of the blend, which is both a safety hazard and a process nuisance. However in spite of its significant impact, shear has not been studied systematically. Typically, shear is applied (often unintentionally) both in the blender and in feed frame. In both these environments the granular flow is poorly understood and we do not know either the intensity or the uniformity of shear that is imparted to the system. As a result, knowledge of shear effects is only qualitative, and no guidelines exist for controlling the amount of shear needed by a given blend or applied in a given system.

In order to carefully examine this issue, a novel "controlled shear environment" (75) was developed in collaboration between Rutgers and MCC, and was used it to study homogenization of MgSt under carefully controlled, homogeneously applied shear rates. The device, shown in Figure 6, is capable of imposing known amounts of shear *homogeneously and at a controlled rate*, making it possible to design experiments where the relationship between measured forces and observed flow and mixing phenomena is clear (Fig. 6). The device is an annular Couette flow cell, which is essentially two concentric cylinders separated by a narrow annular gap. Both cylinders are supplemented with equally spaced interlocking pins in order to achieve a homogeneous shear field in the flow region. Samples weighing approximately up to 1 kg can be exposed to different shear intensities for controlled periods, thus providing an ideal environment for investigating the effect of shear on tablet hardness, dissolution, density, and flow properties.

Experiments were performed in order to examine the effect of total shear and MgSt content on blend flow properties, MgSt homogeneity, bulk density and tablet hardness, using a blend of 58–60% Fast-flo lactose, 40% Avicel 102, 0–2% MgSt. Blends were sheared at various rates in the range from 10 to 245 RPM (corresponding to shear rates between 1.25 and 300 s^{-1}) for a total of 10–2000 revolutions corresponding to 750–150000 total dimensionless shear units), and were subsequently sampled. Bulk density, flow properties, and rate of water uptake by sheared blends were subsequently characterized. Moreover, selected samples were compressed under conditions simulating operation of commercial presses, and the tablets were then tested for crushing hardness.

Figure 7 shows the bulk density of the resulting samples. The bulk density increases and then reaches a plateau, indicating that the cohesion of the blend is diminishing (flowability is increasing) as a result of the applied strain.

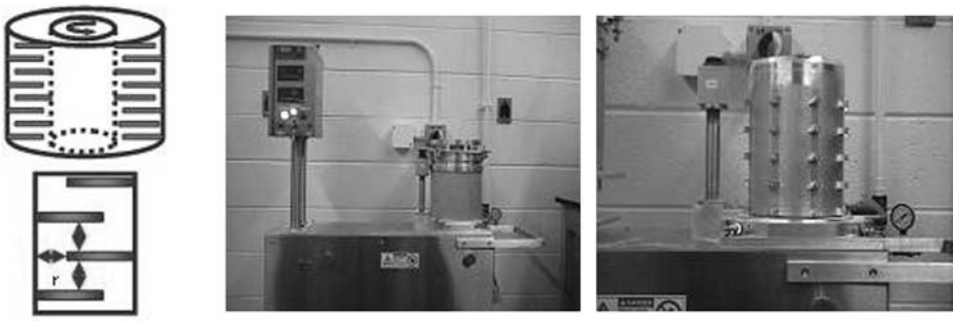


FIGURE 6 The figure shows the schematic and actual picture of the shear instrument. The inner cylinder rotates at a constant speed transmitting shear to the blend in a controlled and uniform fashion. The rheometer displays the total torque, rotation speed and can be attached to a computer to get continuous data.

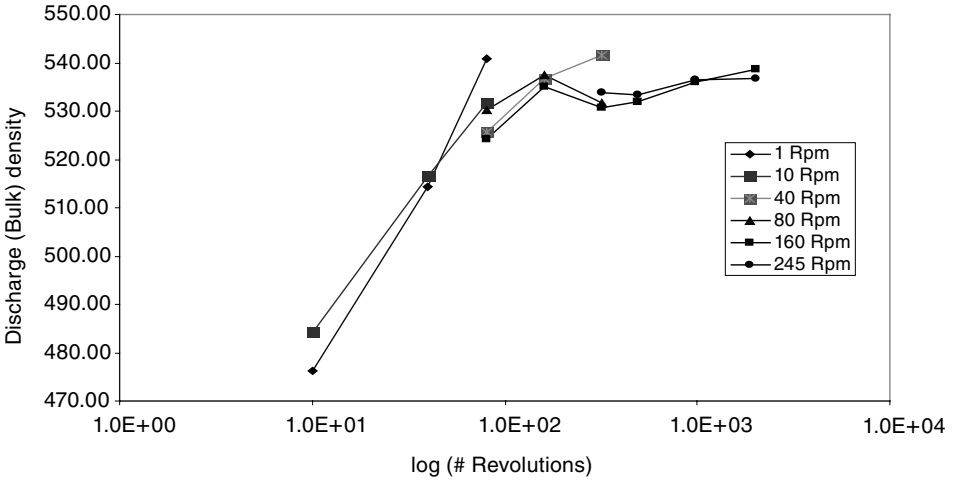


FIGURE 7 The figure shows the effect of total shear on the discharge bulk density of the mixture: 59% Fast Flo Lactose, 40% Avicel 102 and 1% MgSt. The bulk density increases as the total shear is increased and finally reached a constant value. *Abbreviation*: MgSt, magnesium stearate.

Tablet hardness is consistently and reproducibly affected by the total amount of shear imposed on the blend. Figure 8 demonstrate how the hardness of tablets made by MCC Presster, strongly depends not only on the MgSt concentration (as expected) but also on the level of shear. The effect of total shear on tablet hardness (Fig. 8) is determined by shearing three samples of identical composition (1% MgSt) at low, medium and high total shear. The results show a decrease in hardness as the total shear is increased.

Finally, and perhaps most importantly, the hydrophobicity of blends of constant composition is dramatically affected by the total strain applied to blends of constant

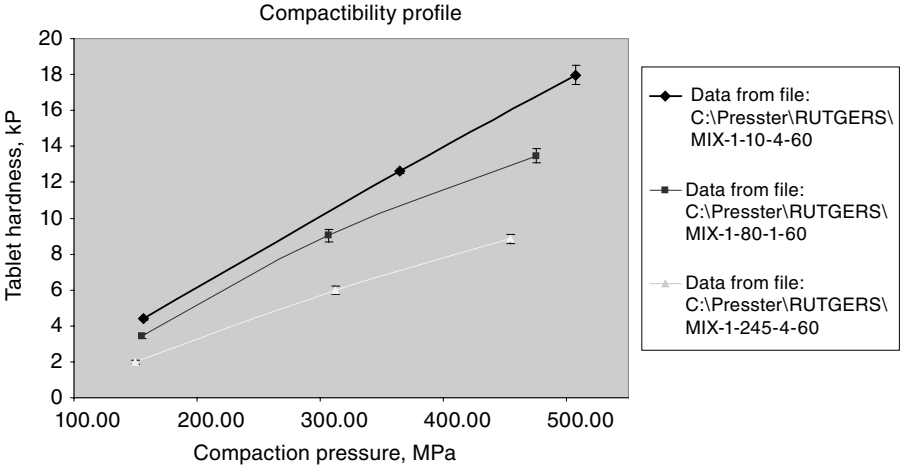


FIGURE 8 The figure shows the tablet crushing hardness of mixtures sheared to three different levels of total shear (3000 shear units, 6000 shear units, 73500 units) in the device. As shear increases a marked decrease in tablet hardness is observed. Simulated press: Fette PT3090 61 station at 60 RPM.

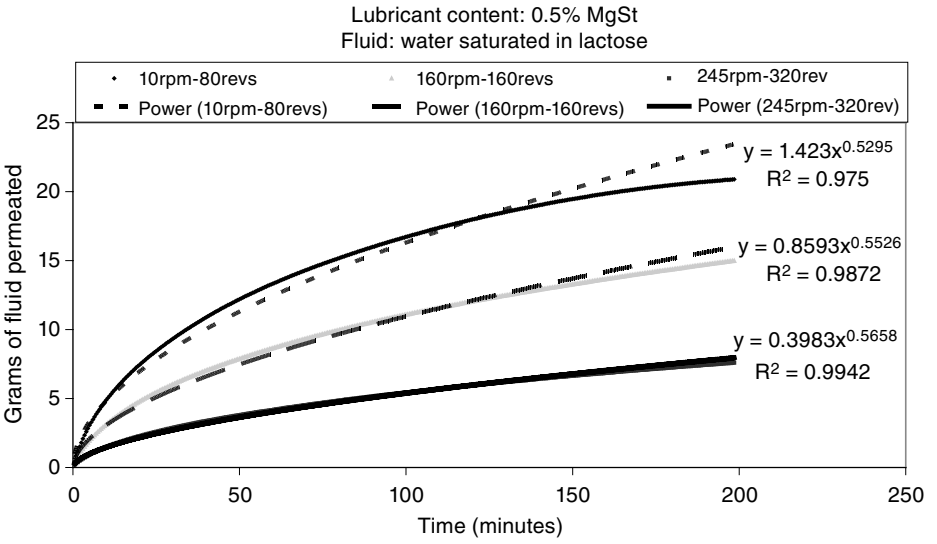


FIGURE 9 The figure shows that sheared blends become increasingly hydrophobic as the total strain imposed on them increases. The rate of uptake of lactose-saturated water by a blend of Lactose, Avicel, and 0.5% MgSt decreases nearly three fold when the total strain is increased from 80 revolutions to 320 revolutions in the controlled shear device. Even more extreme changes are measured at higher concentrations of MgSt. *Abbreviation*: MgSt, magnesium stearate.

composition. This was demonstrated by packing the strained blends inside a glass column, and putting them in contact with a solution saturated in Lactose (the only readily dissolvable ingredient present in the blend). Changes in surface tension can be quantified by measuring the rate of fluid uptake by the powder column. When the powder is hydrophilic, the solution readily penetrates the powder. However, for strained blends, the rate of fluid uptake greatly diminishes, demonstrating that the strained blend has became substantially more hydrophobic.

These results demonstrate that the properties of both blends and finished products depend strongly on shear and strain, highlighting the need for taking into account these variables during process scale up.

EMERGING APPROACHES—CONTINUOUS PROCESSING—SCALE-UP BY TIME EXTENSION

General Comments

In the batch manufacturing practices currently used for most pharmaceutical products, the entire batch is mixed at once and subsequently it is compressed into tablets (or encapsulated). The two most common problems affecting the quality of the finished product, segregation and agglomeration, are often made worse by the usual batch approach. If the material segregates, as is often the case with free-flowing systems, then the entire mixture is exposed to the segregation process, often resulting in a batch with large variability in composition. In this situation, the "scale of segregation" of the mixture is as large as possible, i.e., the same size as the entire batch. Batch manufacturing is also a bad idea for mixtures that agglomerate. The situation can be particularly complicated for low-dose

direct compression products, which represent an industry-wide trend for newer products. Low dose in practice means that even small fluctuations in composition can strongly affect the statistical homogeneity of the finished product. As actives become increasingly potent and particle sizes decrease, the actives become increasingly cohesive. As a result, the finer cohesive particles will have an increased tendency to agglomerate, resulting in a smaller effective number of larger particles, which can increase the statistical fluctuations in active content. Intense shear is required to comminute cohesive actives and disperse them within the larger bulk of the mixture. Unfortunately, it is nearly impossible to apply shear efficiently and uniformly in large-scale batch equipment, which often results in the survival or re-forming of agglomerates and, consequently, fluctuations in finished product content. In addition, the current "large batch" approach to blending requires an entirely empirical and therefore risky scale-up protocol between the lab, the pilot plant, and the manufacturing facility.

Continuous processing has several additional potential advantages for Pharmaceutical manufacturing. Most germane to this chapter, continuous manufacturing methods *enormously simplify development and scale-up*, because processes can be developed using *the same devices* that will later be used in the manufacturing operation. Process scale-up is achieved by running the equipment for longer times (rather than in larger systems). Technology transfer only requires a lateral 1:1 migration from the laboratory to the production plant, greatly eliminating scale-up uncertainties and further reducing development times. Continuous processes are controlled with respect to a stationary set point, which greatly facilitates modeling and control of the manufacturing process. The accumulated knowledge concerning process linearization and control can be immediately applied to pharmaceutical manufacturing processes to minimize deviations from desired outcomes.

Due to the dramatic reduction in the scale of the blending operation and the possibility of integrating blending and compression (or encapsulation) into a single processing step, the proposed approach greatly decreases the facilities cost of manufacturing.

Finally, continuous approaches significantly change the approach to sampling. Since the process takes place in thousands of small-scale overlapping operations, conventional sampling for batch acceptance is no longer a suitable option. One would only need to monitor the feed rate, which can be done gravimetrically, and the composition of the output (i.e., tablets). Thus, the proposed manufacturing process provides the ideal environment for implementation of PAT methods. In fact, PAT is the only suitable approach for on-line and at-line monitoring.

Interestingly, continuous processing has been utilized extensively by petrochemical and chemical manufacturing. Recent research efforts indicate that a well-controlled continuous mixing process can significantly enhance productivity (76,77). Previous reviews on continuous mixing of solids (78,79) point to the fact that a batch system that can be run in continuous mode can be expected to possess similar mixing mechanisms. This is because in continuous blending systems, a net axial flow is superimposed on the existing batch system to yield a continuous flow. Continuous mixing has also been studied for Zeolite rotary calciners (80), chemical processes (SiC or Irgalite and $AL(OH)_3$) (81), food processes (Couscous/Semolina) (76), and a pharmaceutical system (CaCO3—Maize Starch) (82). The effictiveness of continuous mixing was studied by Williams and Rahman (78) with a salt/sand formulation of different compositional ratios. Williams (83) examined the mixing performance of the drum speed using variance reduction ratio (VRR) of unspecified solids. The VRR was used in a paper written by Weinekötter and Reh (84) to observe how purposely-fluctuating tracers into the processing unit were depressed. Harwood et al. (85) studied the performance of seven

continuous mixers as well as the outflow sample size effect of sand and sugar mixtures. Although no simple correlations were generated, they investigated the mixing performance of different convective mixers and sample sizes. Others have focused on the flow patterns formed by the different convective mechanisms within horizontal mixers. Laurent and Bridgwater (86) examined the flow patterns by using a radioactive tracer, which generated the axial and radial displacements as well as velocity fields with respect to time. Marikh et al. (76) focused on the characterization and quantification of the stirring action that takes place inside a continuous mixer of particulate food solids where the hold up in the mixer was empirically related to the flow rate and the rotational speed.

PAT as a Required Component of Continuous Processes

Development of PAT approaches (i.e., process understanding married to rational monitoring and control) for process scale-up is likely to take place at several levels. At the conceptually simplest level, PAT pre-supposes the development of sensing instruments capable of monitoring process attributes online and in real time for control. Once the analytical method is validated for accuracy at the laboratory scale, it can be used to obtain extensive information of process performance (blend homogeneity, granulation particle size distribution, moisture content) under various conditions (blender speed, mixing time, drying air temperature, humidity, and volume, etc.). Statistical models can then be used to relate the observable variables to other performance attributes (e.g., tablet hardness, content uniformity, and dissolution) in order to determine ranges of measured values that are predictive of acceptable performance.

Typically, for batch processes such as blending or drying, this entails the determination of process end-point attributes. The PAT method then becomes the centerpiece of the scale-up effort. Process scale-up can be undertaken under the assumption that the relationships between observables and performance are independent of scale, and if this assumption is verified in practice, the manufacturing process in full scale can be monitored (typically, to completion) providing a higher level of assurance that the product is likely to be within compliance. Control variables (variables that may be adjusted in near real-time) can then be manipulated within limits or between batches to maintain the desired quality attributes of the product.

For continuous or semi-continuous processes (such as tablet compression), the main role of PAT methods is not process end-point determination, rather, it is to serve as a component of a feed-back or feed-forward control strategy devoted to keeping process (and product) performance within the desired range along the life of the process. This is conceptually more complex and requires a greater level of predictive understanding regarding the dynamic effect of controlled variables on performance attributes (see below). However, once the development of suitable controls is achieved, scale-up itself is greatly simplified for continuous (or semi-continuous) processes, which typically involves running the process for longer times.

At a more sophisticated level of articulation, PAT will involve the use of analytical methods, coupled with modeling approaches, to develop models capable of predicting quantitatively the relationship between input parameters (raw materials properties, process parameters, environmental inputs) and product performance (so called "model predictive control"). In the authors' opinion, this is the true definition of "process understanding". On an early stage, models can be statistical (correlation-based), seeking only to determine directional relationships and co-variances. Over time, predictive mathematical models can be developed once mechanistic relationships between inputs and outputs are established.

Predictive models make it possible to perform true process scale-up, which consists of the use of a predictive model to find quantitative criteria for establishing process similarity across scales. The model is also used to determine the changes in both the design space and the target function across scales, and to predict optimum conditions of manufacturing facilities yet to be built.

Even more, a predictive model allows the designer to explore before hand the effect of uncertainty in raw material properties (and other input variables not controllable in real-time), market conditions, and regulatory constraints, thus making it possible to design flexible manufacturing systems that have built-in capabilities for accommodating changing conditions. The methodology, known as "design under uncertainty" is currently an active area of research in the systems engineering community.

A Case Study: Continuous Mixing

This case study discusses the effects of operating conditions and design parameters on the mixing efficiency using blend formulations that contain Acetaminophen as an example of a pharmaceutical product. Effects of design parameters such as blade design and operating conditions such as rotation rate, the processing angle, and the powder cohesion on the mixing performance are discussed.

Apparatus

The continuous blender device used in the case study is shown in Figure 10. The mixer has a 2.2 KW motor power, rotation rates range from 78 revolutions per minute (RPM) at a high speed to 16 RPM at a low speed. The length of the mixer is 0.74 m and the diameter is 0.15 m. An adjustable number of flat blades are placed within the horizontal mixer. The length of each blade is 0.05 m and the width is 0.03. Convection is the primary source of mixing, the components have to be radially mixed which is achieved by rotation of the impellers (84). The convective forces arising from the blades drive the powder flow. As the blades rotate, the powders are mixed and agglomerates are broken

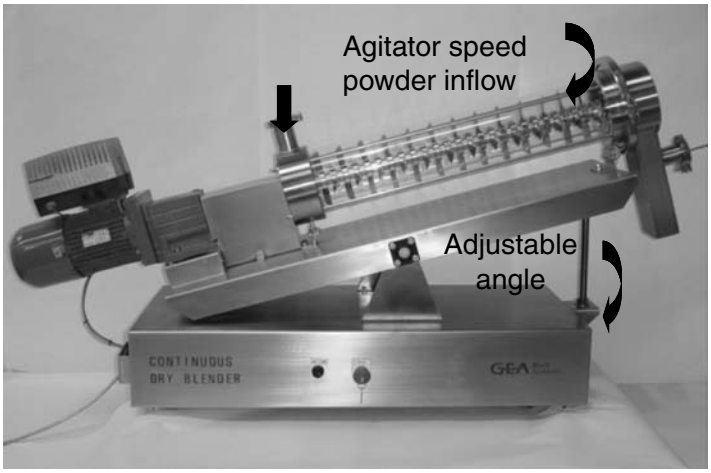


FIGURE 10 A photograph of the continuous powder mixer used in the case study described in this chapter.

up. The powders are fed at the inlet and removed from the outlet as illustrated in Figure 10. The powder is discharged through a weir in the form of a conical screen. This feature ensures that the agglomerates are hindered from leaving the mixer. Thus, by varying the mesh of this screen, different degrees of micro-homogeneity can be accomplished. The particulate clusters become lodged in the screen, were they are broken up by the last impeller, the one closest to the outflow, before departing the blender. The powder ingredients are fed using two vibratory powder feeders. The two vibratory feeders (Eriez) feed powder directly into the mixer inlet. Built-in dams and powder funnels were used to further control the feed rate of each feeder.

Blend Formulations

Case studies consist of one active and one excipient. Model blends were formulated using the following materials: DMV Ingredients Lactose (100) (75–250 μm), DMV International Pharmatose® Lactose (125) (55 μm), and Mallinckrodt Acetaminophen (36 μm). The compositions of the formulations used are as follows: Formulation 1: 3% Acetaminophen, 97% Lactose 100. Formulation 2:3% Acetaminophen, 97% Lactose 125. The formulation is split into two inflow streams both at the same mass flowrate. One flow stream supplies a mass composition of 6% Acetaminophen and 94% of Lactose and the other stream consists entirely of 100% Lactose. Both feeders are identical and process powders with a total a mass rate of 15.5 g/s with a standard deviation of 2.53 g/s. After the feed is processed, the material entering the mixer should contain: 3% Acetaminophen and 97% Lactose.

Mixer Characterization

Two methods are used to characterize the system, the residence time and the degree of homogeneity as described in the next sections.

The residence time distribution is an allocation of the time that different elements of the powder flow remain within the mixer. To determine the residence time distribution, the following assumptions are made: (*i*) the particulate flow in the vessel is completely mixed, so that its properties are uniform and identical with those of the outflow; (*ii*) the elements of the powder streams entering the vessel simultaneously, move through it with constant and equal velocity on parallel paths, and leave at the same time. In this study the residence time is measured as follows:

1. A quantity of a tracer substance is injected into the input stream; virtually instantaneous samples are then taken at various times from the outflow.
2. After the injection, the concentrations of the injected material in the exit stream samples are analyzed using Near Infrared (NIR) Spectroscopy. Sample concentrations are expected to change since the tracer is fed at one discrete time point and not continuously.

The residence time distribution is determined both as a function of time and number of blade passes. The average number of blade passes is used to measure the shear intensity the powder experiences and its effect on blending. The mean residence time is determined using the mass-weighted average of the residence time distribution.

Homogeneity of the output steam is determined by analyzing a number of samples retrieved from the outflow as a function of time. The samples are analyzed to calculate the amount of tracer (in our case Acetaminophen) present in the sample using NIR Spectroscopy. The homogeneity of samples retrieved from the outflow is measured by calculating the variability in the samples tracer concentration. The RSD of tracer

concentration measures the degree of homogeneity of the mixture at the sample. Lower RSD values mean less variability between samples, which implies better mixing. Another important characteristic of the mixer is to what extent variability of feed composition can be eliminated within the unit. In order to measure this characteristic, the VRR is used, which is defined as the ratio of the inflow variance calculated from samples collected at the entrance of the mixer to the outflow variance. Both variances are calculated collecting samples from the inflow and outflow of the mixer. The larger the VRR, the more efficient the mixing system, since inflow fluctuations are reduced. As will be shown in the next section, both metrics (RSD and VRR) lead to the same conclusion regarding which parameters result in better mixing performance.

Effect of Design, Operational, and Material Parameters

The blender has two main design parameters, the number of blades and blade angle, and two operating parameters, processing angle and impeller rotation rate, which affect the shear intensity and powder transport. In addition, powder density and cohesion (among several other variables) also have an impact on flow and mixing. The mixer's function is to simultaneously blend two or more inflow streams radially as the powder flows axially. Choosing the right design parameters, and adjusting the mixers operational parameters, for a given set of material parameters is critical to the system performance. Here, we provide a brief summary of main observations (87).

It is critical to the system performace to choose the right design parameters and adjusting the mixers operational parameters

Number of Blades: Two blade configurations were compared, one having 29 blades, and the other one having 34 blades. For the smaller number of blades, "dead regions" were observed where the powder remained stagnant; samples taken from these locations revealed a large concentration of API. The higher number of blades allowed us to minimize the formation of stagnant zones in the mixer and to increase the intensity of transport mechanisms in the axial direction.

Blade Angle: Another important convective design parameter investigated is the blade angle, which affects powder transport (88). The purpose of the impeller is to propel the powder within the vessel. The motion of the particulates is affected by the blade angle. Varying the blade angle affects the particle's spatial trajectory, thus altering the radial and axial dissipation. Laurent and Bridgwater (88) illustrated that increasing the blade angle promoted additional dispersion forces leading to increasing radial mixing. Five blade angles examined were 15°, 45°, 60°, 90°, and 180°. It was observed that the RSD of the outflow stream was the highest for the lower 15° angle followed by the 45° angle design, and the lowest at the higher 60° angle. Performance collapsed when increasing the angle to (and beyond) 90°.

Processing Angle: Since axial flow is affected by adjusting the processing angle, it is reasonable to assume that the residence time (and residence time distribution) will also be affected. The residence time distribution of Acetaminophen was determined for three processing angles and two rotation rates. The main result observed was that as the processing angle increased to an upward angle of 30°, the residence time increased, RTD became narrower, and RSD and VRR both decreased for all speeds and for both formulations.

Blender Speed: For the two formulations studied here, it was observed that as the speed of the blender increased, the residence time of the API first decreased, and then became constant, indicating that the total level of strain experienced by the API would be higher at higher RPM. The Residence time distribution was much wider at lower speeds

when measured in terms of clock time, but differences were actually minimal when measured in terms of blade passes. Finally, and contrary to our expectations, for the materials examined here, better homogeneity was observed at lower RPM.

Powder Cohesion: Two grades of Lactose varying in particle size, Lactose 100 (130 μm) and Lactose 125 (55 μm), were utilized to examine the effect of the blend cohesion. Surprisingly, decreasing the particle size did not affect the mixing performance of the process at either low or high speed.

SUMMARY AND CONCLUSIONS

While it is a well established *cliché* to end a document such as this by stating that "much remains to be done," this is certainly the case for the QbD methodology in general, and for its applications to process scale up in particular. That said, it might be useful, perhaps, to identify exactly where we are likely to obtain the greatest rate of return on invested efforts:

1. A better understanding of material properties of ingredients and intermediate streams and their impact on process and product performance is clearly at the top of the list. This understanding is a required precondition to the development of instrumental chemistry methods (i.e., sensors, chemometric algorithms, etc.). Without such an understanding, many material variables will go unmeasured simply due to a lack of awareness of their importance.
2. Equal in importance is to develop a deeper predictive understanding of process components, both those discussed here and those that were left out. These process components are mainstays of pharmaceutical manufacturing and will continue to determine process outcome for many years to come.
3. More subtle, but equally critical, is the need to understand process interactions. It is a truism that changes introduced to improve a given stage of the manufacturing process often affect (adversely) the performance of other downstream stages. Many such problems can be avoided, or mitigated, if these interactions along the production sequence are better understood.
4. Finally, while much progress has been achieved by regulatory agencies and by industry in modernizing the conceptual content of the regulatory framework, quite a bit of work remains to be done before the drug approval and licensing process is truly enabling, and supportive, of true process improvement efforts along the product life cycle.

While the full development and implementation of the scientific, educational, and regulatory infrastructure needed to improve pharmaceutical product and process design and optimization will take sustained efforts over many years, the authors believe that the technological, economical, and quality benefits will be clearly enormous, in particular for those companies leading the charge.

REFERENCES

1. Muzzio FJ, Shinbrot T, and Glasser BJ. powder technology in the pharmaceutical industry: the need to catch up fast. Powder Tech 2002; 124:1.
2. Levin M, ed. Pharmaceutical Process Scale-Up. New York: Marcel Dekker, 2002.
3. SUPAC-IR: Immediate Release Solid Oral Dosage Forms Scale-Up and Postapproval Changes: Chemistry, Manufacturing, and Controls, *in vitro* Dissolution Testing, and *in vivo*

Bioequivalence Documentation, 1995. (Accessed June 7, 2007, at www.fda.gov/cder/guidance/cmc5.pdf).
4. SUPAC-IR Questions and Answers about SUPAC-IR Guidance, 1997. (Accessed June 7, 2007, at http://www.fda.gov/cder/guidance/qaletter.htm).
5. SUPAC-SS: Nonsterile Semisolid Dosage Forms. Scale-Up and Postapproval Changes: Chemistry, Manufacturing, and Controls; In Vitro Release Testing and In Vivo Bioequivalence Documentation 1997. (Accessed June 7, 2007, at www.fda.gov/cder/guidance/1447fnl.pdf).
6. SUPAC-MR: Modified Release Solid Oral Dosage Forms. Scale-Up and Postapproval Changes: Chemistry, Manufacturing, and Controls; In Vitro Dissolution Testing and In Vivo Bioequivalence Documentation, 1997. (Accessed June 7, 2007, at www.fda.gov/cder/guidance/1214fnl.pdf).
7. SUPAC-ER: Extended Release Oral Dosage Forms: Development, Evaluation, and Application of In Vitro/In Vivo Correlations, 1997. (Accessed June 7, 2007, at www.fda.gov/cder/guidance/1306fnl.pdf).
8. SUPAC-SS: Nonsterile Semisolid Dosage Forms Manufacturing Equipment Addendum, 1998. (Accessed June 7, 2007, at www.fda.gov/cder/guidance/1722dft.pdf)
9. SUPAC-IR/MR: Immediate Release and Modified Release Solid Oral Dosage Forms Manufacturing Equipment Addendum, 1999. (Accessed June 7, 2007, at www.fda.gov/cder/guidance/1721fnl.pdf).
10. Changes to an Approved NDA or ANDA, 1999. (Accessed June 7, 2007, at www.fda.gov/cder/guidance/2766fnl.pdf).
11. Waiver of In Vivo Bioavailability and Bioequivalence Studies for Immediate-Release Solid Oral Dosage Forms Based on a Biopharmaceutics Classification System, 2000. (Accessed June 7, 2007, at www.fda.gov/cder/guidance/3618fnl.pdf)
12. PAT—A Framework for Innovative Pharmaceutical Manufacturing and Quality Assurance, 2004. (Accessed June 7, 2007, at www.fda.gov/cder/guidance/6419fnl.pdf)
13. Hicks CR and Turner KV. Fundamental Concepts in the Design of Experiments. New York: Oxford University Press, 1999.
14. Horst R, Pardalos PM, Nguyen VT. Introduction to Global Optimization. New York: Springer-Verlag, 1995.
15. Floudas CA and Pardalos PM. A collection of test problems for constrained global optimization algorithms. New York: Springer-Verlag 1990.
16. Muzzio FJ, Robinson P, Wightman C, Brone D. Sampling practices in powder blending. Int J Pharm 1997; 155:153–78.
17. Technical Report No. 25 Blend uniformity analysis: validation and in-process testing. J Pharm Sci Tech 1997; 51(S3).
18. Carstensen JT and Patel MR. Blending of irregularly shaped particles. Powder Tech 1977; 17:273–82.
19. Adams J and Baker A. An assessment of dry blending equipment. Trans Instit Chem Eng 1956; 34:91–107.
20. Brone D, Alexander A, Muzzio FJ. Quantitative characterization of mixing of dry powders in V-blenders. AIChE J 1998; 44(2):271–78.
21. Brone D and Muzzio FJ. Enhanced mixing in double-cone blenders. Powder Tech 2000; 110(3):179–89.
22. Wiedenbaum SS. Mixing of solids in a twin shell blender. Ceramic Age 1963; 39–43.
23. Wightman C and Muzzio FJ. Mixing of granular material in a drum mixer undergoing rotational and rocking motions i. uniform particles. Powder Tech 1998; 98:113–24.
24. Carley-Macauly KW, Donald MB. the mixing of solids in tumbling mixers-I. Chem Eng Sci 1962, 17:493–506.
25. Carley-Macauly KW, Donald MB. 1964, The mixing of solids in tumbling mixers-II. Chem Eng Sci 1964; 19:191–9.
26. Sethuraman KJ, Davies GS. studies on solids mixing in a double-cone blender. Powder Technol 1971; 5:115–8.

27. Poux M. Powder mixing: some practical rules applied to agitated systems. Powder Technol 1991; 68:213–34.
28. Lacey PMC. Developments in the theory of particulate mixing. J Appl Chem 1954; 257–68.
29. Sudah OS, Coffin-Beach D, Muzzio FJ. Quantitative characterization of mixing of free-flowing granular material in tote (bin)blenders. Powder Tech 2002; 126:166–73.
30. Wang RH, Fan LT. Methods for Scaling-Up Tumbling Mixers. Chemical Engineering. 1974; 81(11):88–94.
31. Lloyd PJ, Yeung PCM, Freshwater DC. The Mixing and Blending of Powders. J Soc Cosmetic Chemists 1970; 21:205–20.
32. Roseman B, Donald MB. mixing and de-mixing of solid particles: Part 2: Effect of varying the operating conditions of a horizontal drum mixer. British Chem Engg 1962; 7(1):823.
33. Wiedenbaum SS, Mixing of Solids. In T.B. Drew TB, Hoopes JW, eds. Advances in Chemical Engineering. New York: Academic Press, 1958:209–324.
34. Mort PR. Scale-up of binder agglomeration processes. Powder Tech 2005; 150:86–103.
35. Faure A, York P, Rowe RC. Process control and scale-up of pharmaceutical wet granulation processes: a review. Eur J Pharm Biopharm 2001; 52(3):269–77.
36. Landin M, York P, Cliff MJ, Rowe RC, Wigmore AJ. The effect of batch size on scale-up of pharmaceutical granulation in a fixed bowl mixer-granulator. Int J Pharm 1996; 134: 243–6.
37. Landin M, York P, Cliff MJ. Scale-up of a pharmaceutical granulation in fixed bowl mixer granulators. Int J Pharm 1996; 133:127–31.
38. Landin M, York P, Cliff MJ, Rowe RC. Scaleup of a pharmaceutical granulation in planetary mixers. Pharm Dev Technol 1999; 4(2):145–50.
39. Faure A, Grimsey IM, Rowe RC, York P, Cliff MJ. A methodology for the optimization of wet granulation in a model planetary mixer. Pharm Dev Tech 1998; 3(3):413–22.
40. Faure A, Grimsey IM, Rowe RC, York P, Cliff MJ. Applicability of a scale-up methodology for wet granulation processes in Collette Gral high shear mixer-granulators. Eur J Pharm Sci 1999; 8(2):85–93.
41. Shaefer T, Holm P, Kristensen HG. Melt pelletization in a high shear mixer: II. Power consumption and granule growth. Acta Pharm Nord 1992; 4.
42. Talu I, Tardos G, van Ommen JT. Use of stress fluctuations to monitor wet granulation of powders. Powder Tech 2001; 117:149–62.
43. Landin M, Rowe RC, York P. Characterization of wet powder masses with a mixer torque rheometer. 3. Nonlinear effects of shaft speed and sample weight. J Pharm Sci 1995; 84/5: 557–60.
44. Faure A, Grimsey IM, Rowe RC, York P, Cliff MJ. Importance of wet mass consistency in the control of wet granulation by mechanical agitation: a demonstration. J Pharm Pharmacol 1998; 50(12):1431–2.
45. Faure A, Grimsey IM, Rowe RC, York P, Cliff MJ. Process control in a high shear mixer-granulator using wet mass consistency: The effect of formulation variables. J Pharm Sci 1999; 88(2):191–5.
46. Leuenberger H. New trends in the production of pharmaceutical granules: the classical batch concept and the problem of scale-up. Eur J Pharm Biopharm 2001; 52:269–77.
47. Betz G, Bürgin PJ, Leuenberger H. Power consumption profile analysis and tensile strength measurements during moist agglomeration. Int J Pharm 2003; 252(1–2):11–25.
48. Betz G, Bürgin PJ, Leuenberger H. Power consumption measurement and temperature recording during granulation. Int J Pharm 2004; 272(1–2):137–49.
49. Leuenberger H. Scale-Up in the Field of Granulation and Drying. Drugs and the Pharmaceutical Sciences: Pharmaceutical Process Scale-Up. Vol. 118. New York: Marcel Dekker, Inc, 2002:151–70.
50. Alexanderwerk AG. Compactors and Granulators page. (Accessed July, 17, 2007 at http://www.alexanderwerk.com/index.php?ILNK=Navi_Prod_Compact_WP200&iL=2).
51. Hussain AS. A collaborative Search for Efficient Methods of Ensuring Unchanged Product Quality and Performance During Scale-Up of Immediate-Release Solid Oral Dosage Forms.

Drugs and the Pharmaceutical Sciences: Pharmaceutical Process Scale-Up. Vol. 118. New York: Marcel Dekker, Inc., 2002:325.
52. Kleinebudde P. Roller compaction/dry granulation: pharmaceutical applications. Eur J Pharm Biopharm 2004; 58:317–26.
53. Fell JT, Newton JM. Effect of particle size and speed of compaction on density changes in tablets of crystalline and spray-dried lactose. J Pharm Sci 1971; 60:1866–9.
54. Rees JE, Rue PJ. Time-dependent deformation of some direct compression excipients. J Pharm Pharmacol 1978; 30:601–7.
55. Roberts RJ, Rowe RC. The effect of punch velocity on the compaction of a variety of materials. J Pharm Pharmacol 1985; 37:377–84.
56. Roberts RJ, Rowe RC. The effect of relationship between punch velocity and particle size on the compaction behaviour of materials with varying deformation mechanisms. J Pharm Pharmacol 1986; 38:567–71.
57. Newton JM, Ingham S, Onabajo OO, The effect of strain rate on mechanical strength of tablets. Acta Pharm Technol 1986; 32:61–2.
58. Armstrong NA. Time-Dependent Factors Involved in Powder Compression and Tablet Manufacture. Int J Pharm 1989; 49:1–13.
59. Holman LE, Leuenberger H. Effect of compression speed on the relationship between normalized solid fraction and mechanical properties of compacts. Int J Pharm 1989; 57:R1–R5.
60. Armstrong NA. Considerations of compression speed in tablet manufacture. Pharm Tech 1990; 9:106–16.
61. Garr JSM, Rubinstein MH. The effect of rate of force application on the properties of microcrystalline cellulose and dibasic calcium phosphate mixtures. Int J Pharm 1991; 73:75–80.
62. Nokhodchi A, et al. The effects of compression rate and force on the compaction properties of different viscosity grades of hydroxypropylmethylcellulose 2208. Int J Pharm 1996; 129:21–31.
63. Monedero M, Jimenez-Castellanos MR, Velasco MV, Munoz-Ruiz A. Effect of compression speed and pressure on the physical characteristics of maltodextrin tablets. Drug Dev Ind Pharm 1998; 24(7):613–21.
64. Garr JSM, Rubinstein MH. An investigation into the capping of paracetamol at increasing speeds of compression. Int J Pharm 1991; 72:117–22.
65. Mann SC, Roberts RJ, Rowe RC, Hunter BM, Rees JE. 1983. The effect of high speed compression at sub-atmospheric pressure on the capping tendency of pharmaceutical tablets. J Pharm Pharmacol 1983; 35:44.
66. Bateman SD, Rubinstein MH, Thacker HS. Pre- and main compression in tableting. Pharm Tech Int 1990; 2:30–6.
67. Ritter A, Sucker HB. Studies of variables that affect tablet capping. Pharm Tech 1980; March:57–62.
68. Bateman SD, Rubinstein MH, Wright P. The effect of compression speed on the properties of ibuprofen tablets. J Pharm Pharmacol 1987; 39:66.
69. Mann SC. An investigation of the effect of individual segments of tableting cycle on the capping and lamination of pharmaceutical tablets. Acta Pharm Seuc 1987; 24:54–5.
70. Zlokarnik M. 2006, Scale-Up in Chemical Engineering. New York: Wiley-VCH Verlag GmbH, 2006:174–9.
71. Tsygan L, Murphy S, Levin M. New Dimensionless Performance Factors of Rotary Tablet Presses for Scale-Up of Time-Dependent Formulations. Baltimore: Poster, AAPS General Meeting, 2004.
72. Ruegger CD. An investigation of the effect of compaction profiles on the tableting properties of pharmaceutical materials. Ph. D. Thesis, 1996.
73. Ho AYK, Jones TM. Punch travel beyond peak force during tablet compression. J Pharm Pharmacol 1988; 40:75.
74. Ho AYK, Jones TM. Rise time: a new index of tablet compression. J Pharm Pharmacol 1988; 40:74.

75. Mehrotra A, Llusa M, Faqih A, Levin M, Muzzio FJ. Influence of shear intensity and total shear on properties of blends and tablets of lactose and cellulose lubricated with magnesium stearate. Int J Pharm 2007; 336:284–91.
76. Marikh K, Berthiaux H, Mizonov V, Barantseva E. Experimental study of the stirring conditions taking place in a pilot plant continuous mixer of particulate solids. Powder Tech 2005; 157:138–43.
77. Muerza S, Berthiaux H, Massol-Chaudeur S, Thomas G. A dynamic study of static mixing using on-line image analysis. Powder Technol 2002; 128:195–204.
78. Williams J, Rahman M. Prediction of the performance of continuous mixers for particulate solids using residence time distributions, Part II: Experimental. Powder Tech 1972, 5,5;307–16.
79. Pernenkil L, Cooney C. A review on the continuous blending of powders. Chem Eng Sci 2006; 61:720–42.
80. Sudah O, Chester AW, Kowalski JA, Beeckman JW, Muzzio FJ. Quantitative characterization of mixing processes in rotary calciners. Powder Technol 2002; 126:166–73.
81. Weinkötter R, Reh L. Characterization of particulate mixtures by in-line measurments, Particle Particle Syst Charact 1994; 11(4):284–90.
82. Kehlenbeck V, Sommer K. Possibilities to improve the short-term dosing constancy of volumetric feeders. Powder Tech 2003; 138:51–6.
83. Williams JC. Segregation of particulate materials–a review. Powder Tech 1976; 15(2): 245–51.
84. Weinekötter R, Reh L. Continuous mixing of fine particles. Particle and Particle Syst Charact 1995; 12(1):46–53.
85. Harwood C, Walanski K, Luebcke E, Swanstrom C. The performance of continuous mixers for dry powders. Powder Tech 1975; 11:289–96.
86. Laurent BFC, Bridgwater J. Convection and segregation in a horizontal mixer. Powder Tech 2002; 123:9–18.
87. Portillo PM, Ierapetritou M, Muzzio FJ. Characterization of Continuous Convective Powder Mixing Processes, Powder Technology 2007; to appear.
88. Laurent BFC, Bridgwater J. Performance of single and size-bladed powder mixers. Chem Eng Sci 2002; 57:1695–709.

5
Dissolution and Drug Release Testing

Vivian A. Gray
V. A. Gray Consulting, Inc., Hockessin, Delaware, U.S.A.

INTRODUCTION

Dissolution Testing is a critical part of the characterization of the drug product. The test involves an elaborate sample preparation step, where the product dissolves under controlled conditions using prescribed equipment. This chapter will describe the equipment, sources of error when performing the test, how to validate the method and qualify the equipment, and lastly how to develop methods from simple dosage forms to the more novel dosage forms of today.

HISTORY OF DISSOLUTION TESTING

In the late 1800s, pill absorption was related to dissolution, and the earliest experiments with in vitro–in vivo correlations occurred in the 1930s. In the 1950s, disintegration testing became official in USP XV. The Kefauver–Harris drug amendments were passed in 1962 to ensure drug effectiveness as well as safety. A USP-NF Panel was created to examine physiologic availability and evaluate mechanisms to help assure drug effectiveness. The Panel recommended the need for dissolution testing and the rotating basket apparatus was chosen based on salicylic acid tablet performance. During the 1970s, there were 12 official monographs in USP using baskets. In the early 1980s, the USP proposed a single-point method, 75% in 45 minutes with water as medium. This specification was, in retrospect, mainly for the BCS Class I (highly soluble/highly permeable) compounds (1). In the 1990s, testing using profiles came into the mix with FDA requiring profiles in all the dissolution and drug release guidances. The FDA also pushed for specifications that were tighter than the 75% in 45 minutes, and instead required 80% in 30 minutes. This was to assure there was manufacturing control. Today dissolution issues center around the poorly soluble drugs (BCS Class II—poorly soluble/highly permeable), since this type of product has become the norm. The call is for more clinically relevant specifications, and in particular, in vitro and in vivo correlations when appropriate. There are many novel dosages forms now seeking regulatory approval, these products require unique methods and apparatus. The concept of quality by design (QbD) is presently affecting the way analysts view the dissolution test. Does it add value?

THEORY

There are three stages in the dissolution process. The first is the disintegration of a gross tablet to particles of various sizes. This can be measured by the Disintegration Test in USP General Chapter <701> (2). This stage also includes the rupturing of the capsule shell. Then there is the deaggregation step, where there is a breakdown of the dosage form into discrete particles that increases the surface area, providing solid-liquid interface and beginning dissolution. The dissolution process continues, and the rate is measured by the dissolution test.

The dissolution rate is represented mathematically by the Modified Noyes and Whitney Equation (3).

$$\text{Rate} = kDS/vh\,(C_s - C_t)$$

where D is the diffusion rate constant, S is surface area, v is volume of the dissolution media, h is thickness of the saturated layer, C_s is concentration of the API at saturation, k is the dissolution rate constant, and C_t is the concentration of the bulk solution. Special attention should be paid to the thickness of the saturated layer as this is where the influence of paddle or basket speed on the dosage unit boundary layer is evidenced. If sink conditions are met, the concentration of the bulk solution should be the concentration of the drug at saturation, diluted by at least a factor of three. It is clear from the equation that the drug substance surface area and hence particle size are very important factors in the dissolution rate. The typical dissolution test measures the rate at which a drug substance dissolves from the dosage unit. The term "in vitro release" is more appropriate in the case of an extended-release (ER) product, since drug is released from a matrix then dissolved in the media. The dissolution rate may be defined as the amount of active ingredient in a solid dosage form dissolved in unit time under standardized conditions or liquid-solid interface, temperature, and media composition. The dissolution results are typically expressed as a cumulative percent dissolved, Q, of the label claim, over time intervals, until at least 80% dissolution is obtained.

When approaching the dissolution of drug product, there are three aspects to consider: the solubility of API, which is typically an equilibrium process; the dynamic process of the dissolution rate; and lastly, but of major influence, the effect of excipients, and the manufacturing process. The later may enhance or impede the dissolution.

REGULATORY AND COMPENDIAL ROLE IN DISSOLUTION TESTING

The Food and Drug Administration

A discussion of dissolution testing begins with the primary regulatory agency in the United States, the Food and Drug Administration (FDA). The role of the FDA regarding dissolution extends beyond the obvious role of approving drug products, thus approving dissolution and drug release tests. The FDA by law is the enforcer of the USP standards put forth in the Compendia. FDA has published many guidances related to dissolution. They have led the scientific debate and issues by cosponsoring workshops with the American Association of Pharmaceutical Scientists (AAPS), USP, and other organizations. The formation of task force groups to address current issues has been a very powerful tool in drafting science-based regulations. For example, the task force on gelatin-coated product cross-linking (4) was able to propose addition of enzyme to dissolution medium.

The FDA labs perform off-the-shelf testing and validation of NDA methods. The compliance officers perform inspections; a major concern for the pharmaceutical industry is the FDA issuance of recalls, many of which are based on dissolution results. Also along these lines, the FDA issues 483 warning letters, some of which are concerned with dissolution issues.

The FDA Guidances

The main FDA guidances related to dissolution and drug release are listed below:

1. Dissolution Testing of Immediate Release Solid Oral Dosage Forms.
2. Extended release oral dosage forms: Development, evaluation, and application of in vitro/in vivo correlations.
3. SUPAC-IR: Immediate-release solid oral dosage forms: scale-up and post-approval changes: chemistry, manufacturing, and controls, in vitro dissolution testing, and in vivo bioequivalence documentation.
4. SUPAC-MR: Modified-release solid oral dosage forms: scale-up and post-approval changes: chemistry, manufacturing, and controls; in vitro dissolution testing and in vivo bioequivalence documentation.
5. SUPAC-SS: Nonsterile semisolid dosage forms: scale-up and post-approval changes: chemistry, manufacturing, and controls, in vitro release testing and in vivo bioequivalence documentation.
6. Waiver of in vivo bioavailability and bioequivalence studies for immediate-release solid oral dosage forms based on biopharmaceutics classification system.

United States Pharmacopeia

The influence of USP on dissolution testing has been critical; many initiatives for dissolution testing, including equipment prototypes and the acceptance criteria, came from USP as the various committees and staff worked with the pharmaceutical industry as well as equipment manufacturers to promote accurate and reproducible dissolution tests. USP has several General Chapters devoted to the area of dissolution and drug release, but first a discussion of disintegration is needed.

General Chapter Disintegration <701>

Disintegration testing has been in existence since 1950 (USP XV). The test was introduced when it was realized that tablets that were made very hard (so they would not chip) also would not disintegrate in the gastrointestinal tract. In 1997, an important discovery by Hoag (5) showed that many vitamin products containing folic acid were not meeting the standard of dissolving within an hour. The disintegration test was mandatory for oral dosage forms for 40 years, but its elimination and replacement with dissolution testing became a standard-setting issue in 1981 (6). This was because the disintegration test was not believed to correlate with in vivo performance (7). The apparatus is seen in Figure 1. From 1990 to 1995, the disintegration tests in the USP were replaced with dissolution tests and the disks were removed.

Now it appears that the disintegration test is re-emerging as the test of choice for fast-dissolving products that have a disintegration test that can relate results to dissolution rates. This is shown in the ICH document Q6A, Decision Tree # 7 (8). As the debate of added value for the dissolution test continues, it may be that more disintegration tests will be the regulatory test for products where disintegration is the only critical release mechanism.

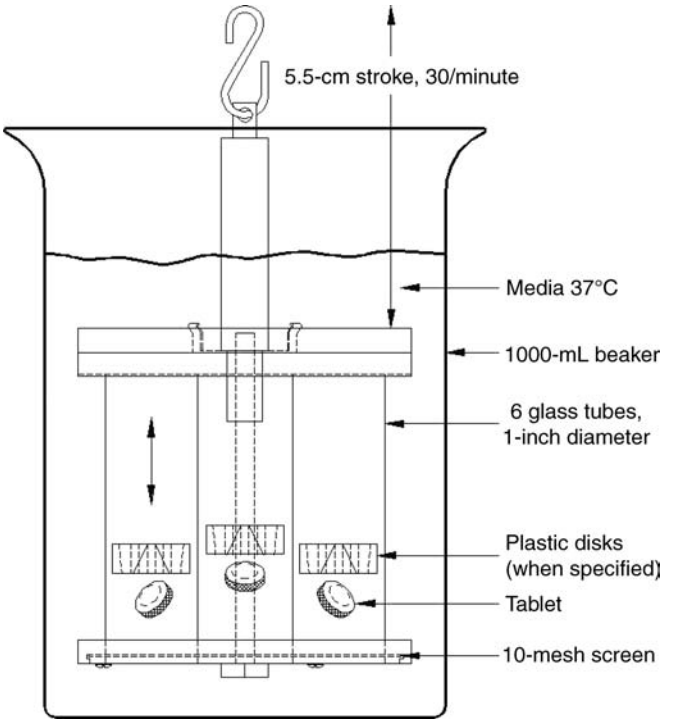


FIGURE 1 USP disintegration apparatus.

The disintegration test is the method now being cited in the Nutritional Supplements section of the USP, with General Chapter <2040> as the recommended procedure.

General Chapter <711> Dissolution

This General Chapter describes the dissolution procedure to be used when testing a monograph product (9). Other than the official test procedure and diagrams of equipment, this chapter contains special notes and instructions on various topics. One of the more recent changes is the allowance of enzyme addition to the second dissolution test when a capsule or gelatin-coated product fails the dissolution test. This addition is an outcome of the FDA gelatin task force mentioned in the section on FDA. The chapter also includes special statements on deaeration/bubbles, calibration, apparatus dimensions, filters, sinkers, and automation. By the early 1990s, the exemptions for chewable tablets and soft gelatin capsules were removed.

In April 2006, the Chapter was officially harmonized with *Japanese Pharmacopoeia* (JP) and *European Pharmacopoeia* (EP). There are now elements of the General Chapter <724> Drug Release within <711>. Those elements are the ER Apparatuses 3 and 4. Apparatuses 5–7 remain in <724>, with that chapter now applied to transdermal dosage form testing.

General Informational Chapters

The content of USP General Chapters above <1000> is considered "informational," somewhat like a guidance. However, if these chapters are referenced in CMC filings, they

take on official status and must be followed. General Informational Chapter <1088> In Vitro and In Vivo Evaluation of Dosage Forms was the precursor to the FDA guidance, Extended Release Oral Dosage Forms: Development, Evaluation, and Application of In Vitro/In Vivo Correlations. Within this chapter, there is immediate/extended release in vitro evaluation or method development instructions. The chapter's main focus is the in vivo evaluation of modified dosage forms and how to perform in vivo-in vitro correlations.

The General informational Chapter <1090> In Vivo Bioequivalence Guidances mainly tells how to conduct bioequivalence tests and contains bioavailability protocols for certain products. This chapter merely repeats what is available from FDA and may be revised to serve some other purpose, probably that of interchangeability.

A very important chapter for all testing procedures is the General Informational Chapter <1225> Validation of Compendial Methods. This chapter is not very informative for dissolution testing methods, and only targets a typical analytical finish to the test, that being chromatographic analysis, mainly by HPLC.

The New General Informational Chapter <1092> the Dissolution Procedure: Development and Validation

This chapter was official in August 2006 (10). This chapter is of utmost importance for dissolution testing and will be explored in greater depth in later sections. The chapter originated with an article written for the Pharmacopeial Forum (11) introducing the concept of a general dissolution chapter that gave guidance on method development and validation of those methods. It was based on industry practices on these topics. The original authors were Vivian Gray, Lew Leeson, Cindy Brown, and Jennifer Dressman; as it progressed to a proposal for USP, the feedback from the USP Expert Biopharmaceutics Committee and comments from PhRMA and other entities were incorporated. The chapter also encourages new technology and automation by instructing on how to validate these analytical methods.

USP Expert Committees and Panels

The standards related to dissolution and drug release issues are addressed by the USP Biopharmaceutics Expert Committee, which is elected every five years according to the revision cycle. The committee members for 2005–2010 are Thomas Foster (Chair), Clarence Ueda, Vivian Gray, Lew Leeson, Eli Shefter, Diane Burgess, Nhan Tran, Leon Shargel, Bryan Crist, Alan Parr, Johannes Kraemer, William Simon, James Polli, and Mario Gonzalez. There are also various Advisory Panels that are selected to address pertinent issues. In 2007, several Advisory panels are working on topics of performance verification testing (previously referred to as calibration) and performance testing for all forms of dosage form delivery.

Other Dissolution Regulatory Documents

The International Federation of Pharmaceutical Scientists issued Guidelines for Dissolution Testing of Solid Oral Products in 1996 (12), and there are regulatory documents from both Europe (13) and Japan (14) that address dissolution topics. There are also Dissolution General Chapters in the WHO International Pharmacopoeia, EP, and JP.

The International Conference on Harmonization (ICH) mandated that the USP, EP, and the JP harmonize the general chapters on dissolution, disintegration, and drug release.

The ICH document "Q6A Decision Trees #7: Setting Acceptance Criteria for Drug Products Dissolution" contains three decision trees. The first discusses the types of drug release acceptance criteria that are appropriate and mentions disintegration testing in lieu of dissolution testing. The second decision tree points to specific test conditions and acceptance criteria that are appropriate for immediate release; the topic of a dissolution test with or without discriminatory power is specifically addressed. The third decision tree deals with appropriate specifications for extended release. The subject of in vitro-in vivo correlations and relationships is covered.

COMPENDIAL EQUIPMENT REVIEW AND SOURCES OF ERROR

The most important aspect of the dissolution equipment is that it provides undisturbed homogenous mixing leading to complete or near complete dissolution and also is designed so that the visual observations are easily obtained. Each aspect of equipment can be a source of error. The major components of the equipment are shown in Figure 2. There is the dissolution tester "head" containing the drive belt, spindle assemblies, and electronics for the mechanical aspects of the equipment. Then there is a water bath that includes a circulator and inlet screen where the vessels are placed, and a top plate containing insert holes for the vessels. Sometimes the vessels are "jacketed" and heated through heating elements instead of water (15). The stirring mechanisms are shafts inserted in the spindle assemblies. These shafts are one entity with either a paddle stirring device (Fig. 3) or a basket attached (Fig. 4). The vessels are inserted into the water bath and filled with dissolution medium. The paddle apparatus is referred to as USP Apparatus 2 and the basket apparatus as USP Apparatus 1. Most commonly they are simply referred to as the "basket" and "paddle."

As a regulatory test, dissolution must be accurate and practical. Justification would be provided for atypical conditions. The test should have low variability and a good profile. Test results should show changes in the formulation and, ideally, an in vivo-in vitro relationship should exist.

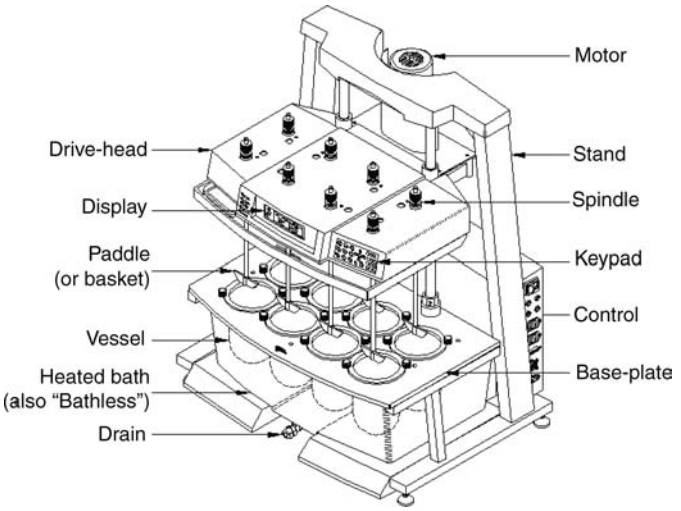


FIGURE 2 Example of modern dissolution test equipment.

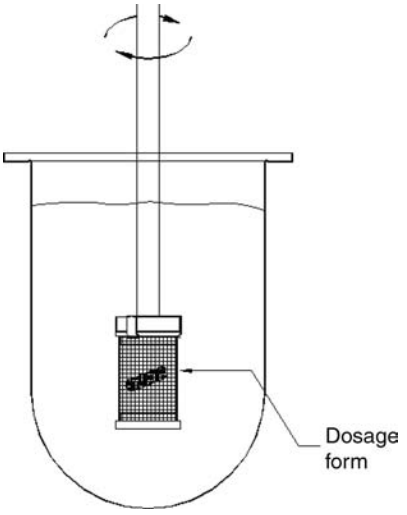


FIGURE 3 USP Apparatus 1: Basket.

The essentials of the test are accuracy of results and robustness of the method. Aberrant and unexpected results do occur, however, and the analyst should be well-trained to examine all aspects of the dissolution test and watch the equipment in operation.

When performing dissolution testing, there are many ways that the test may generate erroneous results (16). The testing equipment and its environment, sample handling, formulation, in-situ reactions, automation, and analytical techniques may be the cause of errors and variability. The physical dissolution of the dosage form should be unencumbered at all times. Certain aspects of the equipment calibration process, as well as a close visual observation of the test, may reveal these errors.

Knowledge of drug properties, especially solubility in surfactants or as a function of pH, is essential. One could anticipate precipitation of the drug as the solution pH changes or as the amount of drug increases. Be aware that complete dissolution of the drug in the standard solution may be more difficult than expected. It is customary to use a

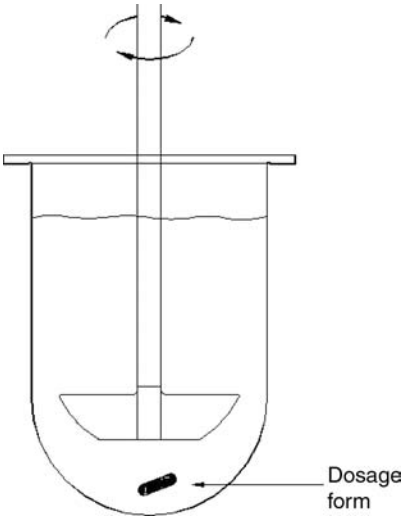


FIGURE 4 USP Apparatus 2: Paddle.

small amount of alcohol to dissolve the standard completely. A history of the typical absorptivity range of the standard can be very useful to determine if the standard has been prepared properly.

Highly variable results indicate that the method is not robust, and this can cause difficulty in identifying trends and the effects of formulation changes. Two major causal factors influence variability, mechanical and formulation. Mechanical causes can arise from the dissolution conditions chosen. Carefully observe the product as it dissolves. An apparatus or speed change may be necessary.

The formulation can have poor content uniformity, and reactions or degradation may be occurring in situ. The film coating may cause sticking to the vessel walls. Upon aging, capsule shells are known to form pellicles, and tablets may become harder or softer, affecting the dissolution and disintegration rate depending upon the excipients and drug interaction with moisture.

Equipment Variables

The major components of dissolution equipment are the tester, water bath, paddles, baskets and shafts, vessels, samplers, and analyzers.

Mechanical aspects, such as media temperature, paddle or basket speed, shaft centering and wobble, and vibration can all have a significant impact on the dissolution of the product. Mechanical and chemical calibration should be conducted periodically, usually every 6 months, to ensure that the equipment is working properly.

The USP General Chapter on Dissolution <711> contains a requirement for the analyst to perform the Apparatus Suitability Test using USP Calibrator Tablets. USP Calibrator Tablets come with certificates identifying appropriate ranges. The Apparatus Suitability Test is designed to detect sources of error associated with improper operation and inadequate condition of the equipment (17–19). Two calibrators are used, USP Prednisone tablets, 10 mg, and USP salicylic acid tablets, 300 mg. Use of each of these types of Calibrator Tablets involves unique considerations.

The salicylic acid tablets should be brushed before use to remove fine particles. This should be done in the hood to avoid breathing the irritating dust. Whole tablets are used, but the tablets can be chipped or nicked. Since this tablet dissolves through erosion and is pure compressed salicylic acid, minor chips or nicks have no significant effect on the dissolution rate. The buffer should be prepared according to USP Reagent (Buffers) section.

Deaeration

The Prednisone tablets use deaerated water as the medium. There are numerous methods for deaeration of medium (20–23). Automated methods are also available. The method described in *USP 29* uses heat, filtration, and vacuum. Helium sparging is also a typical method for deaeration. The level of dissolved oxygen and other gases is related to the presence of bubbles. Bubbles are common and will cause problems in non-deaerated medium. USP General Chapter on Dissolution <711> states that bubbles can interfere with dissolution test results and should be avoided. Dissolved air can slow down dissolution by creating a barrier; bubbles may adhere to either the tablet surface or to basket screens or particles can cling to bubbles on the glass surface of the vessel or shafts. The test should be performed immediately after deaeration. It is best not to have the paddle rotating before adding the tablet, since paddle movement aerates the medium. When preparing standard solutions, the reference standard must be dried properly, preferably on

the day of use. Care should be taken to ensure that the drug powder is completely dissolved. In the case of Prednisone Reference Standard, the powder becomes very hard upon drying, making it slower to dissolve. Dissolving the powder first in a small amount of alcohol helps to eliminate this problem.

Vibration

Vibration interference is a common problem with dissolution equipment (23–25). Careful leveling of the top plate and lids is critical. Within the spindle assembly, the bearings can become worn and cause vibration and wobble of the shaft. The drive belts should be checked for wear and dirt. The tension adjustments for the belt should be optimized for smooth operation. Surging of spindles, though difficult to detect without closely scrutinizing the tester operation, can cause spurious results. Vessels need to be locked in place so they are not moving with the flow of water in the bath.

External vibration sources might include other equipment on bench tops such as shakers, centrifuges, or sonicators. Local construction in the area or within the building is a common, though often overlooked, source of vibration. The testers should not be near hoods or significant air-flow sources. Heavy foot traffic and door slamming should be avoided.

Water Bath

These days, the water bath itself is rarely a source of vibration because the design has been changed to eliminate noisy circulators near the bath. Measuring the temperature of the medium in all the vessels, rather than just one, can assure the temperature uniformity. The bath water level should always be maintained at the top of the vessels to ensure uniform heating of the medium. Last, the water bath should contain clean water so observations of the dissolution test can be performed clearly and easily.

USP Apparatuses 1 and 2

The basket and paddle can be sources of error if not closely inspected before using. Obviously, dimensions should be as specified. In cases of both baskets and paddles, shafts must be straight and true. The paddles are sometimes partially coated with Teflon. This coating can peel and partially shed from the paddle, causing flow disturbance of hydrodynamics within the vessel. Paddles can rust and become nicked or dented; this can adversely affect dissolution hydrodynamics and be a source of contamination. Thorough cleaning of the paddles is important to preclude carry over of drug or medium.

The baskets need special care and examination. They can become frayed, misshapen, or warped with use. Screen mesh size may change over time, especially when used with acidic medium. There are different designs for attaching baskets to shafts. The attachment can be with clips or with O-rings. These attachment variations can affect dissolution results, depending upon the product; therefore, this factor should be taken into consideration when evaluating the method for ruggedness (24,26). Baskets are especially prone to gelatin or excipient buildup if not cleaned immediately after use. Off-center shafts are often critical factors in failed calibration, especially with the USP Prednisone Calibrator tablets.

Glass Vessels

Vessels have their own set of often-overlooked problems. The method of manufacturing of the glass is proprietary. Vessels are probably manufactured from large glass tubing. The

vessel bottom is probably hand blown and molded. Depending upon techniques of the molding process, irregular surfaces can occur, and the uniformity of vessel bottom roundness can vary. Cheaply made vessels are notorious for this problem. There have been extensive studies on the effects of the vessel shape on dissolution results (19,27–29). Close examination of newly purchased vessels is very important, since surface irregularity can cause dissolution results to differ significantly. Another common problem with vessels is residue buildup, either from oily products or sticky excipients. Insoluble product that is not rinsed well from previous testing can cause contamination. Vessels that become scratched and etched after repeated washing and should be discarded. Lids need to be in place to prevent evaporation. As mentioned before, vessels should be locked down to avoid vibration.

Calibration Failures

In assessing calibration failure, one should examine the system by changing one parameter at a time. Do not retest until passing results are obtained. Retest one position only if it is associated with a unique problem, but repeat the entire calibration if adjustments are made to the tester. Good manufacturing practices (GMP) dictate that all adjustments should be documented and all maintenance recorded.

USP Apparatus 3

The Reciprocating Cylinder (Fig. 5) is used mainly as a research tool where the need to change pH is prominent. As seen by the design, the dosage unit can be moved from row to row, and in each row the vessels may contain media of different pH or components. The equipment has a special use for beaded products; the beads are contained by the screens in the upper and lower parts of the cell, yet the reciprocating motion allows good mixing (30–32).

Sources of error when using this apparatus are mainly associated with the loss of media through evaporation and the achievement of sink conditions when the drug is poorly soluble. This lack of sink conditions may be overcome when the product goes from row to row. The elements that need careful study are that the screen mesh size is appropriate for the product, that products do not adhere to the screen, and that the dip rate is constant. When using surfactant, there can be considerable foaming.

USP Apparatus 4

This unique equipment is also known as the flow-through cell (Fig. 6). The drug product is positioned in a cell where the dissolution medium is constantly dissolving and flowing over the tablet. The liquid passes through a filter at the top of the cell and is then collected in a reservoir. Because of this constant flow of media, an ER product or a poorly soluble product can continually be in a sink environment.

Sources of error when using this apparatus are centered on the pump and flow rate reliability and the clogging of the filters. Other considerations related to the flow of liquid through the cell would be the position of the tablet holder the quantity of glass beads used, and tubing lengths, material, and diameters. A special edition of Dissolution Technologies, May 2005, was devoted to methods using Apparatus 4.

USP Apparatus 5

This apparatus is commonly known as the Paddle over Disk and is devoted specifically to the transdermal patches. As shown in Figure 7, there are two patch-holding designs, the watch glass assembly and the screen disk. The screen disk appears to be the official USP

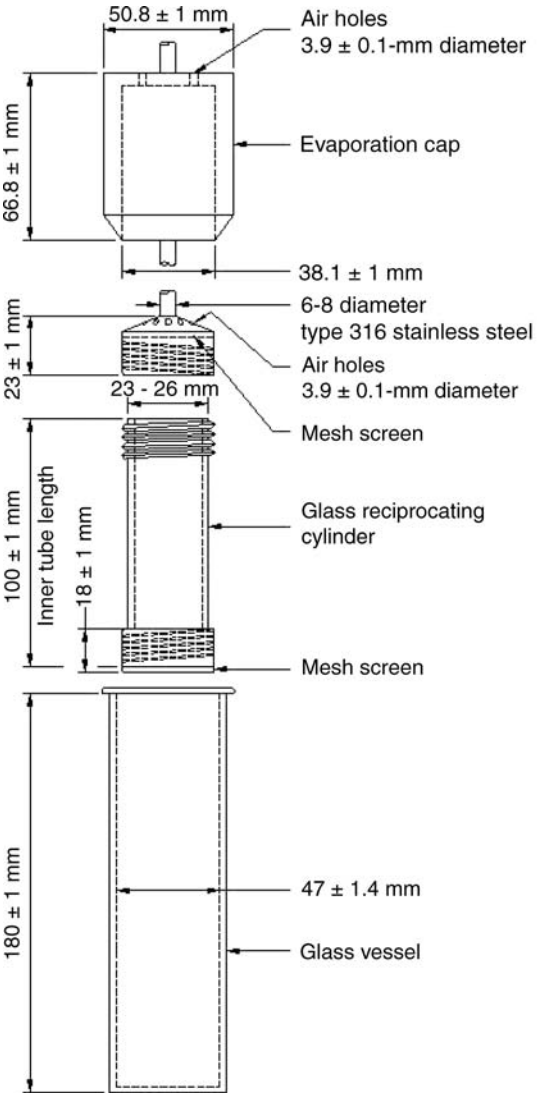


FIGURE 5 USP Apparatus 3: Reciprocating cylinder.

apparatus, but if one reads the general chapter closely, the water glass assembly is also an option. FDA has published articles claiming that the water glass is the only apparatus needed for the transdermal patch.

Sources of error for this apparatus would be similar to those mentioned earlier with Apparatus 2, and the positioning and attachment of the patch to the device chosen are critical.

USP Apparatus 6

As with Apparatus 5, this apparatus is exclusively used for transdermal patches. As shown in Figure 8, the patch is adhered to the cylinder in such a way that the "active" side of the patch is facing the medium.

Sources of error for this equipment would also be centered on the same attributes as for Apparatus 2. The straightness of the shaft would be of the most importance along with the proper and firm adherence of the patch to the surface.

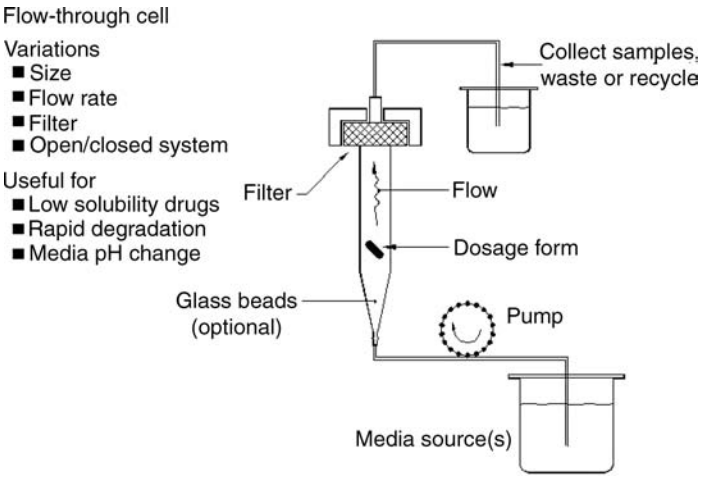


FIGURE 6 USP Apparatus 4: Flow through cell.

USP Apparatus 7

Apparatus 7 is commonly known as the Reciprocating Holder. This apparatus has five designs (Fig. 9). It operates in a reciprocating motion as in Apparatus 3 and also goes from one beaker/vessel to another. There are three designs for use with transdermal patches; the other two designs are for specially designed tablets, called an osmotic pump. These tablets usually have a laser hole where there is a push/pull effect of drug from a polymeric matrix. The hole must be exposed to the medium in a uniform manner; hence the design is a rod-like shaft where the dosage form is glued to the tip of the rod. Another variation is a spring-like cage at the end of the rod that houses the dosage unit.

Sources of error are similar to Apparatus 3 where reciprocation is the agitation principle. The accuracy of the indexer is also a critical parameter.

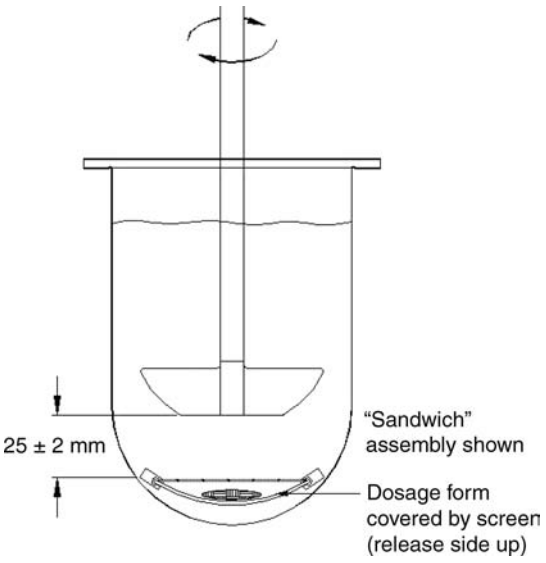


FIGURE 7 USP Apparatus 5: Paddle over disk with "sandwich" or "watch glass" assembly shown.

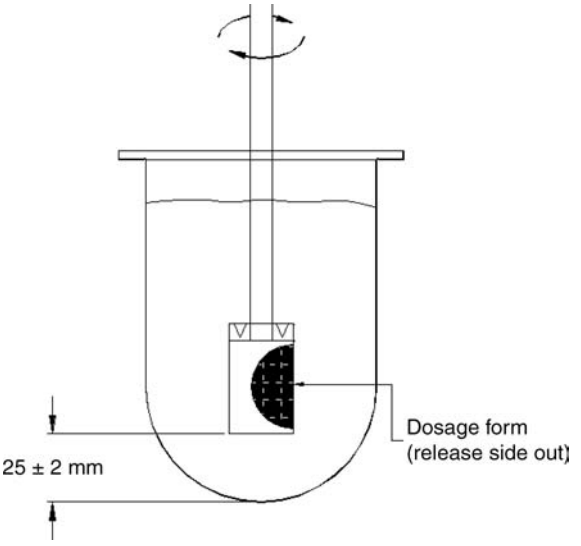


FIGURE 8 USP Apparatus 6: Rotating cylinder.

Method Considerations

The best way to avoid errors and data "surprises" is to put a great deal of effort into selecting and validating methods. Some areas of testing are especially troublesome. Sample introduction can be tricky and, unfortunately at times, uncontrollable. Products can have a dissolution rate that is "position dependent." For example, if the tablet is off-center, the dissolution rate may be higher due to shear forces. Or if it is in the center, coning may occur and the dissolution rate will go down. Film-coated tablets can be sticky and pose problems related to tablet position. Little can be done except to use a basket (provided there is no gelatinous or excipient build up) or a sinker.

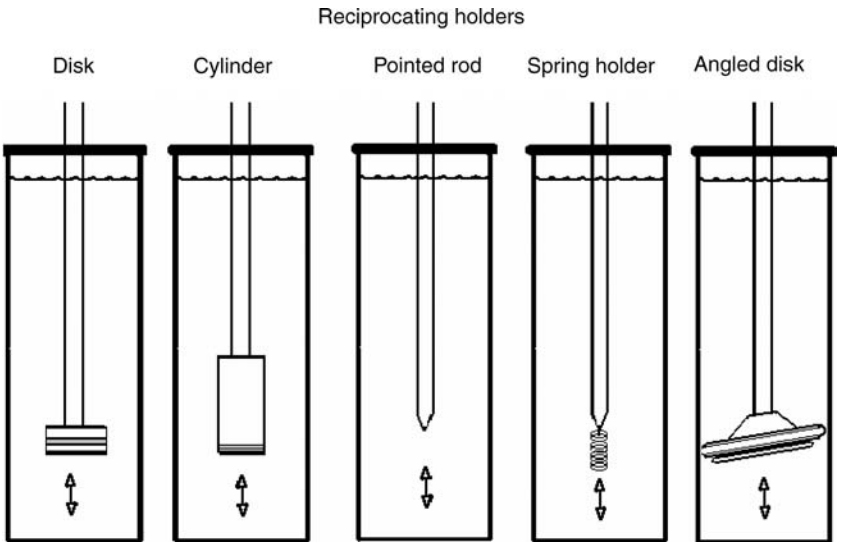


FIGURE 9 USP Apparatus 7: Five designs.

Suspensions can be introduced in a variety of ways: manual delivery using syringes or pipettes, pouring from a tared beaker, or automated delivery using calibrated pipettes. Each method has its own set of limitations, although automated methods may show less variability. Mixing of the suspension sample will generate air bubbles; therefore, the mixing time of suspension samples must be strictly uniform to reduce erroneous or biased results.

Media Attributes

The medium is a critical component of the test that can cause problems. One cause of inaccurate results may be that too great a volume of medium has been removed through multiple sampling without replacement, thereby adversely influencing sink conditions.

Surfactants can present quite a cleaning problem, especially if the concentration is high (i.e., over 0.5%). In the sampling lines, surfactants such as sodium lauryl sulfate (SLS) may require many rinsings to assure total elimination. The same is true for carboys and other large containers. This surfactant has other limitations, for quality can vary depending upon grade and age, and the dissolving effect can consequently change depending upon the surface-active impurities and electrolytes (33). The foaming nature of surfactants can make effective deaeration very difficult. Some pumps used in automated equipment simply are not adapted to successful use with surfactants. One caution when lowering a basket into a surfactant medium is that surface bubbles can adhere to the bottom of the basket and decrease the dissolution rate substantially. When performing HPLC analysis using surfactants as the medium, several sources of error may be encountered. The auto-injectors may need repeated needle washing to be adequately cleansed. Surfactants, especially cetrimide, may be too viscous for accurate delivery. Surfactants can affect column packing to a great degree, giving extraneous peaks or poor chromatography. Basic medium, above pH 8, may cause column degradation

Observations

One of the most useful tools for identifying sources of error is close observation of the test. A trained analyst can pinpoint many problems because he or she understands the cause and effect of certain observations. Accurate, meaningful dissolution occurs when the product dissolves without disturbance from barriers to dissolution, or disturbance of vessel hydrodynamics from any source. The particle disintegration pattern must show freely dispersed particles. Anomalous dissolution usually involves some of the following observations: floating chunks of tablet, spinning, coning, mounding, gumming, swelling, capping, "clam-shell" erosion, off-center position, sticking, particles adhering to apparatus or vessel walls, sacs, swollen/rubbery mass, or clear pellicles. Along with good documentation, familiarity with the dissolution behavior of a product is essential in quickly identifying changes in stability or changes associated with a modification of the formulation. One may notice a change in the size of the dissolving particles, excipients floating upward, or a slower erosion pattern. Changes in the formulation or an increase in strength may produce previously unobserved basket screen clogging. If the contents of the basket immediately fall out and settle to the bottom of the vessel, a spindle assembly surge might be indicated. If the medium has not been properly deaerated, the analyst may see particles clinging to vessel walls. The presence of bubbles almost always indicates that deaeration is necessary.

Sinkers

Sinkers are defined in USP as "not more than a few turns of a wire helix…." Other sinkers may be used, but the analyst should be aware of the effect that different types of

sinkers may have on mixing (34). Sinkers can be barriers to dissolution when the wire is wound too tightly around the dosage unit.

Filters

Filters are used on almost all analyses; many types or different materials are used in automated and manual sampling. Validation of the pre-wetting or discard volume is critical for both the sample and standard solutions. Plugging of filters is a common problem, especially with automated devices.

Manual Sampling

Manual sampling techniques can introduce error by virtue of variations in strength and size of the human hand from analyst to analyst. Therefore, the pulling velocity through the filter may vary considerably. Too rapid a movement of liquid through the filter can compromise the filtration process itself.

Automation

While automation of dissolution sampling is very convenient and labor saving, errors often occur with these devices because the analysts tend to overlook problem areas. Sample lines are often a source of error for a variety of reasons: unequal lengths, crimping, wear beyond limits, disconnection, carryover, mix-ups or crossing, and inadequate cleaning.

The volume dispensed, purged, recycled, or discarded should be routinely checked. Pumping tubes can wear out through normal use or repeated organic solvent rinsings and may necessitate replacement.

The use of flow cells may generate variability in absorbance readings. Air bubbles can become caught in the cell, either introduced via a water source containing bubbles or by inadvertently entering into poorly secured sample lines. Flow rate and dwell time should be evaluated so the absorbance reading can be determined to have reached a steady plateau. Cells need to be cleaned frequently to avoid buildup of drug, excipient, surfactant, or buffer salts from the dissolution medium.

Cleaning

Cleaning of equipment needs to be stressed as it is an overlooked source of error and contamination. The analyst should take special care to examine this aspect when validating the method. In many laboratories where different products are tested on the same equipment, this is a critical issue that, if inadequately monitored, may be a cause of inspection failures.

CALIBRATION OF COMPENDIAL AND NONCOMPENDIAL EQUIPMENT

Calibration of Apparatuses 1 and 2

As mentioned above, the calibrator tablets for Apparatuses 1 and 2 are used routinely. Historically, the calibrator tablets were first needed because representatives from the FDA, USP, and then PMA (now Pharma) all agreed that vibration (internal and external) was influencing the dissolution results of products (35). The USP was charged with the responsibility of adopting calibrator tablets. In the late 1970s, the calibrator tablets were

put in place and were required in <711> USP General Chapter on Dissolution. Now in 2008, we have not been able to assess vibration in any other way except calibrator tablets. In a PhRMA study (36) assessing the value of the calibrator tablets, one conclusion was that "… some type of calibrator tablets should be maintained until enhanced mechanical calibration is further defined (e.g., establishing a definitive vibration tolerance)." We have to give credit to many of the equipment manufacturers who have diligently designed testers that have less and less internal vibration. However, even well-designed equipment that is used for years for 1 hour, or 8 hours, or even 24 hours a day will eventually show signs of wear. Also, the external environment can subject the equipment to vibration from heavy foot traffic, nearby construction, and nearby equipment on the same bench top, to name a few sources. We also have to acknowledge that not all equipment on the global market is solidly designed. With no mechanical means to test vibration other than calibrator tablets, removing calibrator tablets from the equipment performance assessment raises great concern. It is well-documented fact that vibration affects the dissolution results (23–25,37–39), and in some cases, the results are biased high giving a false passing result. The consequences of false passing results should be of great regulatory concern.

There is another aspect of the equipment that is only detected at the present time by calibrator tablets, and that is vessel asymmetry. The glass dissolution vessel is not made from a mold but most probably made from a combination of individual hemispheric shapings from standard tubing (27). The irregularities in the vessel shape can cause a change in the fluid flow pattern and hence change the dissolution results. In the early days of dissolution testing, the FDA lab scientists pointed this out in a publication in 1982 (28). Since then, it has been substantiated in other publications and practical lab experience in many reputable laboratories (19,24,29) As of yet, there are no available mechanical means of detecting flaws in the vessel design, although there may be some devices on the horizon. Until then, the calibrator tablets are the only appropriate tool for detecting this problem.

Calibration of Other Official Apparatus

In the past, there were two calibrator tablets for Apparatus 3, Chlorpheniramine Maleate tablets and Theophylline Beads. Now the Chlorpheniramine Maleate tablets are the only calibrator tablets required. Mechanical parameters are stated in the <711> general chapter. The Apparatuses 5 and 6 are partially covered by having the equipment pass the calibration using Apparatus 2—as this shows the tester and vessels are able to generate accurate results.

Apparatuses 4 and 7 do not have calibrators; however, mechanical parameters are shown in General Chapter <711>. This equipment along with modifications can be qualified in the same manner as non-compendial equipment.

Non-Compendial Equipment Calibration

Some examples of non-compendial equipment are the rotating bottle, mini paddle, mega paddle, peak vessel, diffusion cells (Franz and Enhancer), chewing gum apparatus, and some Apparatus 4 cell designs. Standard equipment should be the first choice, and it should always be justified why official equipment is not suitable.

If the equipment is a commercial product, the installation and operational qualifications can be obtained from the equipment vendor (40). This would include the vendor specifications and tolerances for the equipment. For an in-house design, this becomes more difficult. The first objective would be to look for adjustments and moving

parts. Obtain a baseline of operational parameters, such as agitation rate (rpm), dip speed, flow rate, temperature, alignment, and/or volume control. After enough historical data have been obtained, examine the data for reproducibility, assessing the variability of the various components. If the analyst is satisfied that the equipment performs consistently, then chose ranges or limits based on this data. Then develop a per-run performance checklist based on these parameters. To calibrate or more correctly show performance qualification for non-compendial equipment where a calibrator tablet is not available, there could be an in-house calibrator tablet designated. This should be a product that is readily available with a large amount of reproducible historical data generated on the equipment. It must be a well-characterized and stable product, which ensures that all components of the test are considered, this being the analyst, equipment, and method.

Mechanical parameters such as volume control, alignment, temperature, vibration, flow rate (dip rate, agitation rate, RPM), oscillation frequency and distance, and timing of indexer may be sufficient without the development of a PVT. It should be determined if there is some unique aspect of the equipment that can only be detected using a calibrator tablet. Currently, with Apparatuses 1 and 2, vibration and vessel irregularities are detected by the USP calibrator tablets, with no other practical measuring tools available to the analyst.

For any equipment, hydrodynamics is a big concern. The dissolution fluid-flow characteristics should consist of a predictable pattern that is free of irregularities or inconstant turbulence. Observations of the product dissolution behavior are critical when choosing a dissolution apparatus. If aberrant or highly variable data can be attributed to the apparatus, then it may be unsuitable for that product.

When using non-compendial equipment, the transferability to another site or laboratory should be considered. Non-compendial equipment for quality control testing or at a contract laboratory could present problems of ruggedness. This imposes that ruggedness be thoroughly evaluated before considering transferring product testing using another piece of similar equipment located elsewhere. The non-compendial equipment must have documentation or a log book for tracking the repairs, problems, maintenance, and product performance. Regular calibration, mechanical or chemical, should be documented and the time interval determined. A standard operating procedure (SOP) on operation, maintenance, and calibration should be included. Training and training documentation are critical. The cleaning of any equipment is important. Be alert to parts that may be hard to clean and lead to contamination or residue buildup.

GOOD MANUFACTURING PRACTICES IN DISSOLUTION TESTING

In the dissolution laboratory, GMP issues are pervasive, since there is so much equipment, documentation, and validation involved in testing many products in different stages of development (41). Multiple users of equipment, reagents, and solutions, performing testing on the same and different products add complexities to the laboratory operations. Each lab could have 10–40 testers with associated autosamplers; HPLCs including detectors, pumps, autoinjectors, and columns; UV spectrophotometers and autosippers; deaeration equipment; and fully automated testing equipment, all with logbooks and calibration, maintenance, and operation procedures. The test requires extensive notebook documentation and witnessing as the profile test can have numerous data points with observations and pre- and post-equipment checks. The variety of products requires constant validation and re-validation as formulations change and new test methods are written and revised. Constant monitoring of adherence to GMP is necessary to assure

compliance and successful audit results. Internal audits need to be a regular part of the laboratory operations. The training and documentation of training is becoming more critical in the modern lab where turnover can be high and the type of products quite different.

Metrology

Metrology is an important function associated with the dissolution laboratory. The tracking of equipment identification, repairs, and the calibration status may be performed by personnel outside the dissolution group. This involves frequent communication between the groups, especially in the realm of calibration timelines. Calibration of equipment at its due date is a good indication of the efficiency of the laboratory operations. Missed or late calibration dates can accumulate and give the appearance of poor management of resources and priorities, even if the equipment is labeled appropriately. The status of equipment, whether it is out of service for repairs, calibration, or under investigation, should be very clearly and boldly marked as to avoid any ambiguities as to the equipment condition and usability. Special circumstances, such as use for only one apparatus or new equipment waiting for validation, should be labeled accordingly.

Logbooks or any notebooks associated with or assigned to equipment have to be current and contain the most useful information, that is, observations of problems, how the problems were remedied, calibration results and failures, corrective action, and routine maintenance or performance checks. It is assumed that there is a custodian for each piece of equipment and that this person enters the information into the logbooks. This becomes somewhat cumbersome when someone other than the custodian uses the equipment. Communication becomes critical so the analyst knows when the equipment has had problems in the past. The accurate and current logbook can offer insight into the cause of aberrant data and support the repair, replacement, or upgrading of equipment.

The operational procedures need to have enough detail so an analyst can use the instrument to obtain accurate results without having to rely on verbal hints and reminders from the more experienced users.

Notebook Documentation

There will certainly be a current SOP for documentation in notebooks. The dissolution test does lend itself to inserts or templated work sheets, and such practices are very useful for several reasons. The analyst has many things to remember such as the rpm and temperature checks (before and after the run), the correct speed and apparatus, sinkers or no sinkers, deaeration or no deaeration, observations, sample and equipment IDs, and sample and reagent preparation. This is only a partial list of all the items that should be recorded. A templated list where one fills in the blanks or makes a check mark can serve to keep the information in an organized manner, which will aid the witness tremendously. It also causes the analyst to double check that all aspects of the test have been performed properly. The treatment of inserts or templated worksheets has to be clearly spelled out in the SOP, and quality assurance personnel should have complete confidence that the documentation would meet all compliance concerns.

The recording of sampling times is the subject of much discussion. Does the analyst record in real time every pull (using a traceable calibrated timepiece, of course), or does he/she refer to a test method and presume adherence to the prescribed sampling interval? With manual sampling, this can be a labor-intensive task. Fortunately, with autosampling this is alleviated as the instrument printout tells when the sample was taken.

In the dissolution lab where the testing may require multiple users for the same standard solution and/or medium preparation, there may be special notebooks that are used specifically for this purpose. The specific preparation and date are entered into the notebook; as other analysts use the solution or medium, the date and analyst initials are also entered. The analyst refers to the multi-user notebook number and page in his/her notebook as part of the write-up of the experiment. The witness has to refer to this separate notebook when checking the data. The multi-user notebook will probably need an exception to the SOP for the notebook policy, because most notebooks are for a single analyst.

The role of the witness should not be underestimated. The best witness is an analyst who has performed the test previously and can accurately pick up omissions, mistakes, and out-of-trend results. The witness, in addition to having in-depth familiarity with the method, has some training on the witnessing process. A checklist of things to watch for would be useful.

Equipment Qualification and Method Validation

One of the most frequently sighted areas for 483 warning letters is the lack of validation or improper validation. With the frequent use of autosamplers and fully automated systems in the dissolution laboratory, test method validation using manual versus automation is paramount. The equipment also needs to be validated, with a focus on the unique performance aspects of the specialized equipment. There are two parts to this issue. The instrument itself should go through performance checks that are part of the routine operation of the instrument, usually thought of as operation qualification (OQ). Presumably the installation qualification (IQ) was performed previously when the instrument was newly acquired. When the OQ and IQ are satisfactorily completed, then and only then, can validation be performed using the product. Validation of the use of a simple autosampler may be a straightforward manual and automated run performed concurrently, comparing the results with predetermined acceptance criteria based on the inherent variability of the product. A fully automated system is much more complicated and requires a validation report as part of the validation documentation. Any automated system validation should address contamination from previously tested compounds (cleaning validation) and buildup of surfactant. Pump dwell times, sample lines, and filter checks are often problem areas.

Test methods should reflect the discoveries of a thorough validation. A “critical factors” section is a major component of the method. This part will point out certain aspects of the analysis that require special attention. For example, standard preparation may be addressed. In dissolution testing, the standard may be difficult to dissolve in aqueous medium. Instructions as to the proper amount and addition order of a small amount of alcohol may be very critical to the proper dissolution of the drug substance. The following are examples of critical factors: the deaeration method; sinker type and, if hand made, the instructions; standard preparation if alcohol is used, including sonication time; cleaning instructions for vessels and/or autosamplers; special precautions for cleaning autoinjectors when surfactants are used; septum replacement for auto-injector vials; filter type and discard volume; apparatus speed if not the typical speed; special instructions for the rotation of paddles before the test begins (this may be required for suspensions); exact mixing procedures for dosage forms that need reconstitution; typical absorptivity values (UV) or response factors (HPLC); and precautions to protect from light. Of course, this information is in the method, but a highlighted critical factors section will alert the analyst to aspects of the test that are out of the ordinary.

Audits

Frequent internal audits are a means to keep analysts aware of GMP issues. An internal audit by the dissolution lab personnel is a very good way to monitor GMP and serves as a training tool for the analysts doing the monitoring by compelling them to consider their own work habits. Analysts feel less threatened by observations from lab members than from outside personnel. Internal audits can be done routinely as a part of objectives or performance standards. A checklist is an important aid to this process. The auditor should immediately inform the group of his/her findings without mentioning names; e-mail is a good communication tool. The offenders will usually correct the problem areas. One area that should be routinely inspected in the dissolution lab is sources of vibration, especially external vibration. The counter tops should be examined to see if the dissolution bath is in close proximity to shakers, hoods, or centrifuges. Local construction is a source of vibration and can be overlooked. Observe if there is heavy foot traffic and opening and slamming of doors nearby. It would be a good idea to make vibration a part of the audit checklist.

Other internal audits are performed by QA or teams of section analysts. Routine audits are a necessity to ensure that GMPs are followed, since it is common knowledge that keeping up with all the details is tedious and sometimes ignored, especially in a high-paced testing environment.

Training

In the dissolution lab, training can be labor-intensive and drain resources. However, the area of training is scrutinized by regulatory agencies, so it must be performed adequately and documented. Training is a two-part issue. One part is the training of a new analyst to performing dissolution testing properly, and the other is the training on compound-specific test methods. There is some question as to the role of using the calibration of the equipment as a training tool. The bath calibration is a challenging task and certainly will demonstrate the proficiency of the person performing the test. The difficulty is in using the training to perform an actual calibration, since a failure would pose problems. The training could be done in tandem with an actual calibration performed by a well-trained analyst. There are other aspects of training for dissolution testing, for example, observations. In no other analysis are observations so critical. Training in terminology and what to look for during a dissolution test can be extremely useful in explaining aberrant data and exploring the correct method during method development. The training of a new analyst should be assigned to one person who should track when and if all the training elements are complete. The completion of training should be entered into training records that are kept by a system that is regulated by a training SOP.

Training on a particular method can also be viewed two ways. Some believe an analyst can take a method and perform the test without doing a "training test." Others take a more conservative approach and insist that the analyst perform a training sample test, the results of which should agree with those obtained by an experienced analyst. It is probably best to consider the experience level of the second analyst and the difficulty or uniqueness of the test. A training test may not be needed for a project where the test is routine; however, training test may be appropriate for a test that requires detailed observations or complicated sample introduction (e.g., suspensions).

METHOD VALIDATION

The level of validation depends on the phase of product development. For scouting, the linear range of standards may be sufficient, but as the need for "reportable" data

approaches, the validation parameters increase. This discussion of validation will cover "full validation" of a product that is very far along in the development process, at the end of Phases 2 or in 3. The new USP General Informational Chapter <1092> The Dissolution Procedure: Development and Validation (10) should be used as the pre-eminent reference. This chapter was created, reviewed, and revised according to the general practices throughout industry by industry dissolution experts and should be relied upon for the best information on this subject.

There are two parts to the validation aspects. The most important is the product performance with the method, including robustness, ruggedness (intermediate precision), recovery (accuracy), selectivity (placebo interference), sample stability, sampling method, filtration, comparison dissolution results of manual versus automated, carryover in automation, and sinker validation (42). The other part is the determinative-step validation; this is the validation of the analytical method that is used for the sample aliquot analysis. This determinative step validation is covered thoroughly in the literature (43) and will not be covered in any detail in this chapter. However, certain aspects are critical to determination of the dissolution results: linearity, precision, and standard stability.

During the assessment of product performance with the dissolution method, some primary criteria have to be achieved before proceeding with the method validation. The variability and profile must be satisfactory; the method must be able to detect formulation and process changes. In other words, the method is meaningful, and results can be interpreted without being confounded by other factors. There should be no significant analytical solution stability problems.

Product Performance Validation Parameters

The validation begins with linearity and precision, with the interference of the placebo being well understood. Recovery experiments are next using typical 50%, 100%, and 125% points, or lowest expected profile concentration. The placebo mixture should include all excipients, the capsule shell, coating blend, ink, and sinker. The recovery experiment can be performed in vessel or a flask on the bench top with preheated medium. During recovery experiments, the order of addition (drug vs excipient) may be on a case-by-case basis depending on the physical characteristics of the excipients and drug substance. The drug is preferably added as a powder, but in circumstances where the amount of drug is very low or weighing may be inaccurate (hydrostatic nature), the drug may be first dissolved in an alcoholic solution and spiked into the vessel or flask. This is also decided case-by-case. Poorly soluble drugs may require more vigorous evaluation of the experimental steps. The spiked organic solutions (2% alcohol or less of final analyzed solution) may need longer mixing times and higher initial apparatus speed if performed in a vessel, especially if a powder is used. The usual criterion is 97–103% of the theoretical value.

The selectivity experiment should use the same placebo mixture as used in the recovery experiment. The placebo mixture should be stirred for at least one hour at high rpm. The wetting properties should be noted. There should be an equivalent amount of placebo mixture for highest and lowest strength and, when compared to the 100% standard, the acceptable interference should be not more that 2%.

For sample stability, the sample should be analyzed on day one, and then at intervals from 3 to 12 days. This stability interval depends on how many days may transpire before a re-reading of the sample is allowed by approvals mandated by SOPs. The usual criterion is 98–102% of the fresh sample reading. If UV analysis is the

analytical method of choice, an analysis of the UV samples by HPLC may be instructive, just in case there are hidden stability issues.

Filter validation is performed on both sample and standard solutions using 100% solution, although a range is more comprehensive. For standards solutions, compare filtered with unfiltered. For sample solution, compare filtered versus unfiltered but centrifuged sample solution. Be sure to use 100% dissolved sample, because lower time points may give ongoing dissolution during the centrifugation. The usual criterion is within 98–102% of the unfiltered standard and unfiltered/centrifuged sample solutions.

Robustness

The robustness is the most interesting validation parameter. This is where the really important variables are uncovered. This is vastly important as the dissolution test can be very technique-dependent for some compounds, especially those of low solubility. The impact of small changes within the dissolution test constitutes the robustness parameter. The most critical aspects are typically deaeration and medium concentration and pH. A comparison of deaerated media versus non-deaerated medium is one of the first method validation studies to be performed. It is not wise to generate lots of data using non-deaerated media only to discover many tests later that the presence of bubbles has an affect. When evaluating the effects of media concentration, levels that are 80%, 100%, and 120% of the chosen media may be used. Varying the medium pH by ± 0.5 pH unit will adequately assess the effects of pH. There are other optional changes: paddle height (± 0.5 cm), water bath temperature ($\pm 1^{\circ}$C), sample times (± 2 min), and rpm ($\pm 4\%$). Assessing the relationship of the dosage unit position in vessel (center versus off-center) to the dissolution results and variability is more challenging. And lastly, determine vibration sensitivity, which is usually discovered serendipitously, and rarely are experiments designed to assess this problem. The usual criterion for robustness is 3–5% of method conditions. It should be also pointed out that basket attachment design may affect the dissolution rate. This has been referenced (24,26) and deals with clipped (official USP design) versus o-ring attachment design. If both attachment methods are used or may be used in a transfer lab, it must be part of validation. There may be wide differences when different attachment types are used and therefore a troublesome method transfer issue.

Intermediate Precision

The ruggedness parameter is often referred to as intermediate precision. This is as close to a method transfer as one can get, so it should be treated as an early indication of possible method transfer issues. Therefore, the test parameters should be varied as much as is feasible, that is a different analyst, tester, spectrophotometer, flow cell, media, standard and buffer preparations, and autosampler, on different days and in another laboratory, if possible. The same sample should be tested using 12 units. All strengths should be tested or bracketed when 3 or 5 strengths are present. The usual criterion will consist of mean values within 3–10% from analyst A to analyst B and depends on time point and product variability.

Automated Methodology

There are special considerations when validating a method that has an automated component. Automation can be in many forms, from basic to fully automated systems. Automated systems can include fiber optics, hollow-shaft sampling, and in-residence probes. There are automated deaeration equipment, on-line UV testing, and robotics automation.

Regardless, the principles validating an automated method involve doing a manual sampling method and comparing the dissolution results to those obtained using an automated method. There are several sources of error that can come from automation; this is why a comparison of automated versus manual sampling is quite critical. The comparison experimental study for highly variability products would include simultaneous manual versus automated sampling at all time intervals. Calculations need to account for the duplicate volume lost. However, a strong caveat against this simultaneous manual versus automated sampling is that it will not assess sampling probe interferences. To better assess this critical parameter, concurrent testing is recommended. One to two runs of each dosage strength should be performed using manual and automated sampling. The usual criterion is 5–10% absolute difference for early time points with more variable data and 3–5% absolute difference for later points with >80% dissolved.

Other considerations in automated dissolution: While offering savings of resources and adding productivity to a laboratory, automation can have several drawbacks. Automated equipment requires setup time and validation. As mentioned, the analyst must show that the results are accurate compared to the manual method. Errors often occur with these devices because the analysts tend to overlook problem areas. Sample lines are often a source of error for a variety of reasons: unequal lengths, crimping, wear beyond limits, disconnection, carryover, mix-ups or crossing, and inadequate cleaning. The cleaning time and carryover procedures need to be evaluated. The volume dispensed, purged, recycled, or discarded should be routinely checked. Pumping tubes can wear out through normal use or repeated organic solvent rinsings and may necessitate replacement.

Time must be devoted to training, maintaining logbooks, calibration, and maintenance. There is down time when the equipment is broken and needs troubleshooting. Analysts may develop an approach where they drop the tablets and leave the testing area, ignoring valuable observations. Automated equipment occupies a large amount of lab space.

In the present atmosphere of computer validation, there is an additional aspect of verifying the software and hardware to meet compliance in this area.

The use of flow cells may generate variability in absorbance readings. Air bubbles can become trapped in the cell, either introduced via a water source containing bubbles or by air entering inadvertently into poorly secured sample lines. Flow rate and dwell time should be evaluated so the absorbance reading can be determined to have reached a steady plateau. Cells need to be cleaned frequently to avoid buildup of drug, excipient, surfactant, or buffer salts from the dissolution medium.

In automation, one of the most prevalent problems is carryover of residual drug in the autosampler lines. What are the proper cleaning/rinse cycles? Does one use an organic rinse, water, or a mixture of both? Also, what are the rinse times and what order? This elimination of carryover is best proven by following a run of the highest strength with a run using only dissolution medium. The typical allowance for carryover is 1% or less of 100% dissolved. Some other aspects of automated systems are accurate determination of the pump dwell times for flow cells, the sample line pull volume, sorption on the tubing, and evaluation of the filter type in the automated system, which is usually different from the filter used in the manual sampling. A frequent 483 warning comes from lack of proper validation, especially of automated methods.

Sinker

The validation of the sinker type is very critical as it has been shown that different sinkers can give different dissolution results. Sinkers other than those described in USP should be

evaluated by performing a concurrent test with the chosen sinker versus the USP wire sinker. One to two runs of each strength is sufficient. The usual criterion is the same as for intermediate precision and manual-versus-automated comparisons, that is, 5–10% absolute difference for early time points with more variable data and 3–5% absolute difference for later points with > 80% dissolved.

Determinative Step Attributes

The determinative step validation is quite straightforward and includes linearity, range, and precision. Up to 5% organic solvents (2% organic component preferred) should be used to enhance the solubility of drug in the final standard solution. The typical range is between 25% and 125% (3–5 points) label claim concentrations. If flow cells are used, a validation should be performed comparing standard absorbances using the flow cell versus those of manually diluted standards. All solutions are made from a common stock, using triplicate readings or duplicate injections. The usual linearity criterion is a correlation coefficient of > 0.997, with a Y-intercept of 2% or less of the 100% level standard.

The determinative step validation of precision is easily determined by using the linearity values. The usual criteria are 1–2% RSD for UV analysis and 2% RSD for HPLC injections. Studies of standard stability are performed by analyzing the standard solution on day one and then at intervals from 3 to 12 days. This stability interval depends on how many days may transpire before a re-reading of the sample is allowed by approvals needed in the SOP for re-running samples. The usual criterion is 98–102% of a fresh standard reading. The system suitability criterion for UV analysis is the precision stated above; however, a database of the typical absorptivity range with historical data is useful. With HPLC analysis there are usually retention time and precision criteria. Response factors are not too reliable but do afford some reassurance of a working system. A robustness attribute for the UV analysis is achieved by varying the wavelength (±2 nm). For the HPLC analysis, there are many ways to ascertain robustness; the most typical are by varying the column brand or age of the column, altering the mobile phase ratio (±10%), and changing pH.

METHOD TRANSFER

Problems that occur during transfer of methods can often be traced to the use of equipment that is not exactly the same, such as baskets/shafts, sinkers, dispensing apparatus, or sampling method. A precise description of medium and standard preparation, including grade/purity of reagents, may be useful. Common errors occur when the standard is made without alcohol and the sonication step is long. The use of alcohol is one of the most important ways to eliminate standard prep errors, and the detailed instructions for such are sometimes overlooked in the method transfer documentation.

The dissolution test involves many variables that can contribute to inaccurate results. The robustness component of validation can be very useful to point to weaknesses in the method and frequent sources of error. Also, there may be ambiguities in written test methods, where a lack of detail can be problematic. For instance, if the product is particularly sensitive to dissolved gases, the deaeration technique is a very important procedure that should be described in detail. Otherwise, there may be variable results from one lab to the next if different deaeration techniques are used. Other aspects of the test that should be described are the basket attachment type and mesh size. The sinker type is important as mentioned before; if it is handmade, the procedure should be included. In

some cases, the sample introduction technique needs to be described, especially in the case of suspensions. In some cases with suspensions, it must be specified if the paddle is running or not when the sample is introduced.

Rigorous method development and validation, proper calibration and operation of equipment, and thorough and frequent observations can assist in preventing and identifying sources of error associated with method transfer.

METHOD DEVELOPMENT

The Basics

As mentioned previously in this chapter, the new USP Chapter <1092> The Dissolution Procedure: Development and Validation (10) is a valuable guide for developing dissolution methods. Its purpose is to elaborate on dissolution validation, provide instructions on method development, and encourage new technology and equipment. There are many sources in the literature that give ample guidance on method development (44,45).

There are certain basic requirements for a good dissolution method. These requirements are low variability, a good profile, and the ability of the test to show changes in the product. Low variability is critical; comparing dissolution curves is meaningless if the standard deviation is so wide that the compared curves are indistinguishable. The test conditions must be such that any significant changes in the formulation, manufacturing process, drug substance, and during stability are revealed.

The hydrodynamic aspect of product mixing in the vessel is very important; this is where visual observations are necessary. Any artifacts such as tablet sticking, coning under the paddle, clogging of the basket screens, and/or floating chucks should be minimized, since these phenomena may affect the dissolution results. One should become very familiar with the Biopharmaceutics Classification System (BCS), for it is an excellent starting point for developing a dissolution testing method. The four categories are described in Table 1.

Drug Properties

Method development starts with obtaining as much knowledge as possible about the drug substance. In today's climate of QbD, this knowledge is paramount. As dissolution analysts, you may not have that much control over how much is known about the drug, but at least know the basics. The key properties of the compound are the pK_a, particle size range, solubility as a function of pH and surfactants, stability, the absorption site, and the BCS classification.

Dosage Form Properties

The dosage form properties are the disintegration rate, the functionality of the coating (e.g., enteric coated), modified release (e.g., extended, sustained, delayed), presence of solubility enhancers, and excipients.

TABLE 1 Biopharmaceutics Classification System

Class 1	Class 2	Class 3	Class 4
Highly Soluble Highly Permeable	Poorly Soluble Highly Permeable	Highly Soluble Poorly Permeable	Poorly Soluble Poorly Permeable

Dissolution Profile

Ideally, unless the drug is a BCS Class I drug that is 80% dissolved in 15 minutes using one of the three preferred media (0.1 N hydrochloric acid, acetate buffer pH 4.5, or phosphate buffer pH 6.8), it will be necessary to develop a method that yields a dissolution curve with a reasonable profile shape (Fig. 10). In other words, the dissolution rate should be gradual so that results can be compared using several time points. The similarity factor, f2, discussed in many FDA guidances uses at least three points, with only one point allowed above 85%. This further encourages the analyst to demonstrate a gradual profile. There are many ways to "slow down the profile." One can decrease the apparatus speed or medium flow rate, manipulate the molarity of the buffers and acids used, change the pH, or change the apparatus. One favorite method of the author is to use the 0.01 N hydrochloric acid medium with Apparatus 1 at 50 rpm. This seems to slow down the dissolution rates of many dosage forms, but it is worthwhile only if the product is compatible with pH 2 medium and does not cause clogs in the basket mesh.

Media

Choices of media include acids (hydrochloric acid 0.1–0.001 N); buffers (use USP preparation instructions), namely acetate (pH 4.1–5.5, 0.05 M) and phosphate (pH 5.8–8.0, 0.05 M); and simulated fluids without enzymes (gastric and intestinal). Water may not be appropriate as it affords no buffering capacity, and the pH cannot be measured accurately. The conductivity or pH may vary depending on the water source. However, there are advantages in that water is inexpensive, and disposal is relatively easy. For very poorly soluble compounds, aqueous solutions may be modified to contain a percentage of a surfactant to enhance drug solubility. The need for surfactants and the concentrations used must be justified by showing dissolution profiles at several different

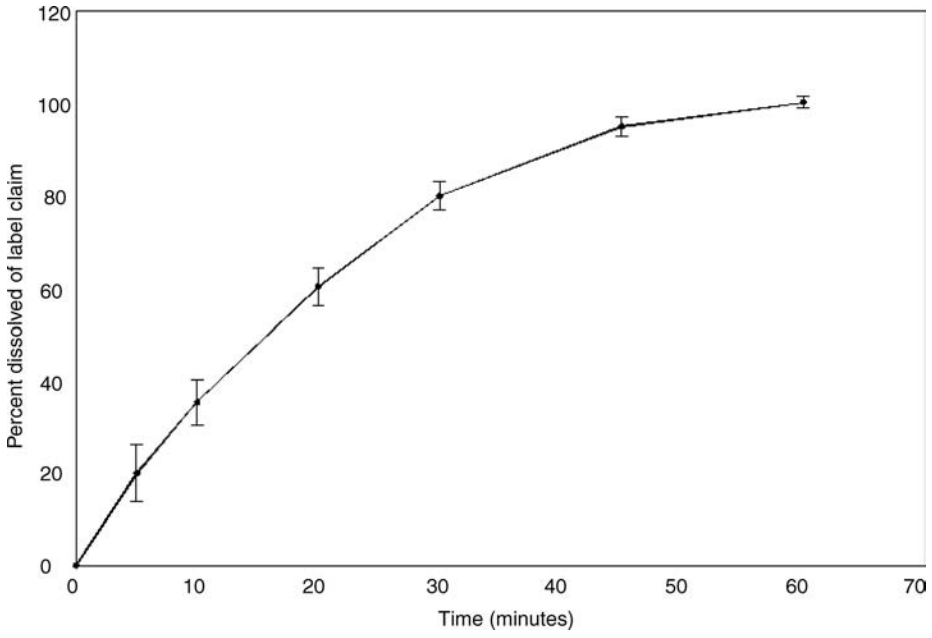


FIGURE 10 Typical dissolution curve for immediate release.

surfactant concentrations. Surfactants can be used either as wetting agents or, when the critical micelle concentration is reached, to solubilize the drug substance. There are many surfactants available. Some examples are SLS, polysorbate 20–80, cetrimide, lauryldimethylamine oxide, bile salts, Brij™, Triton X™, Solutol™, and cremophor. Combinations of surfactants and buffers/acids are also very useful when the pH needs to controlled and solubility is an issue. Molarity changes can change dissolution rate.

Other media are mixtures of aqueous and organic components and buffers above 8 pH. When looking for extensive biorelavance in the dissolution media, fed and fasted, gastric and intestinal media are well discussed in the literature (46–51).

There are analytical considerations when using surfactants (wetting agents/solubilizing agents). SLS is a mixture and therefore can have purity issues (x). Cetrimide may be viscous at certain concentrations and make auto-injection and other handling issues troublesome. The same issues are seen with Tween, where column cleaning is necessary to avoid split or broadening peaks.

Volume

The medium volume is typically 900 mL, with 500 mL for low dosage strengths. The volume may be increased to 1, 2, or 4 L. For the special needs of low dosage strengths, volumes of 200 mL or less may be necessary (52,53).

Deaeration

As mentioned in the validation section, deaeration is a critical variable that needs to be performed if the presence of air bubbles affects the results (17,20–22). Deaeration of surfactants may not be practical due to foaming and may not be necessary (54).

There are a multitude of deaeration methods available: the USP method involving heat, filtration, and vacuum (9); helium sparging; and automated methods.

Speed

The typical rotation speeds for the paddles are 50 rpm (the preferred speed for BCS), 75 rpm to eliminate coning and variability, or 25 rpm or more for suspensions. A speed of 100 rpm or higher requires justification; however, 100 rpm is used frequently with ER products. For the basket, 50–100 rpm is preferred but speeds greater than 100 rpm are sometimes necessary.

Sinkers

Sinkers are a vital part of the dissolution method. As mentioned before, the uniformity is critical, especially when transferring method. According to the USP, other "validated" sinkers can be used with proper validation (9). The point is that different sinkers have significantly different mixing characteristics and can yield different dissolution results. The sinker can be a barrier to dissolution if it is wound too tightly around the product or has too many coils. This is also a problem if an exploding type of disintegrant is used. The sinker may restrict this action and inhibit the dissolution rate.

Filtration

In method development, filter use is necessary for most products, and centrifugation is not preferred because the dissolution can continue, plus centrifugation is time consuming. When selecting a filter, its compatibility with the media and formulation has to be considered, and the usual validation must occur before the filter is used routinely. Filters

are made of many different materials (e.g., nylon, polyethylene, and glass fiber). There are several types and positions of filters: in-line; at the end of the probe or cannula; disk; or in earlier days, a stainless steel filter holder. The pore size of the filters typically is in the range of 0.20–70 μm, with depth or full flow in design.

Time Points

For immediate release, there is the possibility of a five-minute time point where disintegration occurs or is partially completed. This time point may give profile information, especially with suspensions, or be useful in accumulating the necessary three points for an f2 comparison. The other intermediate points are 10, 15, or 20 minutes; any of these points will be useful for a profile and f2 comparison, and in some cases, the specification will be at one of these earlier points. For example, a BCS Class I or a suspension may have a Q-value at these points. The later points of 30, 45, and 60 minutes will be necessary for the typical specification for immediate release, and the test for a poorly soluble drug may go even longer (up to 3 hours in some cases). If complete (100%) dissolution is present at 30 minutes, the 60-minute time point will not be necessary. It is always prudent, however, to keep one extra point past the 100% dissolved point in case there is a decrease in the dissolution rate on stability.

Fast Stir or Infinity Point

After sample has been drawn for the last time point, the rpm may be increased to 150–200 rpm for another 15–30 minutes. This is done to provide a completely dissolved sample in the vessel. Take the sample, and since you will have at least 6 sample readings, there is a data set that is appropriate to compare with the content uniformity data for the product. Comparison of the fully dissolved samples versus label claim will give an early read on recovery and variability. If the content uniformity data are different in either potency or variability, this provides additional information for assessing the method.

Time Points for ER Products

A minimum of three time points are required for ER products. There will be a time point in the first hour or two to measure the potential for dose dumping; a midway point at around 50% dissolved; and a NLT end point where typically at least 80% is dissolved or an asymptote is reached. Other time points may be useful, especially if the test continues for longer than 8 hours. With extended or modified-release dosage forms, it is sometimes difficult to achieve 100% dissolved. This can be caused by the matrix holding on to the drug in such a way that not all of it is exposed to the media and readily dissolved. A fast stir is also not practical with a modified-release product unless 100% dissolved is achievable, then the information would be useful when compared to the content uniformity results.

Poorly Soluble Drugs and Novel Dosage Forms

The classification system is a first step toward dissolution method development. Class II is the most common type of drug and most challenging when developing a discriminating dissolution test. Classes II and IV are the best for in vitro-in vivo correlation because the dissolution is the rate-limiting step in these drugs. For Class I compounds, select one of the three media for the regulatory test but obtain profiles in the other media for future comparisons. The medium with the slowest profile is usually picked for f2 points. To

select media for the poorly soluble drugs, examine the media listed for Class I and, if you are lucky, use any that will afford a good dissolution rate. Usually, however, surfactants are usually needed. Surfactants are cationic, anionic, or nonionic. Chose the one whose chemical nature is most appropriate for the drug substance, starting with a 1–2% concentration, or if predetermined, the concentration needed to achieve sink conditions.

Sink Conditions

Sink conditions are the focus of poorly soluble drugs. There are several options for achieving sink conditions when developing a method. The surfactant concentration can be altered, as previously mentioned, or there can be increased media volume through the use of 2- or 4-L vessels. The use of Apparatus 4 is an option, since infinite sink is obtained with the constant flow of media over the dosage unit.

Establishing and maintaining sink conditions during the dissolution test is an important criterion for the dissolution method, because the true dissolution rate should be measured and not be overlapping in the area of concentration equilibrium. As the solution into which the drug is dissolving becomes more concentrated, the dissolution rate will decrease. In the USP General Chapter < 1088> In Vitro and In Vivo Evaluation of Dosage Units (55) it states, "The quantity of medium used should be not less than 3 times that required to form a saturated solution of the drug substance."

Media

The typical media (0.1 N HCl, pH 4.5 acetate, pH 6.8 phosphate) will usually not give the needed solubility. Simulated Gastric and Intestinal fluids without enzymes are also used but with the same issues. Not until surfactants are used is an appropriate media usually found. SLS is one of the most prevalent. However, there are considerations with this surfactant. As mentioned before, the product is a mixture, so purchasing the most pure form is important. There are also stability problems below pH 2.5. This surfactant will also denature the enzymes typically used in two-tier testing, pepsin and pancreatin, making it difficult to use when a capsule product shows failed dissolution results due to cross-linking. If using SLS in combination with pH 6.8 buffer, it is important to use the phosphate sodium salt and not the potassium salt, because this mixture forms a precipitate at room temperature (56).

Apparatus Selection

Apparatuses 1 or 2 should be the first choice. Apparatus 3 is a good research tool and may be useful for enteric-coated product and some other dosage forms like soft-gel capsules or ER beaded products. Apparatus 4, the flow-through cell, with the open system can provide infinite sink conditions. In both Apparatuses 3 and 4, media can be changed during the test. Apparatus 7 has some utility for extended release, transdermals, and stents/implants.

Novel Dosage Forms

There are many new products with in vitro release delivery systems (e.g., microspheres, liposomes, modified release parenterals, implants, stents, and granules). There is no official methodology, and when the official USP Apparatuses 1–7 are not appropriate for these dosage forms, non-compendial apparatus come into use. These apparatus can include static tubes with dialysis membranes, modifications of Apparatuses 4 and 7, and small-volume apparatus.

Suspensions: In the case of a product where the particles float and are not immediately soluble, there are special considerations. The reconstitution process needs to be evaluated for consistency—is it hand-shaken or is a mechanical shaker used? Surely a patient does not have a mechanical shaker. There are different ways to introduce the liquid sample (57), with many devices available (e.g., Eppendorf pipet, tared beakers, syringes fitted with needles that have tubing at the end).

The paddle may need to be rotated when the sample is introduced to keep the suspension from dropping to the bottom of the vessel in a glob. Is the sample introduced gravimetrically or volumetrically? Air bubbles are a problem for volumetric delivery. These are aspects to consider when developing methods for suspensions. The earlier time point will be most meaningful, since some suspensions do dissolve slowly. On stability, the freeze thaw cycles for suspensions are instructive. The particle size for the conventional suspension is the most important aspect indicated by the dissolution test.

Microspheres/nanoparticles: The dispersion pattern is the problem with these dosage forms. The particles can float and not mix well. There have been several apparatus modifications (e.g., dialysis bags, static tubes, rotating bottle, Apparatuses 4 and 3) (58,59).

Implants/stents: For these slow releasing products, acceleration by increasing the bath temperature from 45˚C to 55˚C is under consideration or the conditions do not yield 100% dissolved. Typical equipment under consideration are the rotating bottle, Apparatus 4 with a special cell design, and Apparatus 7 using a modification of designs for ER dosage forms.

Liquid-filled capsules: Soft gelatin capsules and liquid-filled, hard gelatin capsules were exempt from dissolution testing until the early 1990s when the USP eliminated the exemptions for these products. At this time, USP went out to industry to encourage more dissolution tests, but none were forthcoming, since soft gelatin products that are lipid filled are not apt to dissolve very well in typical media. As an interim move, USP instated a rupture test. For an example of this test see the Ergoloid Mesylates Capsules monograph (60). This was a visual test that included water media with the paddle at 50 rpm. The tolerance was the time, usually 30 minutes, when the rupture of the capsule should have occurred. For the aqueous soluble fill, this was a good indicator of dissolution, since the solution will readily be available for absorption. However, with oil-filled capsules, the rupture time is only half the story, leading to a push for a dissolution test for these products. Methods have been developed for these liquid-filled capsules and are sometimes quite a stretch. Media composed of 5–10 % SLS have been noted; other surfactants [cremophor, Solutol™ (BASF, Ludwigshqfen, Germany)] have been successfully used. Sometimes the dose strength is very low, necessitating the use of LC/MS detection and small-volume apparatus. Apparatus 3 and the paddle have also been used with some success. More attention is now focused on these challenging dosage units as reflected by an article to be published in 2008 from USP on the subject.

Analytical issues: With novel dosage forms, the release is usually extended over a period of time, and even if the drug is moderately soluble, it is usually in a matrix that will control the release. With a very slow release time, the prevalent media will still be surfactants. At times, there are extrusion issues with polymeric formulations, making filtering difficult and necessitating protection for the HPLC column. Fiber optics have been used successfully in some cases. A special edition of Dissolution Technologies was devoted to the subject of fiber optics in dissolution testing in November 10(4) of 2003.

Two-Tier Testing

When pellicles or cross-linking occur with capsules, the dissolution test may fail. In USP <711>, the addition of enzymes is now allowed for these products, but there are still some outstanding issues. The instructions state to add pepsin for water or media with a pH of less than 6.8. Pancreatin is added for media at or over pH 6.8. The problem is that pepsin is not optimally active at a pH between 4 and 6. This has yet to be resolved.

Method Examples from USP Monographs

The USP contains interesting methods that are not the typical procedures. This is good to know because as methods for more challenging products are developed, these variations of the typical procedures may be useful alternatives. For example, in the immediate-release carbamazepine tablet monograph, there are multiple dissolution tests, a test for a 100-mg chewable tablet, and a procedure that calls for the use of Apparatus 3. Also in this monograph are instructions to use methanol in the standard solution to facilitate dissolution of the poorly soluble carbamazepine. The Apparatus 3 method includes the addition of two drops of simethicone to each vessel; presumably, this is because the speed of 35 dips per minute with a surfactant media will generate foaming. The Diltiazem HCl Tablet monograph includes two time points with a long time point, 3 hours, for the final Q. The early time point of 30 minutes and Q of not more that 60% is to detect dose dumping example of a suspension dissolution test is seen in the Indomethacin Oral Suspension monograph. The sample addition technique includes transferring the sample to the media surface, with instructions to be sure the sample is free of air bubbles. There is an early specification, 80% (Q) in 20 minutes.

The dissolution test for Theophylline, Ephedrine HCl, and Phenobarbital Tablets is an example of pooled dissolution testing. This type of dissolution test is found in some monographs with multiple active ingredients, an HPLC finish, and a well-known history of uncomplicated dissolution results that were not highly variable. The pooled dissolution procedure combines one aliquot from each of six vessels into a common flask where is it only necessary to analyze one sample. The acceptance criteria are tighter, with $Q + 10\ \%$ rather than $Q \pm 5\ \%$, using the average dissolution result rather than individual results. Pooled dissolution was intended to save resources, especially mobile phase and time, with just one injection per time point. It was implemented in about 60 USP monographs. However, some companies did not want to re-validate their dissolution analytical methods or were automated to sample six vessels, so no additional dissolution tests were converted to pooled dissolution. A suppository dissolution test is found in the Indomethacin Suppository monograph. This dissolution test uses paddles at 50 rpm with a 60-minute Q, using pH 7.2 phosphate buffer as the media. An example of delayed-release testing in the Aspirin Delayed-Release Tablet monograph uses Method B <711> with a longer buffer stage, going to 90 minutes; in addition, the analysis is measured at the isosbestic point for aspirin and salicylic acid. An example of ER testing is seen in the Theophylline Extended-Release Capsules monograph. Here there are many drug release tests listed in product-dosing intervals. Some tests use the Delayed-Release Method A; there are many different media, apparatus, speeds, and timepoints. Why so many tests? This is because the ER formulations have different release mechanisms; however, they are all approved products that are bioequivalent to a reference product. In the Nifedipine ER Tablets test, Apparatus 7 is used. The test requires the rod design, one of the five Apparatus 7 designs.

The Ergoloid Mesylates Tablets dissolution test is unusual since the distance between paddle blade and the inside of the bottom of the vessel is maintained at 4.5 ± 0.2 cm during the test, a strange paddle height. To date, there has been no explanation of why this is so, other than that the product was approved using this test.

HARMONIZATION

In April 2006, the EP, USP, JP, and BP all harmonized the general chapters on dissolution and drug release. The harmonized chapter combines <711> Dissolution USP Chapter with elements of <724> Drug Release General Chapter. Therefore, Apparatuses 1–4 are described in <711> along with the acceptance tables for delayed- and ER products. Some elements are still not harmonized since the JP does not recognize Apparatus 3 (Reciprocating Cylinder). JP also follows a separate approach to delayed-release products, serial versus concurrent. Harmonizing the name for each release category was not accomplished. The basket wire diameter dimensions are widened to 0.25–0.31 mm to accommodate all regions. This may present method transfer issues when results from baskets at one extreme of the range are compared with results generated at the other end of the range. This needs to be further studied. The specifications are harmonized with the USP Acceptance Criteria required in the other pharmacopeia, with all stages 1–3 present. The other three Acceptance Tables for ER and delayed-release (acid and buffer stage) are included.

CONCLUSIONS

There are challenges to the dissolution test today. The dissolution test has been under scrutiny in several areas: the quality-by-design initiative has called for the end to dissolution testing along with all end-product testing (61–63); there is a push for more clinically relevant specifications (64); the flaws in the hydrodynamic fluid-flow patterns that emerge from the vessel and paddle interaction is being closely examined (65–68); and the use of the calibrator tablets has been questioned (69).

The QbD and PAT initiatives urge companies to know their drugs and drug products much more thoroughly than is the present practice. Nothing is more disheartening than to see a significant change in the dissolution results on stability of a Phase 3 product or on a release batch of a commercial product. It is even more discouraging when an assignable cause is not forthcoming. The increased knowledge expected from PAT may prevent these "surprises," and that would be a welcome change. The dissolution test is sensitive to an infinite number of parameters from characterizations of the drug to formulation changes and, most importantly, manufacturing parameters. To be able to show changes in these many parameters is the power and the frustration of the dissolution test. The power of the test outweighs the frustrations because of the simple reason that the dissolution test is the only test that has some degree of relevance to the drug's therapeutic effect in vivo.

To eliminate dissolution as an end-product test would be problematic from two angles. Can you be sure you have found all of the infinite sources of potential change in the final product with your early testing? How do you measure the stability of the finished product unless you test it at release and then over its shelf life? What is the value of eliminating a proven indicator of stability?

The need to have more clinically relevant dissolution specifications and methods is laudable. The method development stage is extremely critical for this to be accomplished. Many a naïve manager views the dissolution test as a simple test until a problem occurs,

only to find the staff may not be experienced or versed in the test nuances or sources of error (16). A separate dissolution group is the optimal way to handle dissolution method develop and even routine testing. A group allows better training, increased experience your product line, and useful collaboration to take place. Also, a separate lab that is devoted to dissolution testing will help avoid problems that can come from equipment problems stemming from vibration and other related issues.

Finding the appropriate method and specifications, especially with the typical low solubility, takes time and resources. Cutting corners at this stage is very risky. The robustness and variability of the method should be examined thoroughly. As mentioned earlier guidance on method development is abundant throughout the literature, other forms of instruction on method development are the FDA guidances, The new USP Chapter <1092>, the AAPS in Vitro Release and Dissolution Testing Focus Group, books (70–72), and websites with chat room bulletin boards or Q and A possibilities (73,74).

Early in method development, the variability should be examined. High variability is problematic making trend analysis and f2 calculations difficult. Most importantly at this stage, the source of variability should be isolated and understood. The physical dissolution process should be observed for any anomalous stirring; the test should show gentle homogenous mixing. Observation of the hydrodynamic flow of the fluid is very important at this point. Any coning (a concentrated gathering of excipients and drug under the paddle), tablet-sticking, air bubbles, or off-center placement of the dosage form should be noted and the dissolution rate examined to see if there is a correlation. If so, all efforts should be taken to minimize this anomalous behavior. Our ultimate nightmare is a recall due to dissolution failure. At the method development stage, all aspects of the mechanical or physical dissolution test that can affect the results should be illuminated and minimized, so that if a dissolution test failure occurs later on, the failure can, with confidence, be attributed to some change in the dosage form.

When the time comes to set specifications, the sponsor and FDA must collaborate to make the specifications appropriate. A most critical step in the approval process is the fine line of setting a specification that will not allow bioinequivalent batches to pass, yet not be too tight as to fail good (meaning fully effective in vivo) batches that may change slightly. In some instances, a specification is too borderline, and over time, the product goes more and more to stage 2—this may be a scenario that will produce later failures and recalls. Hence, special care should be taken to understand critical parameters and, especially, the stability behavior of the product.

In later phases of the product, the method development and validation should include robustness of the method. At this time, the aspects of the test that may influence the dissolution rate should be examined. Typical parameters such as temperature changes, changes in media concentration, basket attachment type, paddle height, changes in media pH, and many other aspects should be altered within a small tolerance range to see if the dissolution rate is sensitive to these changes. Other areas such as the presence of air bubbles, dosage form position in the bottom of the vessel, and other potential sources of variability should examined. This helps in understanding where the method is robust or overly sensitive, and detailed instructions can be incorporated into the test method or the test can be modified. The importance of the method development and validation stage cannot be overemphasized—it assists in knowing and characterizing the product well and even in predicting the in vivo behavior when an in vivo–in vitro correlation is developed. Problems with variability, poor mixing, or fluid flow usually can be overcome with appropriate change in apparatus type, speed of rotation, sinkers, or even media choice.

A discussion of the dissolution equipment is important since the dissolution rate is generated by the stirring mechanism interacting with the dosage form in the media. But

always be aware that the dissolution equipment is a machine. The initial quality, care, and maintenance will influence the operation and product dissolution rate generated by that machine. Any machine will wear out over time, a lemon could be purchased, the environment in which it operates will affect its performance, and it needs to be running properly at all times. Presently, calibrator tablets are tested every six months to assess the performance of the dissolution equipment.

It has been suggested in the literature that new apparatus for dissolution testing may be better designed to give less variability and more homogenous mixing or even be more easily correlated to in vivo performance of the product (75,76). There has been new technology that has added to the utility of the dissolution test. Fiber optics is one very useful tool as is increased automation of on-line testing. Different types of premixed media also add to the efficiency of the test. With novel dosage forms, the other official Apparatuses 3, 4, and 7 are becoming more suitable as are modifications of this equipment. There are performance tests that may not use the official equipment for unique dosage forms; this is fitting and should not be resisted if the advantages are truly apparent. However, for the immediate-release and ER dosage forms, typically Apparatuses 1 and 2 can provide appropriate methods with special care and study during the method development stage. There are probably 700 compendial tests that use the present apparatus with those tests being used for any number of product brands. At this time many new products are being approved with the use of either Apparatuses 1 and 2. The investment of resources and scientific data and backing for these apparatus is indisputable. Newly designed equipment will have to go through the same rigors and qualification as the present apparatus and will, by virtue of the testing the dissolution rate, be sensitive to the same parameters that influence the present equipment. The imposition on the industry of purchasing new equipment would not be welcome. From the podium, the regulatory agencies have many times discouraged the proliferation of new equipment types.

A more thorough understanding of drug substance and product in the early development stages as recommended will benefit the industry without doubt. The more careful training and experience of analysts is of paramount importance so that sources of variability are minimized and sensitivity to critical parameters is maximized during the method development stage. New equipment that significantly adds to the development of a proper in vitro release test is a worthy endeavor. Until there are appropriate mechanical means to detect vibration and vessel asymmetry, the calibrator tablets are our best tool. However, a search for better ways to characterize the equipment should continue (77).

REFERENCES

1. Waiver of In Vivo Bioavailability and Bioequivalence Studies for Immediate-Release Solid Oral Dosage Forms Based on Biopharmaceutics Classification System, Guidance for Industry, Washington, DC: U. S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research, 2000.
2. Disintegration <701>. United States Pharmacopeia and National Formulary; 30th ed. Vol. 1. Rockville, MD: United States Pharmacopeial Convention, Inc., 2007: 276–7.
3. Noyes A, Whitney W. The rate of solution of solid substances in their own solutions. J Am Chem Soc 1897; 19:930.
4. Gelatin Capsule Working Group. Collaborative development of the two-tier dissolution testing for gelatin capsules and gelatin-coated tablets using enzyme-containing media. Pharm Forum 1998; 24(5):7045–50.

5. Hoag SW, Ramachandruni H, Shangraw RF. Failure of prescription prenatal vitamin products to meet USP standards for folic acid dissolution. J Am Pharm Assoc 1997; 37(4):397–400.
6. USP Policy on Dissolution Standards. Pharm Forum 1981; 7:1225–6.
7. Wagner J, Pernarowski M. Biopharmaceutics and Relevant Pharmacokinetics. Drug Intelligence Publications. 1971.
8. Verified on June 9, 2007, http://www.ich.org/cache/compo/276-254-1.html.
9. Dissolution <711>. United States Pharmacopeia and National Formulary; 30th ed. Vol. 1. Rockville, MD: United States Pharmacopeial Convention, Inc., 2007:277–84.
10. The Dissolution Procedure: Development and Validation <1092>. United States Pharmacopeia and National Formulary; 30th ed. Vol. 1. Rockville, MA: United States Pharmacopeial Convention, Inc., 2007:579–84.
11. Gray VA, Brown CK, Dressman JB, Leeson LJ. A new general chapter on dissolution. Pharm. Forum 2001; 27(6):3432–9.
12. FIP Guidelines for Dissolution Testing of Solid Oral Products, 1997, Dissolution Technologies 1997; 4(4):5–14.
13. Note for guidance on the investigation of bioavailability and bioequivalence, CPMP, The European Agency for the Evaluation of Medicinal Products Evaluation of Medicines for Human Use, 2001.
14. Verified on June 9, 2007, http://www.nihs.go.jp/drug/DrugDiv-E.html,
15. Brinker G, Goldstein B. Bathless dissolution: validation of system performance. Dissolution Technologies 1998; 5(2):7–14, 22.
16. Gray V. Identifying sources of error in calibration and sample testing. Am Pharm Rev 2002; 5:8–12.
17. Nithyanandan P, Deng G, Brown W, Manning R, Wahab S. Evaluation of the sensitivity of the USP Prednisone Tablets to dissolved gas in the dissolution medium using USP Apparatus 2. Dissolution Tech 2006; 13(3):15–8.
18. Deng G, Hauck WW, Brown W, Manning R, Wahab S. Perturbation study of the dissolution apparatus variable–a design of experiment approach. Dissolution Tech 2007; 14(1):20–6.
19. Liddell MR, Deng G, Hauck WW, Brown WE, Wahab S, Manning R. Evaluation of glass dissolution vessel dimensions. Dissolution Tech 2007; 14(1):28–33.
20. Moore TW. Dissolution testing: a fast, efficient procedure for degassing dissolution medium. Dissolution Tech 1998; 3(2):3–5.
21. Queshi SA, McGilveray IJ. Impact of different deaeration methods on the USP dissolution apparatus suitability test criteria. Pharm Forum 1994; 20(6):8565–6.
22. Degenhardt OS, Waters B, Rebelo-Cameirao A, Meyer A, Brunner H, Totli NP. Comparison of the effectiveness of various deaeration techniques. 1998, Dissolution Tech 2004; 11(1):6-5.
23. Collins CC. Vibration: what is it and how might it effect dissolution testing. Dissolution Tech 1998, 5(4), 16–8.
24. Crist B, Spisak D. Evaluation of induced variance of physical parameters on the calibrated USP dissolution apparatus 1 and 2. Dissolution Tech 2005; 12(1):28–34.
25. Vangani S, Flick T, Tamayo G, Chiu R, Cauchon N. Vibration Measurements on dissolution systems and effects on dissolution Prednisone Tablets RS. Dissolution Tech 2007; 14(1):6–14.
26. Gray V, Beggy M, Brockson R, Corrigan N, Mullen J. A comparison of dissolution results using O-ring versus clipped basket shafts. Dissolution Tech 2001; 8(4):8–11.
27. Scott P. Geometric irregularities common to the dissolution vessel. Dissolution Tech 2005; 12 (1):18–21.
28. Cox DC, Wells CE, Furman WB, Savage TS, King AC. Systematic error associated with apparatus 2 of the USP dissolution test II: effects of deviations in vessel curvature from that of a sphere. J Pharm Sci 1982; 71:395–9.
29. Tanaka M, Fujiwara H, Fujiwara M. Effect of the irregular inner shape of a glass vessel on prednisone dissolution results. Dissolution Tech 2005; 12(4):15–9.
30. Borst I, Ugwu S, Beckett AH. New and extended application sfor USP drug release apparatus 3. Dissolution Tech 1995; 2(2):1–8.

31. Takiar NB, Hollenbeck RG. In vitro evaluation of drug release from modified release delivery systems: initial experiences with calibrators for the USP dissolution apparatus 3. Dissolution Tech 1997, 4(3):5–8.
32. Rohrs BR. Calibration of the USP 3 (reciprocating cylinder) dissolution apparatus. Dissolution Tech 1997; 4(2):11–8.
33. Crison JR, Weiner ND, Amidon GL. Dissolution media for in vitro testing of water-insoluble drugs, effect of surfactant purity and electrolyte on in vitro dissolution of carbamazepine in aqueous solutions of sodium lauryl sulfate. J Pharm Sci 1997; 86(3):384–8.
34. Soltero RA, Hoover JM, Jones TF. Standish M. Effects of sinker shapes on dissolution profiles. J Pharm Sci 1989; 78(1):35–9.
35. Sarapu AC, Lewis AR, Grostic MF. Analysis of PMA collaborative studies of dissolution test calibrators. Pharm Forum 1980; 6:172–6.
36. PhRMA Dissolution Calibration Subcommittee. Dissolution calibration: recommendations for reduced chemical testing and enhanced mechanical calibration, Pharm Forum 2000; 26: 1149–66.
37. Beyer W, Smith D. Unexpected variable in the USP/NF rotating basket dissolution rate test. J Pharm Sci 1971; 60:2350–1.
38. Hanson W. Solving the puzzle of random variables in dissolution testing. Pharm Tech 1977; 1:30–41.
39. Thakker K, Naik N, Gray V, Sun S. Fine tuning of the dissolution apparatus. Pharm Forum 1980; 6:177–85.
40. Gray V. Compendial Testing Equipment; Pharmaceutical Dissolution Testing, Dressman J, Krämer J. eds. Boca Raton, FL: Taylor and Francis, 2005:41–3.
41. Gray V, Miller B. Current good manufacturing practices in the dissolution laboratory. Pharm Can 2002; 3(3):19–21.
42. Solving practical problems, method development, and method validation. In: Hanson R, Gray V, eds, Handbook of Dissolution Testing, 3rd ed. Hockessin, DE: Dissolution Technologies, 2004:136–9.
43. Validation of Compendial Methods <1225>. United States Pharmacopeia and National Formulary; 30th ed. Vol. 1. Rockville, MD: United States Pharmacopeial Convention, Inc., 2007; 680–3.
44. Brown CK, Chokshi HP, Nickerson B, Reed RA, Rohrs, B, Shah PA. Dissolution testing of poorly soluble compounds. Pharm Tech. 2004; 28:56–43.
45. Solving practical problems, method development, and method validation; In: Hanson R, Gray V, eds. Handbook of Dissolution Testing, 3rd ed., Hockessin, DE: Dissolution Technologies, 2004:128–36.
46. Dressman JB, Reppas C. In vitro-in vivo correlations for lipophilic, poorly water-soluble drugs. B.T. Gattefosse 2000; 93:91–100.
47. Nicolaides E, Symillides M, Dressman JB, Reppas C. Biorelevant dissolution testing to predict the plasma profile of lipophilic drugs after oral administration. Pharm Res 2001; 18(3):380–8.
48. Horter D, Dressman JB. Influence of physicochemical properties on dissolution of drugs in the gastrointestinal tract. Adv Drug Del Rev 2001; 46:75–87.
49. Kostwicz ES, Brauns U, Becker R, Dressman JB. Forecasting the oral absorption behavior of poorly soluble weak bases using solubility and dissolution studies in biorelevant media. Pharm Res 2002; 19(3):345–9.
50. Lobenberg R, Kramer J, Shah VP, Amidon GL, Dressman JB. Dissolution testing as a prognostic tool for oral drug absorption: Dissolution behavior of glibenclamide. Pharm Res 2000; 17:439–44.
51. Marques M. Dissolution media simulating fasted and fed states. Dissolution Technologies 2004; 11(2):16–7.
52. Craig DJ, Tuis A, Dansereau R. Is the use of a 200 mL vessel suitable for dissolution of low dose drug procedures? Int J Pharm 2004; 269(1):203–9.
53. Klein S. The mini paddle apparatus-a useful tool in the early developmental stage? Experiences with immediate release dosage forms. Dissolution Tech 2006; 13(4):6–11.

54. Flisar KA, Forsyth RJ, Li Z, Martin GP. Effects of dissolved gases in surfactant dissolution media. Dissolution Tech 2005; 12(3):6–10.
55. In Vitro and In Vivo Evaluation of Dosage Forms <1088>. United States Pharmacopeia and National Formulary, 30th ed. Vol. 1. Rockville, MD: United States Pharmacopeial Convention, Inc., 2007:532–8.
56. Stippler E, Kopp S, Dressman JB. Comparison of U.S. Pharmacopeia Simulated Intestinal Fluid TS (without pancreatin) and phosphate standard buffer pH 6.8, TS of the International Pharmacopoeia with respect to their use in the in Vivo dissolution testing. Dissolution Tech 2004; 11(2):6–10
57. Palmieri A, Gray V. Dissolution of Heterogeneous Dosage Forms. Dissolution Theory, Methodology, and Testing. Palmieri, A, ed., Dissolution Tech 2007: 232–240.
58. Kostanski J, DeLuca P. A novel in vitro release technique for peptide-containing biodegradable microspheres. AAPS PharmSciTech 2005; 6(2).
59. Zolnik BS, Raton J-L, Burgess DJ. Application of USP Apparatus 4 and in situ fiber optic analysis to microsphere release testing. Dissolution Technologies 2005; 12(2):11–4.
60. United States Pharmacopeia and National Formulary, 30th ed. Vol. 2. Rockville, MD: United States Pharmacopeial Convention, Inc., 2007:2049.
61. Woodcock J. The concept of pharmaceutical quality. Am Pharm Rev 2004; 7(6):10–5.
62. Hussain AS. Quality by design: next steps to realize opportunities, presentation to the Food and Drug Administration Advisory Committee for Pharmaceutical Science: Manufacturing Science Subcommittee, September 17, 2003.
63. Hussain AS. Biopharmaceutics and drug product quality: performance tests for drug products, a look into the future USP Annual Scientific Meeting "The Science of Quality". September 26–30, 2004
64. Zhang H, Xu L. Dissolution testing for solid oral drug products: theoretical considerations. Am Pharm Rev 2004; 7(5):26–9.
65. Missel PJ, Stevens LE, Mauger JW. Reexamination of convective diffusion/drug dissolution in a laminar flow channel: accurate prediction of dissolution rate. Pharm Res 2004; 21(12): 2300–6.
66. Kukura J, Baxter JL, Muzzio FJ. Shear distribution and variability in the USP apparatus 2 under turbulent conditions. Int J Pharm 2004; 279:9–17.
67. Healy AM, McCarty LG, Gallagher KM, Corrigan GI. Sensitivity of dissolution rte to location in the paddle dissolution apparatus. J Pharm Pharmacol 2002; 54:441–4.
68. Mirza T, Joshi Y, Liu G, Vivilecchia R. Evaluation of dissolution hydrodynamics in the USP, Peak™ and flat-bottom vessels using different solubility drugs. Dissolution Tech 2005; 12:11–6.
69. Buhse L. Measuring and managing method variability, presentation to the Food and Drug Administration Advisory Committee for Pharmaceutical Science: Manufacturing Science Subcommittee, October 25, 2005
70. Dressman J, Krämer J, eds. Pharmaceutical Dissolution Testing. Boca Raton, FL: Taylor and Francis, 2005.
71. Hanson R, Gray V, eds. Handbook of Dissolution Testing, 3rd ed. Hockessin, DE: Dissolution Tech 2004.
72. Palmieri, A, ed. Dissolution Theory, Methodology, and Testing. Hockessin, DE: Dissolution Tech 2007.
73. Verified June 9, 2007, www.dissolution.com; Dissolution Discussion Group.
74. Verified June 9, 2007, www.dissolutiontech.com; Dissolution Technologies
75. Baxter JL, Kukura J, Muzzio FJ. Hydrodynamics-induced variability in the USP apparatus II dissolution test. Int J Pharm 2005; 292:17–28.
76. Qureshi SA. A new crescent-shaped spindle for drug dissolution testing-but why a new spindle? Dissolution Tech 2004; 11:13–8.
77. Gray V. Challenges to the dissolution test, including equipment calibration. Analytical Methods, A Technology Primer, a Supplement to Pharmaceutical Technology, 2006, 4–13.

6
Setting Dissolution Specifications

Patrick J. Marroum
Office of Clinical Pharmacology, Center for Drug Evaluation and Research, U.S. Food and Drug Administration,[*] Silver Spring, Maryland, U.S.A.

INTRODUCTION

The release of the drug substance from the solid dosage form has a major impact on how fast a drug will be absorbed. In certain instances, as is the case with modified release formulations the rate limiting step in the appearance of the drug in the systemic circulation is its release from the formulation. Due to the critical role that dissolution plays in the bioavailability of the drug, in vitro dissolution can serve as a relevant predictor of the in vivo performance of the drug product.

In the vast majority of cases, in vitro dissolution of an immediate release product is one of the most important tools in assuring the batch to batch quality of the drug product. Establishing the appropriate dissolution specifications will assure that the manufacture of the dosage form is consistent and successful through out the life cycle of the product and that each dosage unit within a batch will have the same pharmaceutical qualities that correspond to those that have shown to have an adequate safety and efficacy profile. In the case where dissolution is predictive of the in vivo performance, clinically meaningful dissolution specifications will minimize the variability to the patient and therefore will optimize drug therapy.

In this chapter, an overview of the relevant regulatory guidance on how to set dissolution specifications for IR formulations, MR formulations with or without an in vitro in vivo correlation (IVIVC) will be given. Examples on how to use an IVIVC to set clinically relevant dissolution specifications will be discussed. In addition the issues peculiar to specialized dosage forms such as implants and Drug Eluting stents will be summarized with some recommendations on how to overcome the uniqueness of these dosage forms.

GENERAL PRINCIPLES IN SETTING DISSOLUTION SPECIFICATIONS

Until recently, the dissolution test was considered to be a purely quality control tool to assure consistency from batch to batch. However, with the ability to develop relationship between the in vitro dissolution of a drug product and its in vivo bioavailability, the

[*]The views expressed in this chapter are those of the author. No official support or endorsement by the Food and Drug Administration is intended or should be inferred.

dissolution test became a surrogate for the in vivo performance of the drug product and is used more and more to address the impact of changes in chemistry and manufacturing controls (1,2). Not only that, products can be approved only on the comparability of their dissolution profiles without having to conduct in vivo studies (3). Therefore with the choice of the most appropriate dissolution specifications, one can optimize the therapeutic benefit to the patient by decreasing the variability from one lot to the other.

SHOULD VARIABILITY BE AN IMPORTANT CONSIDERATION IN SETTING DISSOLUTIONS SPECIFICATION?

In the past it was usual and customary to set dissolution specifications based on the variability in the in vitro dissolution data. The end result of such a practice was the possibility of introducing lots on the market that are highly variable resulting in potentially wide fluctuations in plasma levels leading to a variable therapeutic effect and increased incidence of adverse events. Moreover, this practice of setting the limits to ±3 standard deviations tended to reward manufacturers with poor and highly variables formulations. Therefore manufacturers with poorer manufacturing and process controls will have products with relatively wider dissolution specifications compared to manufacturers with very tight controls in their manufacturing. To remedy this, the FDA is no longer accepting such a practice and it now stipulates that variability should no longer be a consideration in setting dissolution specifications. This change in policy would force drug manufacturers to tighten their manufacturing controls and to develop less variable dissolution methods.

USP ACCEPTANCE CRITERIA

The United States Pharmacopea (USP) sets acceptance criteria for the dissolution characteristics. In general the acceptance criteria are composed of 3 levels. Level 1 consists of testing 6 units with the acceptance criteria based on the performance of the individual units. Levels 2 consists of testing 12 units while level 3 tests 24 units. Both levels 2 and 3 use an acceptance criteria based on average performance with limits on the individual units performance.

Table 1 summarizes the USP acceptance table for immediate release dosage forms (4). Table 2 summarizes the USP acceptance criteria for modified release formulation including transdermal delivery systems. Tables 3 and 4 summarize the USP acceptance criteria for the various stages of dissolution testing for delayed release formulations for the acid and buffer phases, respectively.

TABLE 1 USP Acceptance Criteria for Immediate Release Dosage Forms

Stage	Number tested	Acceptance criteria
S_1	6	Each unit is not less than Q + 5%
S_2	6	Average of 12 units ($S_1 + S_2$) is Equal or greater than Q and no unit is less than Q−15%
S_3	12	Average of 24 units ($S_1 + S_2 + S_3$) is equal or greater than Q, not more than 2 units are less than Q−15% and no unit is less than Q−25%

Q is defined as the target % of labeled claim to be dissolved at the specified time point.

TABLE 2 USP Acceptance Criteria for Modified Release Formulations

Level	Number tested	Criteria
L_1	6	No individual value lies outside each of the stated ranges and no individual value is less than the stated amount at the final test time
L_2	6	The average value of the 12 units ($L_1 + L_2$) lies within each of the stated ranges and is not less than the stated amount at the final test time, none is more than 10% of labeled content outside each of the stated ranges and none is more than 10% of labeled content below the stated amount at the final test time
L_3	12	The average value of the 24 units ($L_1 + L_2 + L_3$) lies within each of the stated ranges, not more than 2 of the 24 units are more than 10% of labeled content outside the stated ranges, not more than 2 of the 24 units are more than 10% of labeled content below the stated amount at the final test time, and none of the units is more than 20% labeled content outside the stated ranges, not more than 2 of the 24 units are more than 20% of labeled content below the stated amount at the final test time

TABLE 3 USP Acceptance Criteria for the Acid Phase of Testing for Delayed Release Formulations

Level	Number tested	Criteria
A_1	6	No individual value exceeds 10% dissolved
A_2	6	Average of 12 units ($A_1 + A_2$) is not more than 10% dissolved and no individual unit is greater than 25% dissolved
A_3	12	Average of 24 units ($A_1 + A_2 + A_3$) is not more than 10% dissolved, and no individual unit is greater than 25%

TABLE 4 USP Acceptance Criteria for the Buffer Phase of Testing for Delayed Release Formulations

Level	Number tested	Criteria
B_1	6	Each unit is not less than Q + 5%
B_2	6	Average of 12 units ($B_1 + B_2$) is equal or greater than Q and no unit is less than Q−15%
B_3	12	Average of 24 units ($B_1 + B_2 + B_3$) is equal or greater than Q, not more than 2 units are less than Q−15% and no unit is less than Q−25%

Individual versus Mean Performance

It has been a common practice to propose dissolution specifications based on the ability to pass the specifications at stage 1 of the USP acceptance criteria (all the individual units meet the specifications). This practice would result in having some outlier units drive the specifications. If one accepts the premise that all the units should be able to meet the acceptance criteria, one would result with dissolution specifications that would allow the release of lots with markedly different release characteristics. Such specifications

would not ensure consistency from lot to lot and would not provide the best product to the patient. It is a misconception to believe that if a lot fails to meet the dissolution specification at the stage 1 of USP testing, this signifies that the manufacturing process is not well controlled. In fact from a regulatory point of view, a failure exists when the lot fails to meet the acceptance criteria at stage 3 of testing. In view of the above consideration, setting the dissolution specifications based on average performance (ability to pass stage 2 testing) would result in acceptance criteria that would minimize the probability of the release of lots with atypical performance and therefore ensuring a more consistent therapeutic effect to the patient.

THE CHOICE OF AMOUNT OF DRUG DISSOLVED (Q) FOR IR PRODUCTS

The specification for the amount of drug dissolved is another important consideration in ensuring that the patient always gets the same therapeutic dose from lot to lot. For drugs that exhibit complete dissolution, setting the highest Q value possible would minimize the variability in the dose delivered to the subject. While in an ideal situation, one would like to see a Q value of 100%, from a practical point of view this is not possible due to fact that there is inherent variability both in the content uniformity of the dosage form and in the dissolution test. If one surveys the monographs of older drugs in the USP (2), it can be observed that seldom a Q value of greater than 75% is observed for completely dissolving drugs. However, in recent years, it is more common to see the Q value set at 80% with some cases going up to 85%. Such a specification would not allow the release of lots that on average differ by more than 20% in the amount of drug delivered and thus minimizing the probability of bioinequivalence.

DISSOLUTION TIME SPECIFICATIONS

While the choice of time points is clearly defined for modified release formulation in the 1997 IVIVC guidance, there is much less agreement on the optimal time point for IR formulations. However, for very fasting dissolving products there is considerable debate on how fast the time specification should be. Most sponsors opt not to set specifications faster than 30 minutes even though their product might be completely dissolving in 5 or 10 minutes. It is believed that to set a faster dissolution time specification would not translate into in vivo bioavailability differences. Therefore accordingly, dissolution time points faster than 30 minutes will put an undue manufacturing burden without achieving any benefit. However, at present it is not uncommon that both sponsors and regulators consider dissolution time point specifications as early as 15 minutes for fast dissolving formulations (100% in less than 10 minutes). Such early time points will minimize the introduction of lots with markedly different dissolution characteristics and will ensure a more consistent performance from lot to lot.

SHOULD ALL LOTS MEETING THE DISSOLUTION LIMITS BE BIOEQUIVALENT?

In an ideal situation, one would like to see that all lots allowed to be released by the specifications be bioequivalent. This is not always possible because in certain cases this

will constitute a heavy burden on the manufacturer and one would end up rejecting a large proportion of perfectly acceptable batches. That is why the IVIVC guidance stipulates that at the minimum lots that are on the upper and lower specification limit be bioequivalent to the clinical bio/lot which were used in the clinical trials and whose safety and efficacy has been established (5). This position is deemed not acceptable by some because they believe that all batches found in the market should be bioequivalent. This is somewhat more stringent than the current practice especially for wide therapeutic index drugs. As an example let's take two formulations that are bioequivalent to a clinical formulation but differing in their mean performance by 10% on the upper and lower side of the clinical formulation. These two formulations most probably will not be bioequivalent to each other (since they are 20% different on average and thus would not be able to pass the regulatory requirement of a 90% confidence interval of 80–125%) but will still be acceptable from a safety and efficacy profile point of view due to the fact that a 20% difference in plasma concentrations will not result in any clinical difference in the pharmacological action of the drug product. Therefore for wide therapeutic index drugs, the minimal requirement that these lots be bioequivalent to the clinical/bio lots will provide regulatory relief for manufacturers without introducing into the market lots having inadequate safety and efficacy profiles. However, for drugs exhibiting a narrow therapeutic index, the criteria should be more stringent and should require that all the lots within the dissolution specifications be bioequivalent to each other. It is the opinion of the author that criteria for dissolution specification that take into account the clinical pharmacology characteristics of the drug are more appropriate than criteria that are based solely on the ability to meet a statistical criterion on the plasma concentrations.

FDA GUIDANCE ON DISSOLUTION TESTING OF IMMEDIATE RELEASE ORAL DOSAGE FORMS

In August 1997, the US FDA released guidance on dissolution testing for IR oral dosage forms. This guidance was intended to provide: (a) general recommendations for dissolution testing, (b) approaches for setting dissolution specifications related to the biopharmaceutic characteristics of the drug substance, (c) statistical methods for profile comparisons and a process to determine whether dissolution testing is sufficient to grant a waiver for an in vivo bioequivalence study (6).

RECOMMENDATIONS ON SETTING DISSOLUTION SPECIFICATIONS

According to this guidance, for New Drug Applications, the dissolution specifications should be based on acceptable clinical, pivotal bioavailability, and/or bioequivalence batches. For generic drug applications (ANDAs) the dissolution specifications should be based on the performance of acceptable bioequivalence batches of the drug product. The NDA dissolution specifications should be based on experience gained during the drug development process and the in vitro performance of appropriate test batches. In the case of a generic drug product, the dissolution specifications are generally the same as the reference listed drug (RLD). The specifications are confirmed by testing the dissolution performance of the generic drug product from an acceptable bioequivalence study.

If the dissolution of the generic product is substantially different compared to that of the reference listed drug and the in vivo data remain acceptable, a different dissolution

specification for the generic product may be set. Once a dissolution specification is set, the drug product should comply with that specification throughout its shelf life.

The International Conference on Harmonization (ICH) Q1A guideline (Stability Testing of New Drug Substances and Drug Products) (7) has recommended that for an NDA, three batches (two pilot and one smaller scale) be placed into stability testing. These batches also may be used to set dissolution specifications when a suitable bioequivalence relationship exists between these batches and both the pivotal clinical trial batch and the drug product intended for the market.

Approaches for Setting Dissolution Specifications for a New Chemical Entity

The dissolution characteristics of the drug product should be developed based on consideration of the pH solubility profile and p*K*a of the drug substance. The drug permeability or octanol/water partition coefficient measurement may be useful in selecting the dissolution methodology and specifications. For NDAs, the specifications should be based on the dissolution characteristics of batches used in pivotal clinical trials and/or in confirmatory bioavailability studies. If the formulation intended for marketing differs significantly from the drug product used in pivotal clinical trials, dissolution and bioequivalence testing between the two formulations are recommended.

Dissolution testing should be carried out under mild test conditions, basket method at 50/100 rpm or paddle method at 50/75 rpm, at 15-minute intervals, to generate a dissolution profile. For rapidly dissolving products, generation of an adequate profile sampling at 5- or 10-minute intervals may be necessary. For highly soluble and rapidly dissolving drug products (BCS classes 1 and 3) (8), a single-point dissolution test specification of NLT 85% (Q= 80%) in 30 minutes or less is sufficient as a routine quality control test for batch-to-batch uniformity. For slowly dissolving or poorly water soluble drugs (BCS class 2), a two-point dissolution specification, one at 15 minutes to include a dissolution range (a dissolution window) and the other at a later point (30, 45, or 60 minutes) to ensure 85% dissolution, is recommended to characterize the quality of the product. The product is expected to comply with dissolution specifications throughout its shelf life. If the dissolution characteristics of the drug product change with time, whether or not the specifications should be altered will depend on demonstrating bioequivalence of the changed product to the original biobatch or pivotal batch. To ensure continuous batch-to-batch equivalence of the product after scale-up and postapproval changes in the marketplace, dissolution profiles should remain comparable to those of the approved biobatch or pivotal clinical trial batch(es).

Approaches for Setting Dissolution Specifications for Generic Products

The approaches for setting dissolution specifications for generic products fall into three categories, depending on whether an official compendial test for the drug product exists and on the nature of the dissolution test employed for the reference listed drug. All approved new drug products should meet current USP dissolution test requirements, if they exist. The three categories are:

1. *USP drug product dissolution test available*: In this instance, the quality control dissolution test is the test described in the USP. The Division of Bioequivalence, Office of Generic Drugs, also recommends taking a dissolution profile at 15-minute intervals or less using the USP method for test and reference products (12 units each). The Division

of Bioequivalence may also recommend submitting additional dissolution data when scientifically justified. Examples of this include (*i*) cases in which USP does not specify a dissolution test for all active drug substances of a combination product and (*ii*) cases in which USP specifies use of disintegration apparatus.

2. *USP drug product dissolution test not available; dissolution test for reference listed NDA drug product publicly available*: In this instance, a dissolution profile at 15-minute intervals of test and reference products (12 units each) using the method approved for the reference listed product is recommended. The Division of Bioequivalence may also request submission of additional dissolution testing data as a condition of approval, when scientifically justified.

3. *USP drug product dissolution test not available; dissolution test for reference listed NDA drug product not publicly available*: In this instance, comparative dissolution testing using test and reference products under a variety of test conditions is recommended. The test conditions may include different dissolution media (pH 1–6.8), addition of surfactant, and use of apparatus 1 and 2 with varying agitation. In all cases, profiles should be generated as previously recommended. The dissolution specifications are set based on the available bioequivalence and other data.

Special Cases

Two-Point Dissolution Test

For poorly water soluble drug products (e.g., carbamazapine), dissolution testing at more than one time point for routine quality control is recommended to ensure in vivo product performance. Alternatively, a dissolution profile may be used for purposes of quality control.

Two-Tiered Dissolution Test

To more accurately reflect the physiologic conditions of the gastrointestinal tract, two-tiered dissolution testing in simulated gastric fluid (SGF) with and without pepsin or simulated intestinal fluid (SIF) with and without pancreatin may be employed to assess batch-to-batch product quality provided the bioequivalence is maintained. Recent examples involving soft and hard gelatin capsules show a decrease in the dissolution profile over time either in SGF or in SIF without enzymes. This has been attributed to pellicle formation. When the dissolution of aged or slower releasing capsules was carried out in the presence of an enzyme (pepsin in SGF or pancreatin in SIF), a significant increase in the dissolution was observed. In this setting, multiple dissolution media may be necessary to adequately assess product quality.

Mapping or Response Surface Methodology

Mapping is defined as a process for determining the relationship between critical manufacturing variables (CMV) and a response surface derived from an in vitro dissolution profile and an in vivo bioavailability data set. The CMV include changes in the formulation, process, equipment, materials, and methods for the drug product that can significantly affect in vitro dissolution. The goal is to develop product specifications that will ensure bioequivalence of future batches prepared within the limits of acceptable dissolution specifications. Several experimental designs are available to study the influence of CMV on product performance. One approach to study and evaluate the mapping process includes: (*i*) prepare two or more dosage formulations using CMV to

study their in vitro dissolution characteristics; (*ii*) test the products with fastest and slowest dissolution characteristics along with the standard or the *to be marketed dosage form* in small groups (e.g., $n > 12$) of human subjects; and (*iii*) determine the bioavailability of the products and in vitro–in vivo relationship. The products with extreme dissolution characteristics are also referred to as side batches. If the products with the extreme range of dissolution characteristics are found to be bioequivalent to the standard or the to be marketed dosage form, future batches with dissolution characteristics between these ranges should be equivalent to one another. This approach can be viewed as verifying the limits of the dissolution specifications. Product dissolution specifications established using a mapping approach will provide maximum likelihood of ensuring stable quality and product performance. Depending on the number of products evaluated, the mapping study can provide information on in vitro–in vivo correlations and/or a rank order relationship between in vivo and in vitro data.

Validation and Verification of Specifications

Confirmation by in vivo studies may be needed for validation of an in vitro system. In this situation, the same formulation should be used but nonformulation CMV should be varied. Two batches with different in vitro profiles should be prepared (mapping approach). These products should then be tested in vivo. If the two products show different in vivo characteristics, then the system is validated. In contrast, if there is no difference in the in vivo performance, the results can be interpreted as verifying the dissolution specification limits as discussed under mapping. Thus, either validation or verification of dissolution specifications should be confirmed.

SETTING DISSOLUTION SPECIFICATIONS FOR MODIFIED RELEASE FORMULATIONS

In vitro dissolution specifications should generally be based on the performance of the clinical/bioavailability lots. These specifications may sometimes be widened so that scale-up lots, as well as stability lots, meet the specifications associated with the clinical/bioavailability lots. This approach is based on the use of the in vitro dissolution test as a quality control test without any in vivo significance, even though in certain cases (e.g., ER formulations), the rate limiting step in the absorption of the drug is the dissolution of the drug from the formulation. An IVIVC adds in vivo relevance to in vitro dissolution specifications, beyond batch-to-batch quality control. In this approach, the in vitro dissolution test becomes a meaningful predictor of in vivo performance of the formulation, and dissolution specifications may be used to minimize the possibility of releasing lots that would be different in in vivo performance (9). The IVIVC guidance for modified release formulations makes several recommendations on how to set the most desirable dissolution specifications in the presence and absence of an IVIVC: these can be summarized below.

SETTING DISSOLUTION SPECIFICATIONS WITHOUT AN IVIVC

For drug products without an established predictive IVIVC the following points should be taken when setting the dissolution specifications:

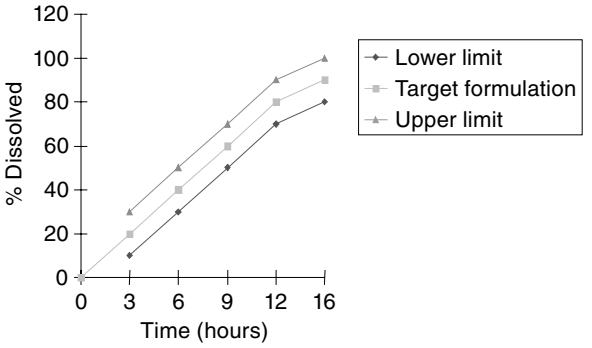


FIGURE 1 Dissolution specifications in the absence of an IVIVC.

(A) The recommended range at any dissolution time point specification is ±10% deviation from the mean dissolution profile obtained from the clinical/bioavailability lots as illustrated in Figure 1. In certain cases, reasonable deviations from the ± 10% range can be accepted provided that the range at any time point does not exceed 25%. Specifications greater than 25% may be acceptable based on evidence that lots (side batches) with mean dissolution profiles that are allowed by the upper and lower limit of the specifications are bioequivalent. Specifications should be established on clinical/bioavailability lots. Widening specifications based on scale-up, stability, or other lots for which bioavailability data are unavailable is not recommended.

(B) A minimum of three time points is recommended to set the specifications. These time points should cover the early, middle, and late stages of the dissolution profile. The last time point should be the time point where at least 80% of drug has dissolved. If the maximum amount dissolved is less than 80%, the last time point should be the time when the plateau of the dissolution profile has been reached. Specifications should be established based on average dissolution data for each lot under study, equivalent to USP stage 2 testing. Specifications that allow all lots to pass at stage 1 of testing may result in lots with less than optimal in vivo performance passing these specifications at USP stage 2 or stage 3. The USP acceptance criteria for dissolution testing are recommended unless alternate acceptance criteria are specified in the ANDA/NDA.

SETTING DISSOLUTION SPECIFICATIONS WHERE AN IVIVC HAS BEEN ESTABLISHED

Optimally, specifications should be established such that all lots that have dissolution profiles within the upper and lower limits of the specifications are bioequivalent. Less optimally but still possible, lots exhibiting dissolution profiles at the upper and lower dissolution limits should be bioequivalent to the clinical/bioavailability lots or to an appropriate reference standard.

Level A Correlation Established

As for the case without the presence of an IVIVC, the specifications should be established based on average data. A minimum of three time points is recommended to establish the specifications. These time points should cover the early, middle and late stages of the dissolution profile. The last time point should be the time point where at least 80% of

drug has dissolved. If the maximum amount dissolved is less than 80%, then the last time point should be the time where the plateau of the dissolution profile has been reached.

However, the dissolution specifications range in this case is no longer determined based on the in vitro performance but on predicted in vivo plasma concentration time profiles. The IVIVC is used to determine the difference in plasma concentration time profiles corresponding to the extreme dissolution profiles that are allowed by the upper and lower limits of the dissolution specifications (as shown in Fig. 2). This is accomplished by calculating the plasma concentration time profile using convolution or other appropriate modeling techniques and determining whether the lots with the fastest and slowest release rates that are allowed by the dissolution specifications result in a maximal difference of 20% in the predicted AUC and C_{max}. An established IVIVC may allow setting wider dissolution specifications. This would be dependent on the predictions of the IVIVC (i.e., 20% differences in the predicted C_{max} and AUC). USP acceptance criteria for dissolution testing are recommended unless alternate acceptance criteria are specified in the ANDA/NDA.

For wide therapeutic window drugs, a specification range narrower than ±10% of the % labeled claim would not be recommended even in the event that such a specification would result in more than 20% difference in the mean predicted AUC and C_{max}. Since the default range without the presence of an IVIVC is 20% sponsors that developed an IVIVC should not be penalized with narrower dissolution specifications specially when such narrower ranges do not provide any therapeutic advantage to the patient but will impose an undue burden from a manufacturing point of view on the sponsor.

Multiple Level C Correlation Established

If a multiple point Level C correlation has been established, establish the specifications at each time point such that there is a maximal difference of 20% in the predicted mean C_{max} and AUC. Additionally, the last time point should be the time point where at least 80% of drug has dissolved.

Level C Correlation Based on Single Time Point Established

This one time point may be used to establish the specification such that there is not more than a 20% difference in the predicted AUC and C_{max}. At other time points, the maximum recommended range at any dissolution time point specification should be ±10% of label claim deviation from the mean dissolution profile obtained from the clinical/

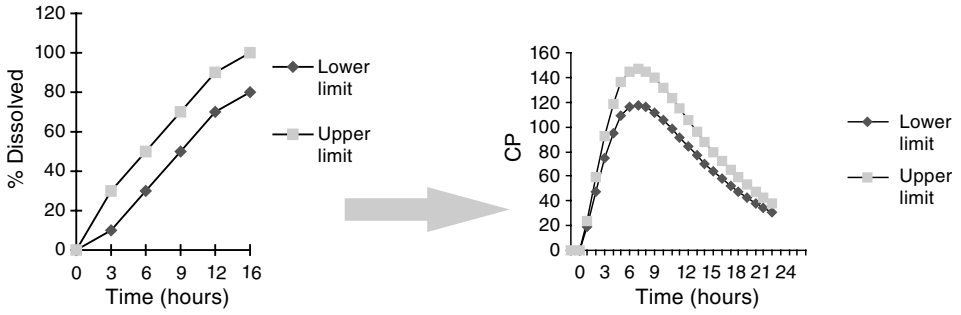


FIGURE 2 Dissolution specifications in the presence of an IVIVC.

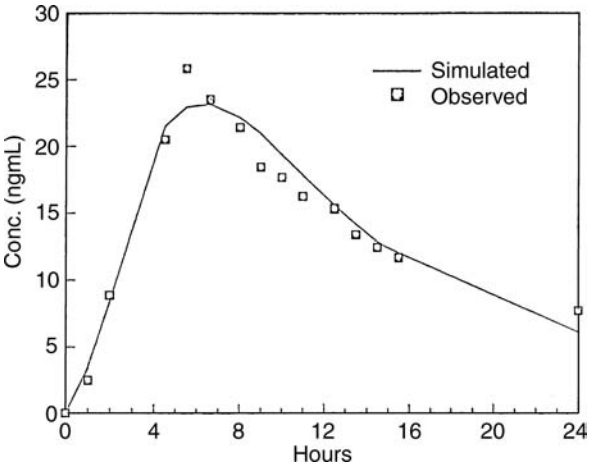


FIGURE 3 Influence of the release rate specifications on plasma levels: Inequivalent plasma profiles.

bioavailability lots. Reasonable deviations from ±10% may be acceptable if the range at any time point does not exceed 25%.

Example on How to Use an IVIVC to Set the Dissolution Specifications

The IVIVC for this modified release drug product was developed using a convolution approach. The sponsor used dissolution as an input function to predict the observed plasma concentrations. The dissolution profiles were fitted to the Weibull function which was used as the input function to predict the plasma concentration time profiles corresponding to the respective dissolution profiles. It is to be noted that any other mathematical function that could describe adequately the dissolution profiles could have been used as an input function. In Figure 3 the straight line describes the predicted plasma profiles and the dotted points are the observed concentrations. This IVIVC was deemed predictive and therefore useful from a regulatory point of view. Figure 4 shows the ranges of the dissolution profiles that correspond to the chosen dissolution limits as well as lots that are bioequivalent. The

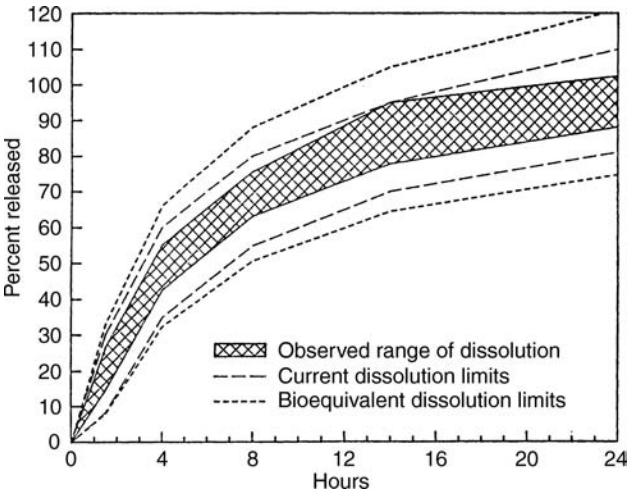


FIGURE 4 Influence of the release rate specifications on plasma levels: equivalent plasma profiles.

dashed lines denote the dissolution limits proposed by the sponsor. The shaded area denotes the dissolution ranges for all the lots that were tested in the NDA. The very upper and lower lines (the dotted lines) denote the limits of dissolution profiles for lots that are predicted to be bioequivalent (12). This is a very good example on how to optimally set the dissolution specifications using all the available data in hand. With the use of modeling techniques, and the presence of a predictive IVIVC, the sponsor was able to set clinically meaningful dissolution specifications in such a way that all the lots within the dissolution specifications are bioequivalent to each other. The end result will be a more consistent therapeutic effect due to decreased variability in the plasma levels.

SETTING SPECIFICATIONS BASED ON RELEASE RATE

If the release characteristics of the formulation can be described by a zero-order process for some period of time (e.g., 5%/hr from 4 to 12 hours), and the dissolution profile appears to fit a linear function for that period of time, a release rate specification may be established to describe the dissolution characteristics of that formulation. A release rate specification may be an addition to the specifications established on the cumulative amount dissolved at the selected time points. Alternatively, a release rate specification may be the only specification except for the specification for time when at least 80% of drug has dissolved.

The FDA guidance introduced this novel approach in setting dissolution specifications for formulations exhibiting a zero order release characteristic. An example of such a formulation is the osmotic delivery system commonly referred to as Gastro intestinal therapeutic systems (GITS). If these formulations are designed to deliver the drug at a constant rate that can be described by a linear relationship over a certain period of time, then one can set a release rate specification to describe the performance of the formulation. This release rate specification can be in addition to or instead of the cumulative dissolution specifications that one usually sets for a modified release product.

A release rate specification will provide for a better control of the in vivo performance of the drug because it is the release characteristics of the formulation that will determine the rate of appearance of the drug in the systemic circulation. This can be described more appropriately by the release rate compared to the cumulative amounts of drug dissolved at a certain interval of time. As an illustration of this point, let's consider the dissolution profiles of two lots of the same formulation (shown in Fig. 5) with similar release rates but are on the upper and lower limits of the cumulative dissolution specifications. Assuming a level A correlation for this product, the predicted plasma concentration time profile corresponding to these two lots are similar, differing only in the time to achieve peak plasma concentration. On the other hand if one examines the case presented in Figure 6 whereby the two lots are very close in their cumulative dissolution profiles (both at the upper limit of the dissolution specifications) but markedly different in their release rates, one can clearly see that the predicted plasma profiles corresponding to these lots are very different and considered not to be bioequivalent (13).

SPECIALIZED DOSAGE FORMS

Specialized dosage forms such as vaginal rings, intra uterine devices and implants present a unique challenge in terms of dissolution testing. These dosage forms are designed to release very small amounts of the drug over extended period of time (days, months, and

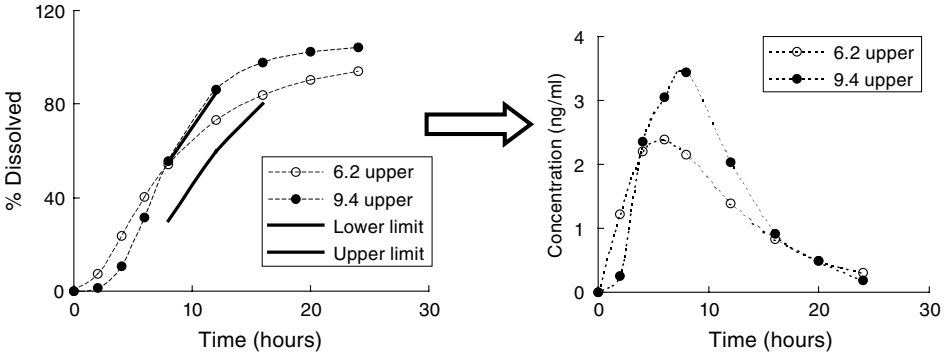


FIGURE 5 Plasma profile observed and predicted from dissolution.

years). Setting dissolution specifications in terms of the cumulative amount of drug released over time might neither be practical nor would it provide the most meaningful way in controlling the quality of the product. Since with these formulations, the rate limiting step for the appearance of the drug into the site of action is the release of the drug from the formulation, it is therefore beneficial to find the dissolution conditions that mimic the release rate in vivo. Once these conditions are established, the dissolution specifications should be based on the observed release rate (in terms of amount of drug or % released versus time). The upper and lower limits should be chosen as per the recommendation given for modified release products in the IVIVC guidance and should not result in more than 20% difference in the predicted PK parameters of interest. Such an approach would not only allow setting specifications with predictable in vivo outcomes but will also alleviate the testing burden in that the release rate specification could be estimated at various time intervals throughout the intended dosing interval.

DRUG ELUTING STENTS

With the recent advances in medical technology, it is more common to see the therapeutic effect of a device be optimized by its combination with a drug. A prime example of such a device is the drug eluting stent. Since these stents are implanted, having consistent

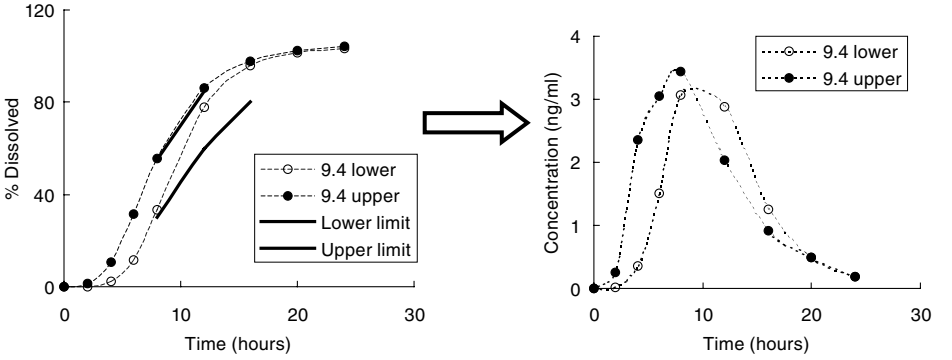


FIGURE 6 Dissolution limits.

elution characteristics throughout the intended duration of action is crucial in maintaining the therapeutic benefit to the patient. Due to the extreme difficulty in estimating the in vivo elution characteristics for such devices setting elution specifications that will be relevant from an in vivo point of view becomes very challenging.

In the case where the measurable plasma levels are indicative of the in vivo elution of the drug from the stent at the site of action and the in vitro conditions result in in vitro elution rates mimicking those observed in vivo, the dissolution specifications should be set in terms of the observed in vitro elution rate.

However, in the situation where the plasma levels are too low to measure, it becomes practically impossible to determine the elution characteristics. In such a case, animal models could be used to determine the elution characteristics of the drug eluting stents (DESs). At different time intervals, the stents could be explanted and the amount of drug remaining on the stent as well as the amount found in the adjacent tissues could be measured. This information can be a valuable guide for the development of the most relevant elution method with the most relevant specifications. In other situations, with the current advances in x-ray computer technologies, it may be possible to non-invasively monitor the local drug release from the DES. Such a capability will go a long way in characterizing the elution behavior in the target population. This will in turn enable one to select the elution method and specifications with the in vivo considerations in mind (14,15).

Another important consideration in setting the elution specifications is the clinical performance of the DES. If the clinical trials showed that there is a correlation between the safety and efficacy profile and elution rates, the specifications should be set in such a way that only DES with elution rates with acceptable safety and efficacy profiles be released to the market. At a minimum, the elution specifications should not release any lots with elution characteristics beyond what was found to be acceptable from a clinical point of view.

CONCLUSION

Dissolution can play a major role in assuring the quality of a drug product. For this reason, the setting of optimal dissolution specifications can minimize the variability to the patient by providing less variable release characteristics. This will lead to more consistent plasma concentrations resulting in a more consistent therapeutic effect. IVIVCs can be a powerful tool in setting clinically meaningful dissolution specifications. The ability to predict plasma concentrations from in vitro dissolution profiles will allow the setting of dissolution specifications that would ensure that all lots released would be bioequivalent to the lots that were shown to be safe and effective thus minimizing the probability of releasing lots with unproven safety and efficacy profiles.

REFERENCES

1. Guidance for modified release solid oral dosage forms, scale up and post approval changes: chemistry and controls: in vitro dissolution testing and in vivo bioequivalence documentation. Center for Drug Evaluation and Research, Food and Drug Administration, July 1997.
2. Guidance for immediate release solid oral dosage forms, scale up and post approval changes: chemistry and controls: in vitro dissolution testing and in vivo bioequivalence documentation. Center for Drug Evaluation and Research, Food and Drug Administration, July 1997.

3. Guidance on BA and BE studies for orally administered drug products—general considerations, Center for Drug Evaluation and Research, Food and Drug Administration, March 2003.
4. Dissolution, US Pharmacopeia, 711, 30, 2007.
5. Guidance for industry, extended release solid oral dosage forms: development, evaluation and application of in vivo/in vitro correlations. Center for Drug Evaluation and Research, Food and Drug Administration, September 1997.
6. Guidance for industry, dissolution testing for immediate release solid oral dosage form. Center for Drug Evaluation and Research, Food and Drug Administration, August 1997.
7. International conference on harmonization guidance for industry Q1A(R2) stability testing of new drug substances and products. Center for Drug Evaluation and Research, Food and Drug Administration, November 2003.
8. Guidance for industry waiver of in vivo bioavailability and bioequivalence studies for immediate-release solid oral dosage forms based on a biopharmaceutics classification system. Center for Drug Evaluation and Research, Food and Drug Administration, August 2000.
9. Marroum PJ. Role of in vivo in vitro correlations in setting dissolution specifications. Am Pharm Rev 1999; 2:39–42.
10. Gillespie WR. Convolution—based approaches for in vivo in vitro correlation modeling, in vitro in vivo correlations. Adv Exp Med Biol 1997; 423:53–65.
11. Gillespie WR. Modeling strategies for in vivo in vitro correlations. In: Amidon G, Robinson JR, Williams RL, eds. Scientific Foundations for Regulating Drug Product Quality, Alexandria, VA: AAPS Press, 1997:275–92.
12. Marroum PJ. Regulatory examples: Dissolution specifications and bioequivalence product standards. In: Amidon V, Robinson JR, Williams RL eds. Scientific Foundations for Regulating Drug Product Quality, Alexandria, VA: AAPS Press, 1997: 305–19.
13. Marroum PJ. In vitro–in vivo correlation: A regulatory perspective with case studies. In: Chilikuri DM, Sunkara G, Young D, eds. Pharmaceutical Product Development In Vitro–In Vivo Correlation, New York, NY: Informa Healthcare, 2007: 177–95.
14. Szymanski-Exner, et al. Noninvasive monitoring of local drug release using x-ray computed tomography: Optimization and in-vitro/in-vivo valiation. J Pharm Sci 2003; 92:289.
15. Hwang, et al. Physiological transport forces govern drug distribution for stent-based delivery. Circulation 2001; 104:600.

7
Mechanical Strength of Tablets

Göran Alderborn and Göran Frenning
Department of Pharmacy, Uppsala University, Uppsala, Sweden

INTRODUCTION

In order to secure that a tablet, i.e., a porous specimen formed by confined compression by moving punches, is elegant and that the correct dose of the drug(s) is administered, a tablet must remain intact during handling between manufacturing and administration. Tablets must thus resist attrition and fracturing and possess a certain mechanical strength after formation. The mechanical strength is related to the micro-structure of the tablet, i.e., the size and the orientation of the particles and pores forming the tablet and the structure of the contacts formed between the particles that provides coherency. Other important properties of a tablet that also must be controlled by the formulation scientist, such as tablet disintegration and drug dissolution, will possibly also depend on the tablet micro-structure. Thus, formulation or process factors that will change the mechanical strength of a tablet will probably also have a parallel effect on other tablet properties. Relationships between the mechanical strength and other relevant pharmaceutical properties of a tablet may in many cases be complex and will not be discussed in this chapter. The inter-dependence between different properties of a pharmaceutical tablet should however be a concern to the reader of this chapter.

The scientific discipline dealing with fracturing of solids is referred to as fracture mechanics and is a part of solid mechanics. In addition to mechanical strength testing, several methods are today used in pharmaceutical research and formulation development as a means to assess fracture mechanics parameters of drugs and excipients (such as the critical stress intensity factor). The solid mechanics discipline deals also with the deformation of a solid body due to an externally applied force. Such deformations occur normally before the solid fracture and they are described by mechanical parameters, such as the modulus of elasticity and the yield stress. The measurements of fractures mechanics parameters and deformations are not scopes of this chapter. The terms used in describing the deformation of solid bodies will however be used in this chapter. The reader is referred to text books on solid mechanics (1,2) to clarify the meaning of these terms.

MECHANICAL STRENGTH TESTING

Pharmaceutical Applications of Strength Testing

The mechanical strength of a solid specimen is associated with the force or stress needed to crack, fracture or erode the specimen. The term mechanical strength is thus used in this

chapter as a collective term of different events that will crack, fracture, fragment, crush, or erode a tablet. In pharmaceutical literature, the term hardness is often misused as a term describing the fracture resistance. The hardness of a specimen is associated with its resistance against local permanent deformation and is measured predominantly by indentation. Thus, hardness is a parallel term to the yield strength of a solid and will show some proportionality to the yield strength (3).

From the requirement that a tablet must remain intact during handling between production and administration and thus must resist fracturing follows that measurements of mechanical strength are an important part of tablet formulation development, process up-scaling and tablet manufacturing. The determination of the mechanical strength of a tablet is carried out of several reasons during both development and manufacturing, such as:

- to aid in the selection of drug candidates and excipients during preformulation and formulation
- to detect batch variations of drugs and excipients in their compaction performance
- to assess the importance of formulation and production variables for the mechanical strength of the tablet
- to control the quality and quality consistency of tablets during production.

A tablet can be mechanically strained in numerous ways, such as by compression, bending and impaction, and the potential number of methods that could be used in mechanical strength testing is thus large. The results differ obviously between the methods and the design of the test method is related to one of three ambitions. Firstly, to mimic the complicated forces that will act on a tablet during processing or handling, such as impaction and attrition during tumbling. Secondly, to load the tablet in a simple and quick but yet reproducible way until fracture, i.e., a method suitable for use as a process control method during tablet manufacturing. Thirdly, to apply the force in such a way that the distribution of stresses evolved within the tablet can be described and approximated. Using the third approach, the fracture strength can be calculated from the stress needed to initiate a crack that grows and fractures the tablet. A method based on such a stress analysis enables the derivation of a measure of mechanical strength that is theoretically independent of the dimensions of the tablet. The most common mechanical strength value used in pharmaceutical scientific work in this context is the tensile strength.

Despite the number of potential test methods for assessing the resistance of a tablet towards fracturing or attrition, two methods dominate in pharmaceutical practice, i.e., the friability test and the fracture resistance test, and our discussion of tablet strength testing will thus focus on these two methods. The common use of these two methods is reflected by the fact that the tests are described in the current issues of the *European Pharmacopoeia* (EP) (4) and the *United States Pharmacopoeia* (USP) (5).

Friability

The term friability is associated with the response of a tablet subjected to impaction and sliding during shaking or tumbling and is thus an indication of the attrition resistance of a tablet. The idea behind attrition resistance methods is to mimic the kind of forces, caused by phenomena such as collisions and sliding of tablets towards each other, which a tablet is subjected to during handling between its manufacturing and its administration. The consequence of such mechanical straining of the tablet may be that single particles or particle clusters can be eroded from the tablet surface or the tablet may even fracture or fragment. For example, tablets without any visible defects can cap (i.e., split into two

pieces along the tablet main axes) during a friability test (6). The result of such phenomena will be a reduction in the tablet weight with a parallel change in the appearance of the tablet. A general definition of the term friability may thus be any change in physical characteristics of tablets that results in a reduction in the mass or in the formation of fragments of the tablet, occurring when the tablets are subjected to mechanical straining during handling. A friable tablet is a tablet which is prone to undergo such change in physical characteristics during handling. As a rule of thumb, a maximum weight loss of the tablets during a friability test of 1% is often applied (compare monographs in the USP and EP).

A multi-fold of methods with equal suitability may be used in the testing of the friability of tablets, such as shaking, gentle milling, tumbling, vibration, and fluidization. The most common experimental procedure to determine friability involves the rotation of tablets in a cylinder followed by the determination of the weight loss of the tablets. The most commonly used friability apparatus consists of a cylindrical drum of specified dimensions, equipped with a curved projection that will cause the tablets to fall along the drum diameter during rotation of the drum (Fig. 1). During testing, tablets will thus be subjected to forces due to rolling, sliding, collision etc. After tumbling for a specified number of rotations, the tablets are sieved, inspected and weighed. The weight loss is most commonly determined after a given number of rotations and this is the approach used in the USP and the EP. Alternatively, the weight loss can be followed over time (6,7) and one application of such a relationship is the assessment of a capping tendency of tablets. The rate of wear of tablets during mechanical straining has also been modeled based on a vibrating sieve method (8,9).

Fracture Resistance

The fracture resistance test involves the application of a force along a given direction of the tablet until the tablet fails, i.e., cracks, breaks or fragments. In pharmaceutical practice, the force is mostly applied by compression and in such a case, the tablet is placed against a platen and the force is applied along some axis of the tablet (i.e., the diameter in case of a cylindrical shaped tablet) by a movable platen or plunger (Fig. 2). The force is continuously increasing until the tablet fails and the force at failure is recorded.

During such compression, the tablet may fail in different ways, i.e., crack, fracture into two separate pieces of similar size or fragment into several differently sized pieces. The test is therefore referred to in pharmaceutical practice in different ways, such as fracture strength, breaking strength, crushing strength, and even hardness. The latter term is not advisable to use as discussed above. In the current issue of the EP, the test is referred to as resistance to crushing of tablets and in the USP, the term tablet crushing strength appears. A common type of failure that occurs during compression testing is a single fracture parallel to the compression load, giving two fragments of similar size. Such mode of failure is often referred to as a tensile failure (10,11). Other terms used to describe the mode of failure of the tablet during compression are double-cleft, triple-cleft and shear/compressive failure (12,13), indicating more complicated fracturing processes.

During testing, care must be taken to ensure that the test is conducted in a reproducible way. This involves a consistent orientation of the tablet by considering the shape of the tablet and break-marks and inscriptions. The force should be applied in a consistent way regarding the rate of movement of the movable platen since also this variable may affect the force at fracture (14).

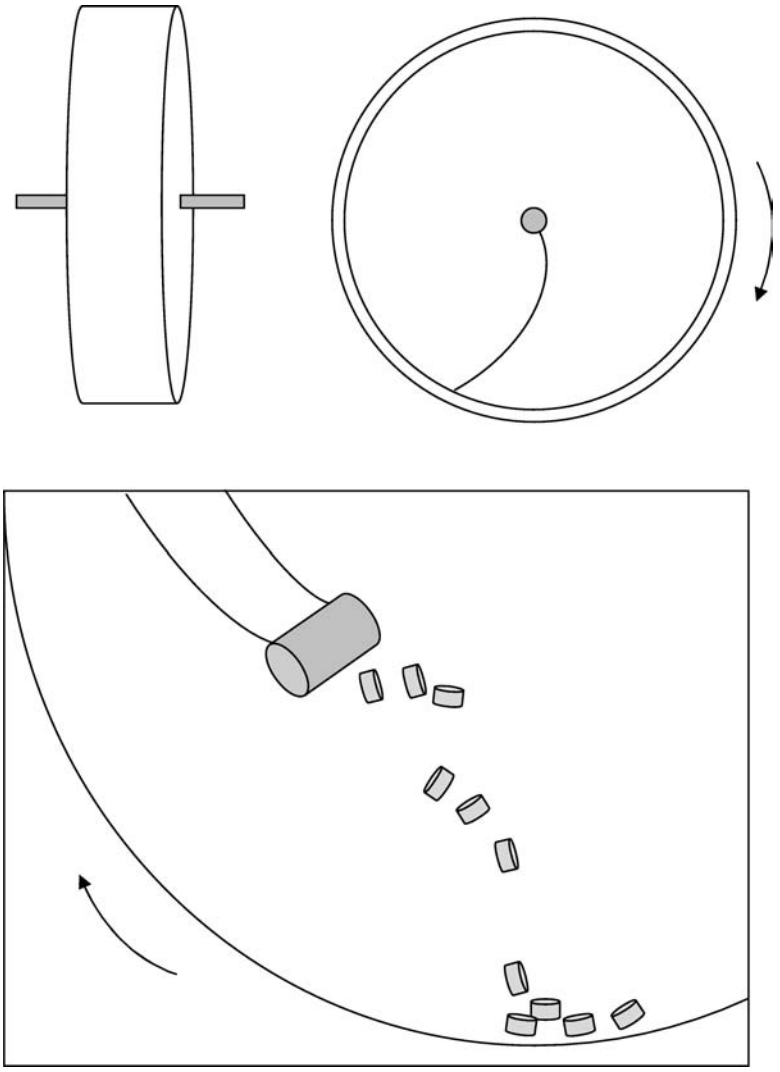


FIGURE 1 Schematic illustration of the most common type of friability apparatus, showing the drum and the curved projection and a close-up illustrating tablets falling from the curved projection.

Due to the simplicity and reproducibility of the test, the method has a broad use during formulation and manufacturing of tablets. Many commercial testers exist thus today and in a recent paper (15), a series of such testers are compared. Different units are in use to indicate the load that causes the tablet to fracture, such as Newton, kilogram (kg), and kilopound (kp). In research papers, the force in Newton is the dominant unit while in formulation development and in production alternative units may also be used. However, the current version of the EP states that the force at fracture should be expressed in Newton. The units kg and kp are units of mass and can thus be converted into Newton. An early instrument for measurement of fracture resistance of a tablet was the Strong-Cobb tester which indicated the load at fracture in Strong-Cobb units, a unit that still may be in use.

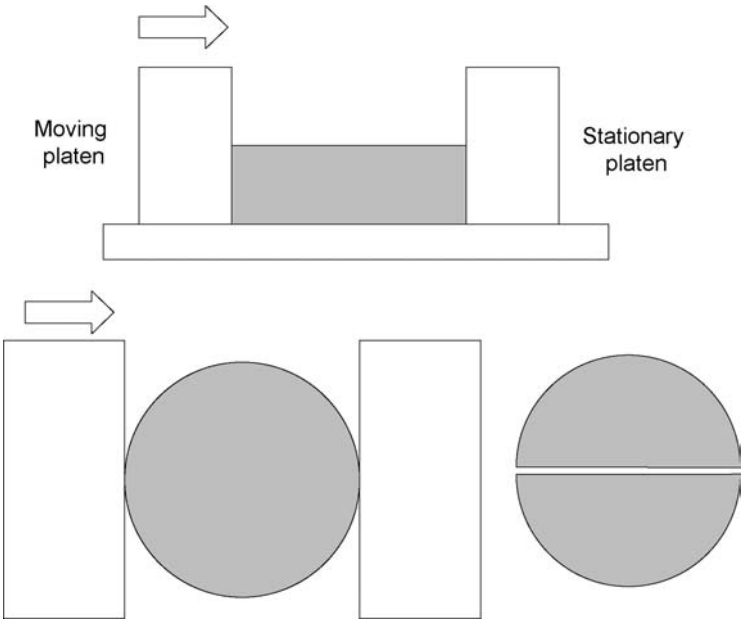


FIGURE 2 Schematic illustration of the diametrical compression test of a cylindrical flat-faced tablet. The illustration shows the side view and the upper view during loading of tablet and a top-view of a tensile failure of the tablet.

Pharmaceutical tablets can generally be described as brittle solids, i.e., the fracture is preceded by a limited deformation of the tablet, predominantly elastic deformation. However, the fact that tablets deform, both elastically and plastically, before fracture has caused an interest in studying also the force–displacement relationship during mechanical strength testing. One application is the calculation of the work of failure, also referred to as toughness (16), as a measure of the mechanical response of a tablet. The use of toughness measurements in formulation development seems today however limited.

Tensile Strength

Tensile Strength by Diametral Compression

The force needed to fracture a tablet is dependent on the dimensions of the tablet. By determining the tensile strength of a tablet, a comparison between tablets of different sizes or even shapes can be done. The most common tensile strength test is based on the diametral compression test discussed above.

The tensile strength test is normally used for plane-faced tablets, i.e., small cylinders. The calculation of a tensile strength is based on the assumption that the tablet fails by a single linear fracture across the diameter of the cylinder, i.e., a normal tensile failure (Fig. 2). The equation was introduced in pharmaceutical practice by Fell and Newton (11) but due to its original development, the procedure is also referred to as the Brazilian test. For a cylindrical flat-faced tablet, the tensile strength (σ) can be calculated as follows (11):

$$\sigma = \frac{2F}{\pi D t} \tag{1}$$

where F is the force needed to fracture a cylindrical flat-faced tablet of thickness t along its diameter D.

The application of the compression test to calculate a tensile strength requires that the tablet fails by a normal tensile failure. It is normally considered that a tensile strength can be calculated from a diametric compression test also in cases when tablets fail by double-cleft and triple cleft failures (see above). However, when a tablet fails by a shear or compressive failure, the tensile strength equation cannot be used.

The equation is derived from a stress analysis in terms of how the principal stresses develop during application of a load (see further below). It has thus been pointed out (17) that the tensile strength equation is not a simple correction for tablet size but is the result of a stress analysis. Further corrections of the tensile strength equation for other indicators of the size or the size-weight ratio of a tablet, such as the relative volume or relative density, is thus not advisable.

The spread in tensile strength of tablets is normally expressed as a range or an arithmetic standard deviation, i.e., it is assumed that the variability in tensile strength can be represented by a normal distribution. It has however been suggested (18,19) that the variation in tensile strength of tablets can be satisfactorily represented by the Weibull function and the variability can thus be described alternatively by the Weibull modulus.

The tensile strength of tablets derived by compression can also be calculated for tablets of other shapes. For convex-faced cylindrical tablets, an equation has been derived by Pitt et al. (20,21) in which both the height of the cylinder and the thickness of the whole tablet are included. More on, the tensile strength for squared-shaped compacts can be calculated and the procedure has been used also in pharmaceutical studies (22). In that study, it was shown that tablets prepared by uni-axial compression have different tensile strength in different directions of measurement.

Tensile Strength by Alternative Methods

As an alternative to diametral compression of the tablet, a tensile strength can be derived by the bending of a tablet, a method also referred to as flexure testing (23). Three- or four-point bending methods are in use in this context.

Finally, another procedure of deriving a tensile strength (6,24,25) is to pull the tablet along the main axes of the tablet until it fails. This test has been denoted an axial tensile strength method and is suggested to be used primarily as a means to detect weaknesses in the compact in the axial direction, which is an indication of capping or lamination of the tablet.

Stress Analysis and the Tensile Strength Test

As mentioned, the equation normally used to calculate the tensile strength of a tablet from a diametrical compression test [Eq. (1) above] may be inferred from a rigorous stress analysis. To benefit the interested reader, the underlying procedure will be described in this section. Before turning our attention to the diametrical compression test, we will say a few words about stress in general. A more thorough discussion may be found in textbooks on solid mechanics (1,26).

Stress

The concept of stress in a continuous body dates back to Cauchy, and expresses the interaction of one part of the body with another part via surface forces or tractions. Consider a deformable body in its current configuration, as depicted in Figure 3, and introduce an imaginary surface through the body, whose orientation is specified by its

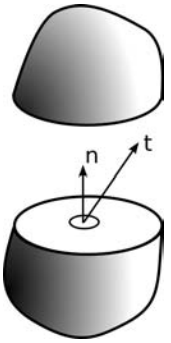


FIGURE 3 Definition of stress.

unit outward normal $\hat{\mathbf{n}}$. The action of the material outside the surface on the adjacent material inside the surface may then be specified in terms of the traction $\mathbf{t} = \mathbf{t}(\hat{\mathbf{n}})$ i.e., the force per unit area. As indicated, the traction depends on the orientation of the surface (and in general, also upon time and location, but these dependences have not been explicitly indicated). Moreover, from the balance of linear momentum (or force in the static case), expressed by Newton's laws, it follows that the traction in fact depends linearly on the surface normal. This linear dependence enables the (Cauchy) stress $\boldsymbol{\sigma}$ to be introduced as a linear transformation between the direction of the surface and the surface force it experiences. Linear transformations of this type that map vectors onto vectors constitute second order tensors and may be represented as matrices. Finally, from the balance of angular momentum (or torque in the static case), it follows that the stress tensor and its matrix representation are symmetric. If we for simplicity restrict ourselves to the two-dimensional case we may thus represent the Cauchy stress as

$$\boldsymbol{\sigma} = \begin{pmatrix} \sigma_{xx} & \tau_{xy} \\ \tau_{yx} & \sigma_{yy} \end{pmatrix} \tag{2}$$

In Eq. (2), σ_{xx} and σ_{yy} represent normal stresses on surfaces whose normals are parallel to the x and y axes, respectively, while $\tau_{xy} = \tau_{yx}$ represent shear stresses on these surfaces (which are equal since the stress tensor is symmetric). These stress components are indicated by solid arrows in Figure 4. Positive normal stresses are tensile while negative ones are compressive (note, however, that an opposite sign convention sometimes is used, most notably in the soil mechanics literature). From the interpretation of

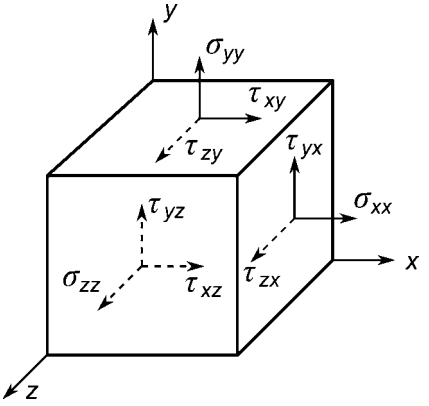


FIGURE 4 Components of the stress tensor. The components needed for a two-dimensional (plane stress) analysis are represented by solid arrows, while the remaining ones are indicated by dashed arrows.

the elements of the stress tensor in Eq. (2) it is realized that the matrix representation (but not the tensor itself) will change if another set of x and y axes are used.

Principal Stress

According to the discussion in the preceding section, the traction on any plane through a certain point in a continuous body may be obtained as the product of the stress tensor and the outward unit normal to the plane, an operation that formally may be represented as $\mathbf{t} = \boldsymbol{\sigma} \cdot \hat{\mathbf{n}}$. The direction of the traction is in general different from the direction of the unit normal, i.e., the surface force has both normal and tangential components. There are, however, exceptional directions, for which the surface normal and traction are parallel, known as principal directions. In fact, since the stress may be considered as a symmetric linear mapping, there are in general three mutually orthogonal principal directions $\hat{\mathbf{n}}_i, i = 1, 2, 3$ (two for the two-dimensional case) and three corresponding principal stresses σ_i, which thus are defined by $\mathbf{t} = \boldsymbol{\sigma} \cdot \hat{\mathbf{n}}_i = \boldsymbol{\sigma}_i \hat{\mathbf{n}}_i$. As mentioned above, the matrix representation of the stress depends on the choice of coordinate axes, and a particularly simple, diagonal representation is obtained if the coordinate axes are chosen to coincide with the principle directions:

$$\boldsymbol{\sigma} = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix}. \tag{3}$$

It should be noted, however, that the principal directions and stresses generally are different at different locations of the body, and that the principal directions determined for one point in general thus do not result in a diagonal representation of the stress also for other points of the body.

Stress Distribution for Diametrical Compression Tests

Let us consider the stress distribution in a tablet of cylindrical shape (diameter D and thickness t) subjected to a diametrical compression test. The traction must vanish on any unloaded surface, and thus in particular on the flat surfaces of the tablet. It is therefore natural to assume that traction components parallel to the normal of the flat surfaces vanish throughout the tablet, an assumption which leads to a state of plane stress, which means that the stress distribution effectively is two-dimensional and that the stress tensor therefore may be represented by a two-by-two matrix as in Eq. (2). For simplicity, we will also assume that the loading may be represented by point loads (i.e., that the contact between the platens and the tablet is a line if the thickness dimension of the tablet is retained). This latter assumption greatly simplifies the solution of the problem, but needs to be relaxed for cases of practical interest, as discussed below. Despite these simplifying assumptions, it may appear to be a formidable task to determine the stress in every point of the tablet. Fortunately, however, the stress distribution may be constructed relatively straightforwardly by superposition of terms representing each point load and a correction that makes the traction vanish on the circumference. We will briefly sketch the procedure. As before, positive principal stresses are tensile and negative ones compressive. Shear stresses do, on the other hand, not present themselves as principal stresses, since shear stresses correspond to tangential tractions which vanish when principal directions are selected as coordinate axes. Knowing the principal stresses and directions at a particular point, it is possible to determine the traction on any plane through that point. In particular, a geometrical construction, referred to as a Mohr diagram, may be used to

illustrate how the normal and tangential (shear) components of the traction depend on the orientation of the plane.

It may be assumed that one point load, i.e., an applied force F, is equilibrated by a radial stress distribution centered at the point of application of the load (Fig. 5A). This in turn means that the traction on any semicircular surface around the load will be in the radial direction, and equilibrium is obtained provided the radial stress is (26,27)

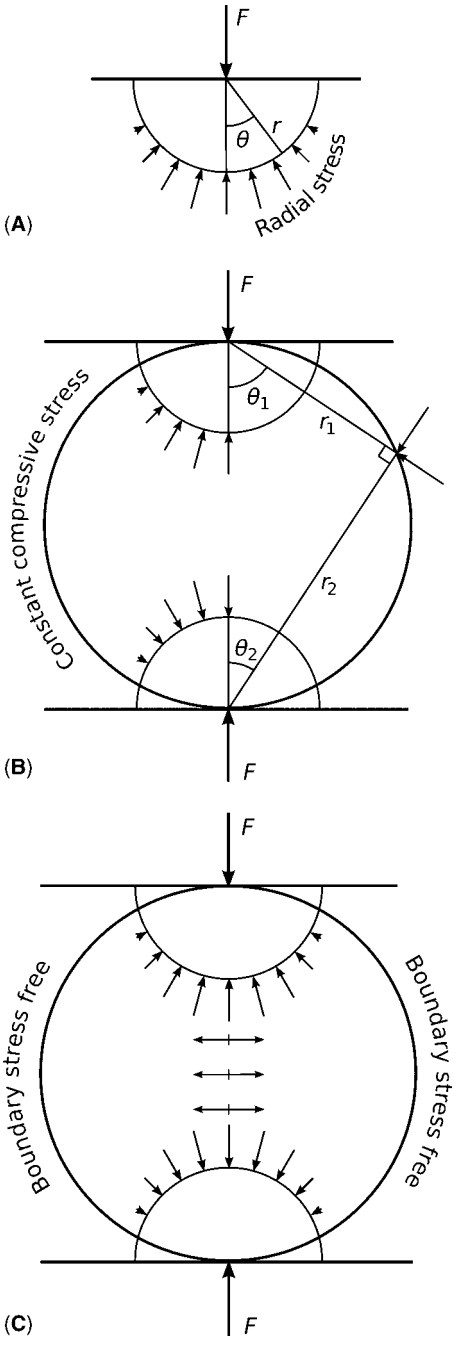


FIGURE 5 Construction of the stress distribution for the diametrical compression test: (**A**) Stress distribution for one point load, (**B**) stress distribution for two oppositely directed point loads, and (**C**) final stress distribution.

$$\sigma_{rr} = -\frac{2F}{\pi t}\frac{\cos\theta}{r}, \tag{4}$$

where r and θ are defined in Figure 5A.

Now consider the situation depicted in Figure 5B, which shows the stress generated by two oppositely directed point loads, as in the diametrical compression test. Since the material response is assumed to be linear, the effect of these two point loads may be obtained as the superposition of the effects of the individual loads. Clearly, the traction on the circumference is non-zero, which means that the obtained stress field cannot be the correct solution. However, whenever the point of interest lies on the circumference, two special conditions are fulfilled: First, the angle between r_1 and r_2 is 90 degrees, and, second, $\cos\theta_1/r_1 = \cos\theta_2/r_2 = 1/D$, where D is the diameter of the tablet. These two conditions between them assure that the contributions from the two point loads are equal and moreover result in a state of hydrostatic compressive stress. Thus, to obtain the desired solution, all that needs to be done is to add a hydrostatic tensile stress that exactly cancels the compressive stress at the circumference, as illustrated in Figure 5C.

The stress on the diameter between the loads is of most interest for the interpretation of diametrical compression test results. With the origin in the center of the tablet (and the x axis to the right and the y axis upwards in Figure 5C), the non-zero stress components are (28)

$$\sigma_{xx} = +\frac{2F}{\pi Dt}, \tag{5a}$$

$$\sigma_{yy} = -\frac{2F}{\pi Dt}\frac{3D^2 + 4y^2}{D^2 - 4y^2}. \tag{5b}$$

Since the shear stress is zero along this diameter, the above stress components also represent principal stresses. As seen, σ_{xx} is positive and thus represents a tensile stress, which is constant along the diameter [compare Eq. (1) above]. On the other hand, the compressive stress σ_{yy} (note the negative sign) increases in magnitude from the value $-6F/(\pi Dt)$ obtained in the tablet centre towards minus infinity when either of the loading points is approached. Since the tensile stress is constant, this analysis indicates that tablet failure could start at any point between the two loads. Moreover, since the minimum compressive stress is three times larger in magnitude than the tensile stress, the compressive strength of the tablet needs to be at least three times larger than the tensile strength in order to ensure a tensile failure.

The above analysis is not completely satisfactory, however, since it predicts an infinite compressive stress at the loading points, as a result of the assumption of concentrated point loads, which would indicate that the tablet fails in compression at either of the loading points and not in tension in the central part. However, for the typically used flat platens, the load is instead distributed over finite areas of contact, which means that the stress is everywhere finite. An approximate analytical solution for this case has been derived by Wright (29), which is compared to the solution obtained for point loads in Figure 6. As may be seen in the figure, the changes in the stress caused by the change in loading conditions is confined to a region in the vicinity of the platens, and the stress along the major part of the diameter between the loads is still well approximated by Eqs. (5a) and (5b). In particular, the tensile stress may still be computed with Eq. (5a).

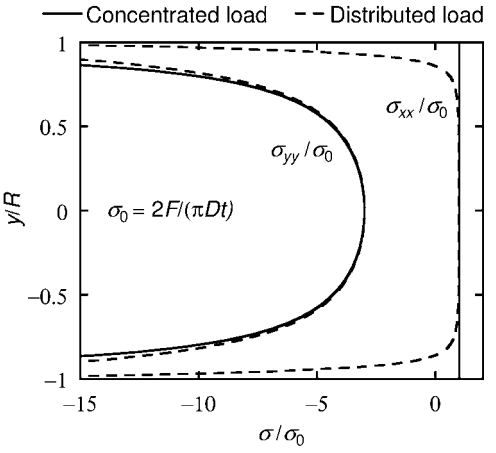


FIGURE 6 Stress along the loaded diameter in diametrical compression tests for concentrated and distributed loads.

AGGLOMERATE TENSILE STRENGTH

Agglomerate Microstructure

Agglomerates may be defined as clusters of primary particles held together by adhesive and/or cohesive forces. The commonly used theoretical approaches to agglomerate strength are therefore based on considerations of the number and strength of bonds between clearly identifiable, distinct primary particles. Although the original particles are fractured and deformed during the formation of a tablet, the literature indicates (30) that the description of a tablet in physical terms as a cluster of primary particles is a reasonable approximation. Theoretical approaches to the strength of dry agglomerates are thus applicable also in the discussion of tablet strength.

A Micromechanical Approach: Rumpf's Theory

Conceptually, it appears natural to consider the agglomerate strength as a function of the strength and number of the bonds between primary particles. The strength of the inter-particle bonds may here be defined as the force required separating the particles from each other, but may also be expressed in terms of surface energy. The inter-particle bonds in any real agglomerate will generally be of different strength, but is usually assumed that a reasonable approximation is obtained by using a representative average value. The influence of contact number on the agglomerate strength does, on the other hand, depend on the way the agglomerate is assumed to fail.

The simplest (though probably not the most accurate approach) is to assume that simultaneous breakage of all bonds in a certain plane through the agglomerate is required for failure. The agglomerate tensile strength may then be obtained as the sum of the strength (expressed in terms of the separation force F) of the individual primary particle bonds in the fracture plane. This assumption underlies the perhaps most widely known expression for agglomerate tensile strength, derived by Rumpf (31,32), who considered a random packing of mono-dispersed spheres and obtained:

$$\sigma_t = \frac{9(1-\varepsilon)QF}{8\pi d^2} \approx \frac{(1-\varepsilon)}{\varepsilon}\frac{F}{d^2}. \tag{6}$$

In this equation, d is the primary particle diameter, ε is the agglomerate porosity, and Q is the coordination number, i.e., the average number of contact points for one primary particle. The second expression in Eq. (6) is obtained by assuming an empirical relationship between coordination number and porosity, of the form $Q \approx \pi/\varepsilon$ (31), which was based on data from Smith et al. (33).

The micromechanical description of agglomerate strength may be refined by considering the dynamics of failure, which has been extensively studied within the realm of fracture mechanics. Let us at this point therefore consider some important fundamental fracture mechanics concepts, such as stress the intensity factor and fracture toughness.

Stress Intensity Factor and Fracture Toughness

The separation of a solid body into two or more fragments is generally regarded to occur through the propagation of one or several cracks through the material (2,34). In real materials, cracks or defects that eventually could evolve into cracks almost always exist. Considering the agglomerate microstructure, it is evident that voids of different sizes are abundant, which could serve as the origin of cracks. Although stress and strain continue to be very important for the description of cracks and failure, additional concepts—like stress intensity factors and energy release rates—are also needed. Generally, a distinction is made between brittle and ductile fracture. Brittle fracture is characterized by the fact that no significant inelastic deformation occurs prior to failure, and the material is thus able to withstand only relatively small elastic straining. Conversely, ductile behavior is characterized by plastic (permanent) deformation that ultimately may lead to failure. Some types of agglomerates are able to deform plastically without fracture, but a brittle behavior is more common, and will therefore be the topic of this section.

It is possible to identify three different modes of fracture, which are sketched in Figure 7 (34,35). Mode I crack opening is caused by tensile stress, whereas the remaining ones (Modes II and III) are caused by shear stress. Mode II is also referred to as in-plane shear and Mode III as anti-plane shear: If one looks at a crack 'from the side' as in Figure 8, the shear stresses are in the plane for Mode II and orthogonal to the plane for Mode III.

Let us consider the situation depicted in Figure 8, which shows a symmetric (Mode I) crack opening. As mentioned, this mode is typical for a tensile failure, but the results are qualitatively the same for Modes II and III as well (a thorough discussion of crack opening modes and crack tip fields may be found in texts on fracture mechanics [e.g., (34,35)]. Since the material is assumed to behave in a brittle manner, we may safely assume it to be linearly elastic (except possibly at a small zone in the very vicinity of the crack tip, where the deformation may be extensive). If we for simplicity restrict our attention to the positive r axis in Figure 8, the non-vanishing components of the stress tensor in the vicinity of the crack tip may be written in the generic form.

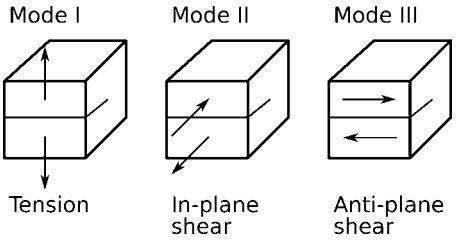


FIGURE 7 Modes of fracture.

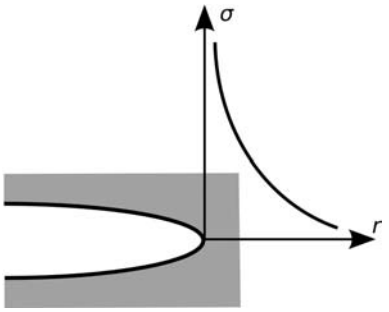


FIGURE 8 Crack tip for a Mode I crack.

$$\sigma_{xx} = \sigma_{yy} = \frac{K_I}{\sqrt{2\pi r}}, \tag{7}$$

where K_I is a constant and r is the distance from the crack tip. From Eq. (7) it is evident that the stress field is singular at the crack tip, i.e., the magnitude of the non-vanishing components of the stress tensor tends to infinity as $1/\sqrt{r}$ when the crack tip is approached. This is also the reason for Eq. (7) being generic: The stress may in general be expressed as a series containing other terms than the one in Eq. (7), but the additional terms are all bounded, which means that the singular term will dominate sufficiently close to the crack tip.

The result expressed by Eq. (7) is typical in the sense that stress concentration generally occurs in the vicinity of cracks and other flaws in a material. Although the stress is infinite at the crack tip itself, it is clear that the amplitude of the stress may be uniquely characterized by the constant K_I, which is known as the stress intensity factor. The stress intensity factor depends on the mode of crack opening, as indicated by the subscript, and also on the size of the crack and the loading conditions, typically being proportional to the applied stress σ and to the square-root of crack size a, i.e.:

$$K_I \propto \sigma\sqrt{a}. \tag{8}$$

It is generally assumed that a crack starts to grow once the stress intensity factor K_I exceeds a certain material-specific value,K_{Ic}, called the critical stress intensity factor or fracture toughness. This, in turn, leads to the well known result that the strength of a material generally is inversely proportional to the square-root of its defect size, i.e.,:

$$\sigma_{\max} \propto K_{Ic}/\sqrt{a} \propto 1/\sqrt{a}. \tag{9}$$

Although we have chosen to use the stress intensity factor as the basic variable in our discussion, it deserves to be mentioned that the same conclusions could have been drawn from a consideration of the energy released when a crack is advanced. In fact, a unique relationship exists between the stress intensity factor and the energy release rate (the energy release rate is proportional to the square of the stress intensity factor, the constant of proportionality being the reciprocal of an appropriate elastic modulus for the material). One may thus equivalently assume that a crack starts to grow once the energy release rate exceeds a certain material-specific value. This is the Griffiths energy criterion for fracture.

A Refined Micromechanical Approach: Kendall's Theory

Contrary to Rumpf, Kendall (36) assumed that agglomerate failure is caused by crack nucleation at flaws followed by crack propagation through the agglomerate, and used

fracture mechanical considerations as described in the previous section to determine the agglomerate strength. We will briefly indicate the procedure.

Describing the primary particles as linearly elastically spheres, the inter-particle contact area was first determined from the equilibrium between the surface energy Γ and the elastic resistance of the spheres. Then, by considering regular assemblies of particles, Kendall derived expressions for the effective Young's modulus and energy release rate. In the latter, the fracture energy Γ_c was used instead of the surface energy Γ, since experiments indicated that the energy release rate otherwise would have been underestimated. Knowledge of the energy release rate and the elastic modulus makes possible the determination of the critical strength intensity factor, and could thus be used to determine the strength of the regular arrangement of particles along the lines indicated in the preceding section. Kendall finally argued that any real agglomerate contains macroscopic flaws that would reduce the agglomerate strength, and again using fracture mechanical arguments expressed the agglomerate fracture strength as:

$$\sigma_f = 15.6\frac{\phi^4\Gamma_c^{5/6}\Gamma^{1/6}}{\sqrt{dc}} \tag{10}$$

In this equation, $\phi = 1 - \varepsilon$ is the solid fraction, d is the particle diameter, and c is the size of the macroscopic flaw. Except for the pre-factor, this expression would also be valid for the tensile strength. Note, however, that the assumptions made during the derivation are consistent with agglomerates without binder.

POWDER COMPACTIBILITY

Powder Compressibility and Compactibility

An associated term to the mechanical strength of a tablet is powder compactibility (also referred to as tabletability and tablet forming ability). The term compactibility was introduced by Leuenberger (37) in order to clearly differentiate between two functional properties of a powder during its processing, i.e., the compressibility and the compactibility of a powder. The compressibility is defined as the propensity of a powder, held within a confined space, to reduce in volume while loaded. The compressibility is normally described by the relationship between tablet relative volume or relative density (porosity) and the compression pressure and several equations for such relationships are reported in the literature (38). The compactibility may be defined as the ability of a powder to form a coherent tablet as a result of compression. The ability of a powder to cohere is normally understood in a broad sense, i.e., a powder with a high compactibility readily forms tablets with a high resistance towards fracturing and without tendencies to cap or laminate.

Due to the importance of the compactability of a powder or a powder blend in the formulation of tablets, aspects of powder compactibility are frequently reported in the literature. The focus of such studies is often on the relationship between powder properties and the mechanical strength of the tablet and the overall objective is often to identify material factors that control powder compactibility. Different approaches to derive measures of the powder compactibility are used in such studies. In this section, we will firstly give an brief overview of measures (categorized as descriptors or indicators) of powder compactibility. In the discussion of compactibility descriptors, we have used a categorization of methods and models for quantification of compactibility published by Sonnergaard (39). In the subsequent section, we will thereafter discuss material properties that control powder compactibility.

Descriptors of Powder Compactibility

Single-Point Values

A simple type of descriptor of powder compactibility is a single-point value. Two types of single-point values are used in the literature. The dominating type is the mechanical strength of tablet formed at a given compaction pressure (40,41) but the mechanical strength of a tablet formed at a certain tablet porosity is an alternative similar type of approach. The second type of single-point value is the compaction pressure needed to form a tablet of a predetermined mechanical strength (42).

For both types of descriptors, the normal application is that the derived descriptor is used as a means to compare materials regarding their tablet forming ability. However, since the dependency of the mechanical strength of tablets on compaction pressure or tablet porosity may vary significantly between materials, a more comprehensive understanding of the powder compactibility is obtained by studying the relationship between tablet tensile strength and the compaction pressure or between tablet tensile strength and tablet porosity. Such relationships are often described graphically but a series of procedures aiming at deriving quantitative measures or descriptors of the compactibility from such relationships have also been used.

Tensile Strength—Tablet Porosity Relationship

The relationship between tablet strength and tablet relative density or porosity is normally non-linear, characterized by a concave shape. The most commonly used expression for the tablet tensile strength-tablet porosity relationship is probably the equation often referred to as the Ryshkewitch equation (43) and it is stated (44,45) that this equation represents well the tensile strength-porosity relationship for a wide range of materials. Tablet porosity is a global tablet property but a change in tablet porosity due to further compression will also change the micro-structure of the tablet, i.e., the size of particles and inter-particulate voids of the tablet and the structure of the inter-particulate contacts. The mechanical strength can thus be expected to show some relationship with tablet porosity. The Ryshkewitch equation can be written in the following form:

$$\ln \sigma = \ln \sigma_0 - k\varepsilon, \tag{11}$$

where ε is the porosity of the tablet, σ_0 is the tensile strength of a tablet of zero porosity and k is a constant, sometimes denoted the bonding capacity. This constant may thus be used as a descriptor of powder compactibility and has, for example, been used in the assessment of the tensile strength of tablets formed from binary mixtures of particles (44) (Fig. 9).

An alternative procedure to describe the relationship between tablet strength and tablet porosity (normally expressed as a tablet relative density) is to use a percolation equation, i.e., a power law of the following form (46):

$$\sigma = S(\rho - \rho_c)^q, \tag{12}$$

where ρ is the relative tablet density (i.e., $1 - \varepsilon$), ρ_c is the percolation threshold (i.e., the relative tablet density at which the tensile strength changes abruptly), S is a constant referred to as a scaling factor and q a scaling exponent. The scaling factor may be used as a descriptor of the compactibility in terms of a measure of how the tensile strength changes with relative density, provided that a proper value of the scaling exponent is used. The percolation threshold may be seen as a single-point descriptor of powder compactibility.

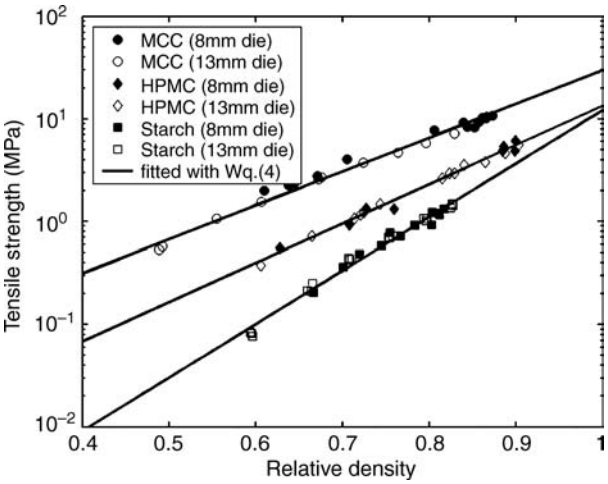


FIGURE 9 Examples of the relationship between tablet strength and tablet relative density for three materials, expressed as a ln—lin relationship in accordance with the Ryshkewitch equation. (From ref. 44).

Tensile Strength—Compaction Pressure Relationship

The relationship between tablet strength and compression pressure may be complex. However, excluding a situation where cracks are formed in the tablet or if capping occurs during compaction, which is often reflected as a sudden drop in the tablet strength—compaction pressure profile (40), the relationship between tablet strength and compaction pressure, i.e., a compactibility profile, can be approximated as a three region relationship: A lower region, where no coherency has been reached, an intermediate region at which the tablet strength increases with compaction pressure, and an upper region where the tablet strength is again independent of the compaction pressure (Fig. 10). This upper plateau corresponds to a porosity of the tablet close to zero, at which the tablet behaves as an elastic body. The regions are separated by lower and upper tablet strength thresholds. This description of the compactibility profile is a percolation approach since the properties of the system change abruptly at the thresholds. In practice, sharp percolation thresholds cannot be expected and a relationship resembling a sigmoidal curve with a significant nearly linear portion could probably be expected. The fitting of strength–pressure relationship by the Weibull function, giving a sigmoidal curve, has also been used in the literature (47). Based on this three region compactibility profile, four compactibility descriptors can be derived, i.e., the upper and lower pressure thresholds, the slope of the linear portion and the maximum tablet strength (denoted σ_{max} in Figure 10).

In the literature, a series of simple descriptors of the relationship between tablet tensile strength and compaction pressure has been used. The slope of a lin–lin relationship has been argued to be the preferable descriptor (39), which is in accordance with the relationship discussed above (Fig. 11). Since it may occur that two materials give a similar slope but different tensile strengths at a given pressure, the combination of the slope from the tablet strength-compaction pressure profile with other descriptors, such as the upper and lower pressure thresholds, gives a more comprehensive description of the compactibility of a powder. The slope from other relationships between tablet tensile strength and compaction pressure, a lin–log (48) and a log–log (49), have also been reported.

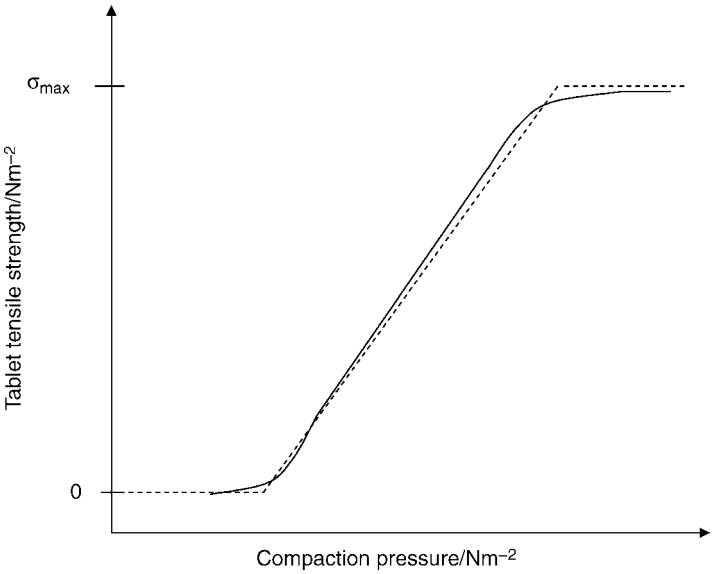


FIGURE 10 Illustration of a sigmoidal compactiblity profile (solid line) and a percolation type of compactibility profile (dotted line).

In addition to empirical descriptions, attempts to mechanistically model the relationship between tensile strength of tablets and the compaction pressure in terms of theoretical or semi-empirical expressions have been presented in the literature, for example by Leuenberger (37) and Alderborn and coworkers (50,51). Both these approaches are based on the modeling of the evolution of the inter-particulate bond structure during compaction. Implicit is thus that the tablet tensile strength has some proportionality to the sum of the bonding forces of the inter-particulate bonds acting over a unit area of fracture surface. In practice, tablets may however fail by a combination of an inter- and an intra-particulate fracture process. The consequent evolution in tablet

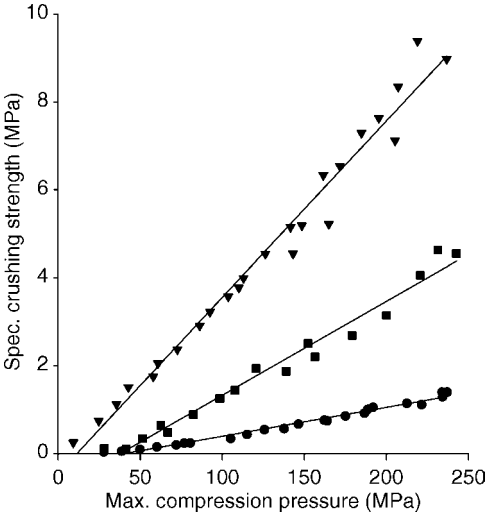


FIGURE 11 Examples of the relationship between tablet strength and compaction pressure for three materials, sodium carbonate (highest compactibility), sodium chloride (intermediate compactibility) and sodium bicarbonate (lowest compactibility). *Source*: From Ref. 39.

tensile strength due to the change in tablet micro-structure is in the models related to an end-point, representing the maximum tablet tensile strength that can be reached for a given material (compare Fig. 10).

Leuenberger assumed that in a tablet, a number of bonding and non-bonding points exists and their relative number depends on the applied pressure during compression and the tablet relative density. The equation has the following form:

$$\sigma = \sigma_{\max}[1 - e^{(\gamma \Pi \rho)}], \tag{13}$$

where P is compaction pressure, $\sigma_{\max}$ is the maximum tensile strength that can be reached and γ is the compression susceptibility which describes the compressibility of the powder and has the unit pressure^{-1}.

Alderborn assumed that the evolution in tablet strength is proportional to the evolution of the effective contact area between particles in a cross section of the tablet. The effective contact area was proposed to be proportional to the product of the number of inter-particulate junctions and the mean area of contact formed at the inter-particulate junctions in a tablet cross section. The contact process between particles during compression can be viewed as the formation of adhesive inter-particulate joints of successively increased dimension with reduced tablet porosity. The equation has the following form:

$$\sigma/\sigma_0 = (P - P_0)/C, \tag{14}$$

where P_0 is the minimum compaction pressure that is required to from a coherent tablet and C is a compression parameter that indicates the effective deformability of the particles during given compression conditions. The significance of the expression is that the evolution in tablet strength is controlled mainly by the plasticity of the particles which also will control the range of compaction pressure in which the tablet strength will evolve with pressure.

Indicators of Powder Compactibility

In addition to different types of descriptors derived from compactibility profiles, indices have been derived that are suggested to describe in some quantitative way the ability of powders to cohere, i.e., indicators of powder compactibility. The most frequently used indicators in formulation development and scientific work are probably the indices of tableting performance derived by Hiestand and co-workers. A comprehensive description of the use of these indices are given elsewhere (52). Primarily two of the Hiestand indices of tableting performance are suggested to reflect powder compactibility, i.e., the bonding index and the brittle fracture index. Both these indices are based on the measurement of tensile strength and hardness of compacts and ratios between these properties give a dimensionless index. The bonding index (BI) is defined as:

$$\mathrm{BI} = \sigma/H, \tag{15}$$

where σ is the tensile strength of the compact and H is the hardness of the compact. The brittle fracture index (BFI) is defined as:

$$\mathrm{BFI} = [\sigma/\sigma_H - 1]/2, \tag{16}$$

where σ_H is the tensile strength of a compact containing a hole or perforation (corresponding to macroscopic defect). The bonding index is proposed to reflect the ability of a

powder to cohere into a tablet of high tensile strength while the brittle fracture index is proposed to reflect the ability of a tablet to resist fracturing, such as capping, during tablet production.

MATERIAL PROPERTIES OF IMPORTANCE FOR POWDER COMPACTIBILITY

Factors Controlling Powder Compactibility

A large number of studies can be found in the pharmaceutical literature as well as within other related disciplines in which factors which affect the mechanical strength of tablets or the compactibility of powders are discussed. These factors can be categorized into three main groups that however are interrelated, i.e., formulation factors, processing factors and environmental factors (primarily relative humidity). Of special interest from a formulation perspective is the physical properties of the particles used in the formulation and in the following section, we will discuss the importance of physical properties of particles for their compactibility. In this discussion, we will make a distinction between two types of particles, referred to as particulate and granular solids. The reason for making the distinction is that the difference in the particle physical structure will affect the behavior of the powder while compacted and the possibilities to modulate or control the compactibility of the powder. The term particulate solids refers in this chapter to a powder consisting of dense particles, i.e., particles that are non-porous or of low porosity and that are not agglomerates of smaller primary particles, while the term granular solid refers to a powder consisting of granules, i.e., particles that are clusters or agglomerates of smaller particles and formed by some particle size enlargement process. Granules normally consist of drug and excipient particles and a binder that is distributed on the surface of these substrate particles.

As stated above, the literature indicates, e.g., that a simplified description in physical terms of a tablet formed from particulate (30,53,54) or granular solids of a normal tablet porosity is a cluster of discrete particles adhered to each other into a coherent specimen. The proposed dominant physical structure of a tablet is shown in Figure 12, showing the upper surface of a tablet formed from microcrystalline cellulose granules. The basic structural parts forming such a coherent cluster are the particles, the voids between these particles and the inter-particulate joints at which the particles adhere to each other. The tablet micro-structure together with the adhesive capacity of the solid surface will control the fracture process (see above) and the tablet strength.

The Compactibility of Particulate Solids

Particle Mechanics

During compression, the powder will reduce in volume and on the particle scale, the processes involved in the compression of particulate solids are particle rearrangement, particle fragmentation and particle reversible and permanent deformation. Fragmentation and permanent deformation of particles are the two processes that will control the evolution in tablet micro-structure in terms of the inter-particulate joints and voids and they are thus sometimes denoted strength-producing compression mechanisms (55). In a simplified way, fragmentation can be described as affecting the number of inter-particulate bonds while permanent deformation relates primarily to the area of contacts developed between particles with a subsequent increased bonding force (50). Reversible

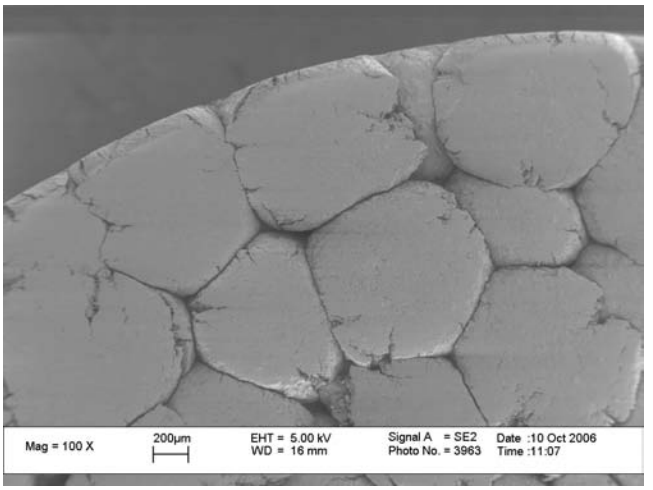


FIGURE 12 A photomicrograph of the upper surface of a tablet formed from microcrystalline cellulose granules, illustrating the proposed physical structure of a tablet.

or recoverable deformation, i.e., elastic and visco-elastic deformation, is traditionally considered as a disruptive rather than a strength-producing mechanism. The functional behavior of a powder during compression, i.e., to what degree the particles will deform and fragment, is possibly controlled by the mechanical properties of the solid (56), i.e., a brittle material is prone to fragment while a tough material is prone to deform during powder compression. Relationships between the molecular and crystalline structure and the mechanics of solids have also been discussed in the literature (57,58).

Although this general conception of the importance of functional mechanics of particulate solids for powder compactibility is widely accepted since decades, there are few reports that have substantiated this conception in experimental terms and have discussed their relative importance.

In a series of papers on the compactibility of lactose powders (59,60), a relationship was observed between the tablet strength and the tablet surface area for tablets formed from different types of crystalline lactose. This finding was later interpreted (61) in terms of a relationship between tablet surface area and the number of inter-particulate contacts in the fracture plane. It was thus suggested that an increased degree of fragmentation of particles during compression will improve the fracture strength of the tablets.

In two consecutive papers, Sebhatu el al. (62,63) investigated the compactibility of amorphous lactose powders. The deformability of the particles, a property that could be modulated for the amorphous particles by their moisture content, was assessed by the yield pressure. By accounting for the yield pressure, a single relationship between tablet strength and compaction pressure was obtained for the powders studied. It was thus concluded that increased degree of deformation of particles during compression will improve the strength of the tablets. The importance of particle yield strength or hardness was later supported (51) by studying the difference in evolution in relative tensile strength of tablets formed from sodium chloride and sucrose (Fig. 13).

The compression behavior of particles will also affect the compactibility of a binary mixture consisting of a main component and a second component added in a low proportion, typically a dry binder, a disintegrant and a lubricant. Such a binary mixture thus formed is often referred to as structured, interactive or ordered mixtures. The additive can

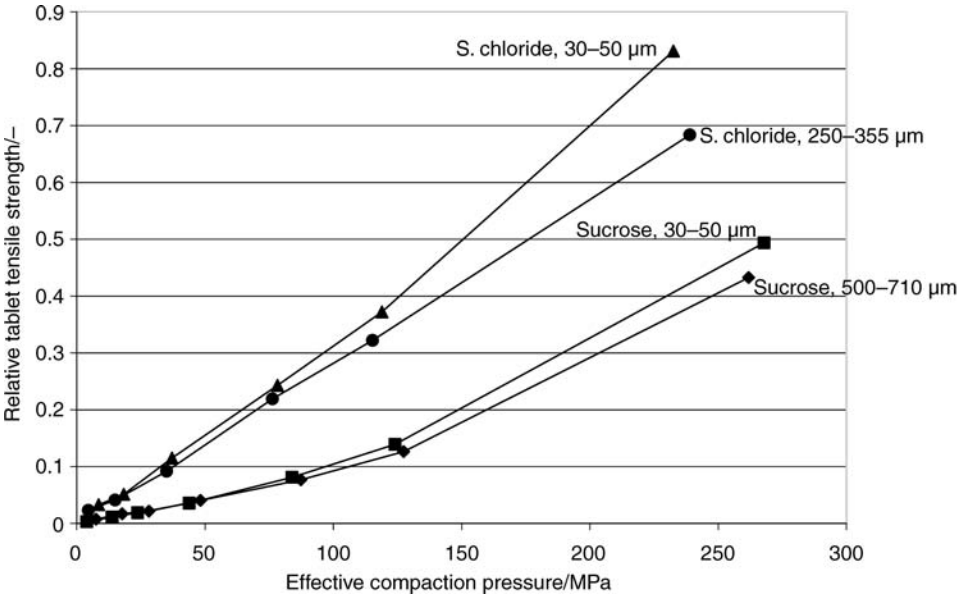


FIGURE 13 The evolution in relative tablet tensile strength with compaction pressure for four powders, i.e., two particle size fractions of sodium chloride and of sucrose. The difference in relative compactibility is explained by a difference in hardness of the two materials. *Source*: From Ref. 51.

either increase or decrease the compactibility of the mixture relative to the compactibility of the main component alone. The compactibility enhancing or reducing effect of the additive is related to the compression mechanics of the main component, primarily its fragmentation propensity (64–67). A material of high fragmentation propensity will show a limited change in compactibility due to the addition of the second component, i.e., show a high dilution capacity, while the reverse applies to a material of low fragmentation propensity.

The literature on the importance of the solid state properties, i.e., crystalline form (68–70), salt form (71) and the crystallinity (63,72,73) of the particles, as well as the moisture content of crystalline or amorphous particles (63,74,75) for the compactibility of powders is large. Variations in solid state and moisture content of powders represent important formulation factors. However, the fundamental role of such variations for the compactibility of a powder is possibly that they affect the bonding between particles through an effect on the compression mechanics, the dimensions or the surface energy of the particles. Relevant reports (63,70,75) concern the effect of crystal structure and moisture content (Fig. 14) on the plasticity of particles and the subsequent evolution of inter-particulate contact area and tablet strength.

Particle Dimensions

Besides the compression mechanics, the micro-structure of a tablet will possibly also be related to size and shape of the original particles. Since the particulate properties are properties that can be altered by processing (crystallization, agglomeration, milling, fractionation etc.), the relationship between particle size, size distribution and shape on one hand and powder compactibility on the other is widely reported on in the literature.

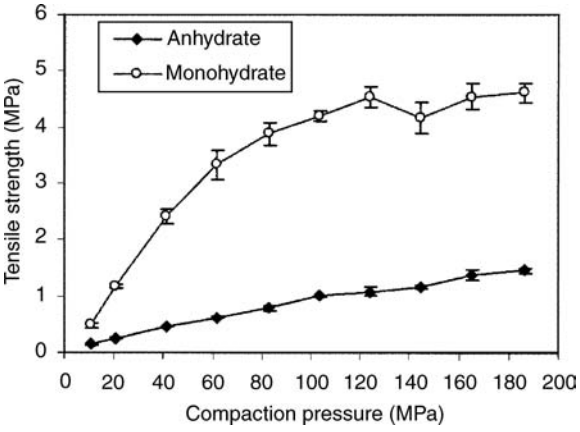


FIGURE 14 Compactibility profiles of the anhydrate and the monohydrate of hydroxybenzoic acid. The difference in compactibility is explained by a difference in plasticity of the particles due to the presence of water molecules in the crystal structure. *Source*: From Ref. 75.

The size of the particles to be compacted is often considered as a significant factor for tablet strength. It seems that the most common type of relationship between original particle size and tablet strength is that a decreased original particle size increases the tablet strength (40,51,60,72,76). A reduced original particle size may also reduce the compaction pressure needed to form a tablet (51). However, complex relationships that deviates from a simple relationship between particle size and tablet strength have also been reported (77).

Regarding the distribution in size of particles for their compactibility, it was recently shown (78) that this factor has a limited effect on the evolution in tensile strength during compression. It was observed, however, that the spread in particle size had an effect on a post-compaction increase in tablet tensile strength, demonstrating the complexity in the factors controlling the strength of a compact. The authors thus concluded that the particle size distribution may have an effect on powder compactibility due to a post-compaction reaction.

It has also been shown in the literature that the particle shape can significantly affect the compactibility of a powder (41,79,80). A general interpretation of data reported in these papers is that for particles which fragment to a limited degree during compression, an increased particle irregularity improved powder compactibility while for particles which fragmented markedly during compression, the original shape of the particles did not affect the tablet strength. Thus, the compression mechanics and the particulate properties may show an inter-dependence of each other. Finally, an attempt has also been made (81) to demonstrate the importance of surface roughness of particles for their ability to form a tablet.

Particle Adhesiveness

The transformation of a powder of low cohesivity into a tablet with strongly cohered particles is based on the formation of inter-particulate bonds or adhesive joints. The bonding process between solid surfaces is essentially an interfacial phenomenon and the surface energy of the solid is thus a factor of importance to consider in parallel to the tablet micro-structure (see above). The relationship between particle surface energy and powder compactibility is difficult to experimentally study since, ideally, it should involve the comparison of the tensile strength of tablets with similar microstructure. Thus, there are only few reports, e.g., (82), that have specifically focused on this

relationship but the reported data may be interpreted in such a way that an increased surface energy corresponds to an increase in powder compactibility. More recently, Li et al. (83) found a relationship between adhesion force, assessed by atomic force microscopy, of some particles and the tensile strength of tablets formed form these particles.

There are, however, several reports that have demonstrated the importance of a change in the property of the surface of the particles that could influence their surface energy for the compactibility of the powder. It is a well-known fact that the addition of a low proportion of a lubricant to a powder, e.g., (65) will reduce its compactibility significantly, i.e., the lubricant will adhere to the surface of the substrate particles and affect the interaction between the particles. Sakr and Pilpel (84) reported that when lactose particles were coated with increasing concentration of surfactant, the compactibility of the powders was subsequently reduced, most profoundly at low concentrations. Berggren et al. (85) compared the compactibility of some powders prepared by spray-drying from lactose solutions with and without the addition of a polymer and a surfactant. It was reported that the surface properties of the particles affected their adhesiveness and thus the tablet strength. Notable is that the presence of a surfactant reduced the powder compactibility.

The Compactibility of Granular Solids

Granule Mechanics

During compression of a granular solid in a confined space, it has been suggested that granules tend to keep their integrity and the tablet formed from the granules can in physical terms be described as a cluster of closely packed granules (53,54,86) with a dualistic pore system (87,88). The pores of such a tablet can be classified as inter-granular (voids between cohered granules) and intra-granular (pores between primary particles forming the granules). The mechanisms reported to be involved in the compression of a granular solid (89,90) are rearrangement, deformation (i.e., a change in shape of the granules), densification (i.e., granules reduce in volume), erosion (i.e., primary particles are abrased from the surface of the granules), cracking (formation of cracks in the granule surface) and fragmentation (i.e., original granules break down into smaller granules). It is recently reported that for pharmaceutical granules (91), the dominating mechanisms, i.e., compression rate controlling mechanisms, involved in the compression process of granules are cracking followed by plastic deformation followed finally by an elastic deformation of the whole tablet within the die.

During fracturing of a tablet structured as a cluster of cohered granules, the failure will often propagate between the granules and break the inter-granular bonds. In such a case, the stress needed to break the inter-granular junctions of the tablet during strength testing will, in simplified terms, be a function of the area of intimate contact established between the granules during the compression process and the strength of the adhesive bonds that coheres the granules. Thus, factors that control the contact process between granules during compression will also affect the tablet strength.

For granules that have sufficient strength to withstand breakage during handling, permanent granule deformation has been proposed to be the single most critical factor for the evolution in tablet strength tablet during compression (53,91,92). Thus, physical properties of granules that control their degree of deformation during compression are thus significant for the fracture strength of tablets. Granule deformation

involves the shearing of the granules and important factors for the readiness of the granules to shear and thus deform during compression are their porosity and their composition in terms of the mechanical properties of the granule forming particles and the presence of a binder.

By using a series of granules of consistent composition but of varying porosity, it has been shown (53,92) that an increase in granule porosity will increase the degree of deformation that is expressed during compression. Thus, an increased porosity facilitated deformation which corresponded to an increased compactibility of the granules (Fig. 15).

The mechanical properties of the granule forming particles will be of importance for the compactibility. It is for example common knowledge that granules formed from a capping prone material will show a poor compactibility (93), an observation that may be related to the elasticity of the primary particles from which the granules are formed. In addition, based on a comparison of the compression behavior and compactibility of granules of different composition but of the same range of granule porosity, it was suggested that the granule deformation propensity was affected by the hardness of the granule forming particles (92).

A material that interferes with and facilitates shearing of the granule can be described as an internal glidant that promotes the deformation propensity of the granule. An example of an internal glidant is a binder that is distributed as a film on the surface of the primary particles (94). Thus, the role of the binder in enhancing the compactibility of a granular solid may be to affect the degree of deformation of granules that occurs during compression, modulated by an increased deformation propensity, as well as to increase the adhesiveness of the granules (see below).

Granule Dimensions

In addition to the deformation propensity of granules, there are indications in the literature that dimensions of granules, i.e., granule size (90) and granule shape (95), may affect the degree of deformation that is expressed during compression although the deformation propensity of the granules seems to be constant. In case of the granule size, the change in degree of deformation was not accompanied by a corresponding change in compactibility while the reverse applied for the granule shape.

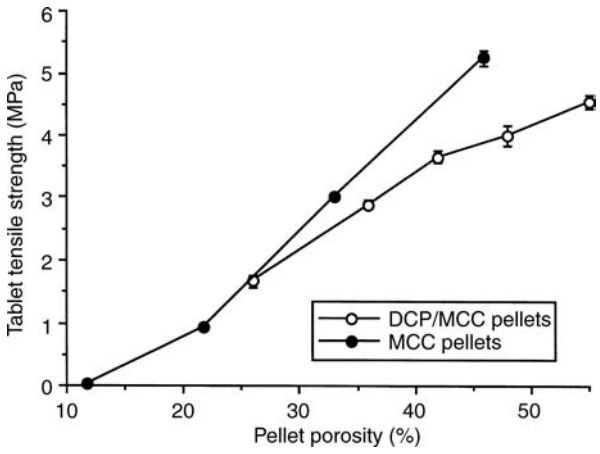


FIGURE 15 The importance of granule porosity for the compactibility of granular solids (formed from microcrystalline cellulose or from a mixture of microcrystalline cellulose and calcium phosphate). The difference in compactibility is explained by an effect of porosity on degree of deformation of the granules that is expressed during compaction. *Source*: From Ref. 92.

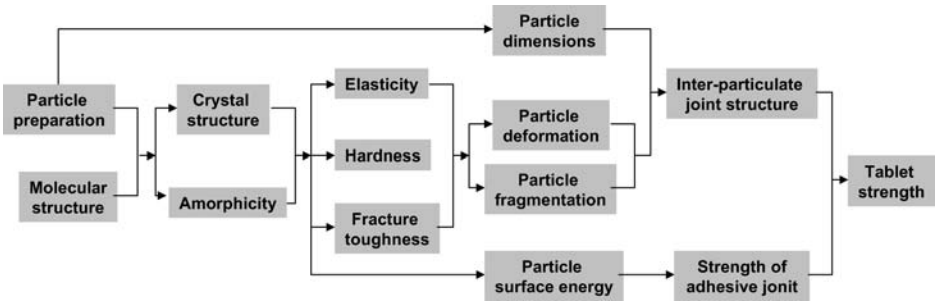


FIGURE 16 Compactibility map for particulate solids.

Granule Adhesiveness

The perception that the adhesiveness of the extra-granular surfaces will be important for tablet strength is demonstrated by the marked effect of the addition of a low proportion of a lubricant to the granular solid (96) for the compactibility of granules. Another example is the effect of intra-granular binder distribution for tablet strength. Since granules change in physical appearance during compression due to deformation, attrition and fracturing, the distribution of the binder within the granules prior to compression may affect the properties of the surfaces involved in bonding at the inter-granular junction of the tablet. It has been reported (97,98) that a peripheral localization of the binder, i.e., a concentration of the binder at the granule surface, may be advantageous for the compactibility of granular solids compared to a homogenous binder distribution. The explanation behind this statement is that the binder can thereby be used most effectively for the formation of inter-granular bonds. However, by comparing the compactibility of granules of similar porosity but of different intra-granular binder distribution (99), it was reported that granules of a homogeneous binder distribution showed higher compactibility than granules of an in-homogeneous binder distribution (i.e., with the binder located primarily at the external surface of the granules). This observation was explained by assuming that, owing to extensive deformation and some attrition of granules during compression, new extra-granular surfaces was formed during compression that originated from the interior of the granules. Such compression-formed surfaces were more adhesive when the concentration of binder increased.

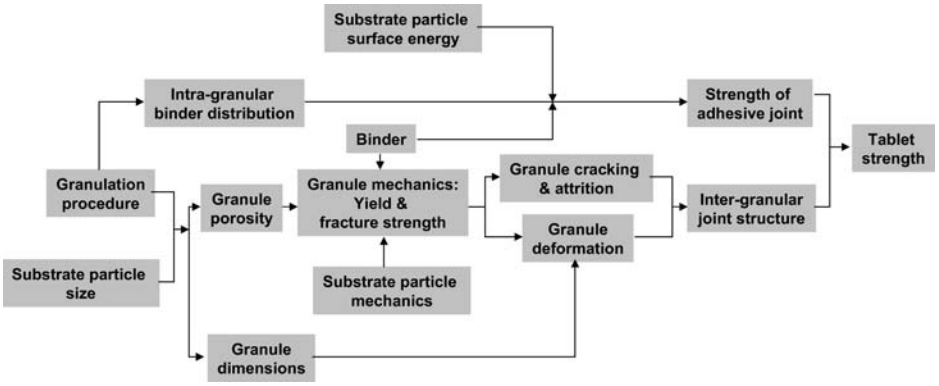


FIGURE 17 Compactibility map for granular solids.

As discussed above, the fundamental roles of the binder for the compactibility of a powder are twofold: Firstly, to modulate the plasticity of the granules and thus affecting the contact area of the inter-granular joints and, secondly, to affect the adhesiveness of the granules so that the strength of the inter-granular joints will be changed (e.g., through local deformation of the binder or through the formation of binder bridges between the granules). A complicating factor in understanding the role of the binder is that the failure may be localized in different ways during the breakage of a tablet formed from binder-substrate granules (100,101), i.e., binder–binder, binder– substrate and substrate–substrate. The spreading of the binder over the substrate particle surfaces and the interaction between binder and substrate will possibly affect the bonding between and breakage of granules (102).

Since choice of binder and final proportion of the binder in the formulation are traditionally important formulation factors for the mechanical strength of tablets, a large number of reports can be found in the literature dealing with the effect of binder and binder proportion on tablet strength (93,103–107). It seems reasonable that in many cases, the effect of these formulation factors on the mechanical strength of tablets is expressed through simultaneous effect on the plasticity and on the adhesiveness of the granules.

Compactibility Maps

In Figures 16 and 17, we have schematically summarized the discussions above on material properties that control the compactibility of particulate and granular solids. These compacibility maps indicate in a qualitative way the relationship between the dominant material properties and the tablet tensile strength.

REFERENCES

1. Courtney TH. Mechanical Behavior of Materials. New York: McGraw-Hill, 1990.
2. Hertzberg RW. Deformation and Fracture Mechanics of Engineering Materials, 4th ed. New York: John Wiley, 1996.
3. Rowe RC, Roberts RJ. Mechanical properties. In: Alderborn, G Nyström C, eds. Pharm Powder Compaction Technology. NY: Marcel Dekker, 1996:165.
4. European Pharmacopoeia 5.0. Strasbourg: The Directorate for the Quality of Medicines of the Council of Europe (EDQM), 2005.
5. United States Pharmacopoeia 28. Rockville, MD, USA: United States Pharmacopeial Convention, 2005.
6. Nyström C, Malmqvist K, Mazur J, Alex W, Hölzer AW. Measurement of axial and radial tensile strength of tablets and their relation to tablet capping. Acta Pharm Suec 1978; 15: 226–32.
7. Hwang R-C, Parrott EL. Rate of wear of pharmaceutical tablets. Drug Dev Ind Pharm 1993; 19:1317–29.
8. Duncan-Hewitt WC. Predicting the relative rate of wear of pharmaceutical compacts using the mechanical properties of their constituent crystals. Powder Technol 1990; 60:265–72.
9. Duncan-Hewitt WC, Grant DJW. The impact fracture wear test: A novel method of tablet evaluation. Powder Technol 1987; 52:17–28.
10. Rudnick A, Hunter AR, Holden FC. An analysis of the diametral-compression test. Mater Res Stands 1963; 3:283–9.
11. Fell JT, Newton JM. Determination of tablet strength by diametral compression test. J Pharm Sci 1970; 59:688–91.

12. Jungersen O. The use of tensile failure of compacts for evaluation of powder compressibility. Acta Pharm Suec 1976; 13:261–6.
13. Davies PN, Newton JM. Mechanical strength. In: Alderborn G Nyström C, eds. Pharmaceutical Powder Compaction Technology. New York: Marcel Dekker, 1996:165.
14. Rees JE, Hersey JA, Cole ET. The effect of rate of loading on the strength of tablets. J Pharm Pharmacol Suppl 1970; 22:64S–9S.
15. Sonnergaard JM, Jensen CG, Poulsen L, Lokin KB. Comparative investigations of tablet crushing force testers. Pharm Ind 2005; 67:108–15.
16. Rees JE, Rue PJ. Work required to cause failure of tablets in diametral compression. Drug Dev Ind Pharm 1978; 4:131–56.
17. Podczeck F. Particle-Particle Adhesion in Pharmaceutical Powder Handling. London: Imperial College Press, 1998; 176.
18. Stanley P, Newton JM. Variability in the strength of powder compacts. J. Powder Bulk Solids Technol 1977; 1:13–19.
19. Kennerley JW, Newton JM, Stanley P. Variability in the mechanical strength of tablets. Acta Pharm Technol Suppl 1979; 7:53–6.
20. Pitt KG, Newton JM, Richardson R, Stanley P. The material tensile strength of convex-faced Aspirin tablets. J Pharm Pharmacol 1989; 41:289–92.
21. Pitt KG, Newton JM, Stanley P. Effects of compaction variables on porosity and material tensile strength of convex-faced Aspirin tablets. J Pharm Pharmacol 1990; 42:219–25.
22. Newton JM, Alderborn G, Nyström C, Stanley P. The compressive to tensile strength ratio of pharmaceutical compacts. Int J Pharm 1993; 93:249–51.
23. David ST, Augsburger LL. Flexure test for determination of tablet tensile strength. J Pharm Sci 1974; 63:933–6.
24. Nyström C, Alex W, Malmqvist K. A new approach to tensile strength measurement of tablets. Acta Pharm Suec 1977; 14:317–20.
25. Jarosz PJ, Parrott EL. Factors influencing axial and radial tensile strengths of tablets. J Pharm Sci 1982; 71:607–14.
26. Malvern LE. Introduction to the Mechanics of a Continuous Medium. Englewood Cliffs, NJ: Prentice-Hall, 1969.
27. Timoshenko SP, Goodier JN. Theory of Elasticity, 3rd ed. New York: McGraw-Hill, 1970.
28. Den Hartog JP. Advanced Strength of Materials. New York: McGraw-Hill, 1952.
29. Wright PJF. Comments on an indirect tensile test on concrete cylinders. Mag Concrete Res 1955; 7:87–96.
30. Nyström C, Alderborn G, Duberg M, Karehill PG. Bonding surface area and bonding mechanism—Two important factors for the understanding of powder compactability. Drug Dev Ind Pharm 1993; 19:2143–96.
31. Rumpf H. The strength of granules and agglomerates. In: Knepper WA, ed. Agglomeration. New York: Interscience 1962; 379–418.
32. Rumpf H. Zur Theorie der Zugfestigkeit von Agglomeraten bei Kraftübertragung an Kontaktpunkten. Chem Ing Tech 1970; 42:538–40.
33. Smith WO, Foote PD, Busang PF. Packing of homogeneous spheres. Phys Rev 1929; 34: 1271–4.
34. Gross D, Seelig T. Fracture Mechanics: With an Introduction to Micromechanics, Mechanical Engineering Series. Berlin: Springer, 2006.
35. Janssen M, Zuidema J, Wanhill RJH. Fracture Mechanics, 2nd ed. London: Spon Press, 2004.
36. Kendall K. Agglomerate strength. Powder Metall 1988; 31:28–31.
37. Leuenberger H. The compressibility and compactibility of powder systems. Int J Pharm 1982; 12:41–55.
38. MacLeod HM. Compaction of ceramics. In: Stanley-Wood NG, ed. Enlargement and Compaction of Particulate Solids. London: Butterworths, 1983:259.
39. Sonnergaard JM. Quantification of the compactibility of pharmaceutical powders. Eur J Pharm Biopharm 2006; 63:270–7.

40. Shotton E, Ganderton D. The strength of compressed tablets. III. The relation of particle size, bonding and capping in tablets of sodium chloride, aspirin and hexamine. J Pharm Pharmacol Suppl 1961; 13:144T–152T.
41. Alderborn G, Nyström C. Studies on direct compression of tablets. III. The effect on tablet strength of changes in particle shape and texture obtained by milling. Acta Pharm Suec 1982; 19:147–56.
42. Stamm A, Mathis C. Verpressbarkeit von festen Hilfstoffen für Direkttablettierung. Acta Pharm Technol Suppl 1976; 1:7–16.
43. Ryshkewitch E. Compression strength of porous sintered alumina and zirconia. J Am Ceramic Soc 1953; 36:65–8.
44. Wu CY, Best SM, Bentham AC, Hancock BC, Bonfield W. A simple predictive model for the tensile strength of binary tablets. Eur J Pharm Sci 2005; 25:331–6.
45. Wu CY, Best SM, Bentham AC, Hancock BC, Bonfield W. Predicting the tensile strength of compacted multi-component mixtures of pharmaceutical powders. Pharm Res 2006; 23: 1898–905.
46. Leu R, Leuenberger H. The application of percolation theory to the compaction of pharmaceutical powders. Int J Pharm 1993; 90:213–9.
47. Castillo S, Villafuerte L. Compactibility of ternary mixtures of pharmaceutical powders. Pharm Acta Helv 1995; 70:329–37.
48. Higuchi T, Elowe LN, Busse LW. The physics of tablet compression V. Studies on aspirin, lactose, lactose-aspirin, and sulfadiazin tablets. J Am Pharm Assoc Sci Ed 1954; 43:685–9.
49. Newton JM, Grant DJW. The relation between the compaction pressure, porosity and tensile strength of compacted powders. Powder Technol 1974; 9:295–7.
50. Eriksson M, Alderborn G. The effect of particle fragmentation and deformation on the interparticulate bond formation process during powder compaction. Pharm Res 1995; 12:1031–9.
51. Alderborn G. A novel approach to derive a compression parameter indicating effective particle deformability. Pharm Dev Technol 2003; 8:367–77.
52. Hiestand EN. Rationale for and the measurement of tableting indices. In: Alderborn G, Nyström C, eds. Pharmaceutical Powder Compaction Technology New York: Marcel Dekker, 1996: 219.
53. Johansson B, Wikberg M, Ek R, Alderborn G. Compression behaviour and compactability of microcrystalline cellulose pellets in relationship to their pore structure and mechanical properties. Int J Pharm 1995; 117:57–73.
54. Santos H, Veiga F, Pina ME, Sousa JJ. Compaction, compression and drug release properties of diclofenac sodium and ibuprofen pellets comprising xanthan gum as a sustained release agent. Int J Pharm 2005; 295:15–27.
55. Benbow JJ. Mechanisms of compaction. In: Stanley-Wood, NG, ed. Enlargement and Compaction of Particulate Solids. London: Butterworths, 1983:168.
56. Roberts RJ, Rowe RC. The compaction of pharmaceutical and other model materials—a pragmatic approach. Chem Eng Sci 1987; 42:903–11.
57. Roberts R, Rowe RC, York P. The relationship between Young's modulus of elasticity of organic solids and their molecular structure. Powder Technol 1991; 65:139–46.
58. Roberts RJ, Rowe RC, York P. The relationship between indentation hardness of organic solids and their molecular structure. J Mater Sci 1994; 29:2289–96.
59. Vromans H, de Boer AH, Bolhuis GK, Lerk CF, Kussendrager KD, Bosch H. Studies on tableting properties of lactose. Part 2. Consolidation and compaction of different types of crystalline lactose. Pharm Weekblad Sci Ed 1985; 7:186–93.
60. de Boer AH, Vromans H, Lerk CF, Bolhuis GK, Kussendrager KD, Bosch H. Studies on tableting properties of lactose. Part III. The consolidation behaviour of sieve fractions of crystalline α-lactose monohydrate. Pharm Weekblad Sci Ed 1986; 8:145–50.
61. Leuenberger H, Bonny JD, Lerk CF, Vromans H. Relation between crushing strength and internal specific surface area of lactose compacts. Int J Pharm 1989; 52:91–100.
62. Sebhatu T, Ahlneck C, Alderborn G. The effect of moisture content on the compression and bond-formation properties of amorphous lactose particles. Int J Pharm 1997; 146:101–14.

63. Sebhatu T, Alderborn G. Relationships between the effective interparticulate contact area and the tensile strength of tablets of amorphous and crystalline lactose of varying particle size. Eur J Pharm Sci 1999; 8:235–42.
64. David ST, Augsburger LL. Plastic flow during compression of directly compressible fillers and its effect on tablet strength. J Pharm Sci 1977; 66:155–9.
65. de Boer AH, Bolhuis GK, Lerk CF. Bonding characteristics by scanning electron microscopy of powders mixed with magnesium stearate. Powder Technol 1978; 20:75–82.
66. Vromans H, Bolhuis GK, Lerk CF. Magnesium stearate susceptibility of directly compressible materials as an indication of fragmentation properties. Powder Technol 1988; 54:39–44.
67. Nyström C, Glazer M. Studies on direct compression of tablets. XIII. The effect of some dry binders on the tablet strength of compounds with different fragmentation propensity. Int J Pharm 1985; 23:255–63.
68. Kopp-Kubel S, Beyer C, Graf E, Kubel F, Doelker E. Einfluss der Polymorphie von Phenobarbital auf Tabletteneigenschaften. Eur J Pharm Biopharm 1992; 38:17–25.
69. Joiris E, Di Martino P, Berneron C, Guyot-Hermann AM, Guyot JC. Compression behaviour of orthorombic paracetamol. Pharm Res 1998; 15:1122–30.
70. Sun C, Grant DJW. Influence of crystal Structure on the tableting properties of Sulfamerazinepolymorphs. Pharm Res 2001; 18:274–80.
71. Sun C, Grant DJW. Compaction properties of L-lysine salts. Pharm Res 2001; 18:281–6.
72. Vromans H, Bolhuis GK, Lerk CF, van de Biggelaar H, Bosch H. Studies on tableting proerties of lactose. VII. The effect of variations in primary particle size and percentage of amorphous lactose in spray dried lactose products. Int J Pharm 1987; 35:29–37.
73. Sebhatu T, Elamin AA, Ahlneck C. Effect of moisture sorption on tabletting characteristics of spray dried (15% amorphous) lactose. Pharm Res 1994; 11:1233–8.
74. Shukla AJ, Price JC. Effect of moisture content on compression properties of two dextrose-based directly compressible diluents. Pharm Res 1991; 8:336–40.
75. Sun C, Grant DJW. Improved tableting properties of p-hydroxybenzoic acid by water of crystallization: A molecular insight. Pharm Res 2004; 21:382–6.
76. Sun C, Grant DJW. Effects of initial particle size on the tableting properties of L-lysine monohydrochloride dihydrate powder. Int J Pharm 2001; 215:221–8.
77. Alderborn G, Nyström C. Studies on direct compression of tablets. IV. The effect of particle size on the mechanical strength of tablets. Acta Pharm Suec 1982; 19:381–90.
78. Fichtner F, Rasmuson Å, Alderborn G. Particle size distribution and evolution in tablet structure during and after compaction. Int J Pharm 2005; 292:211–25.
79. Wong LW, Pilpel N. The effect of particle shape on the mechanical properties of powders. Int J Pharm 1990; 59:145–54.
80. Alderborn G, Börjesson E, Glazer M, Nyström C. Studies on direct compression of tablets. XIX. The effect of particle size and shape on the mechanical strength of sodium bicarbonate tablets. Acta Pharm Suec 1988; 25:31–40.
81. Karehill PG, Glazer M, Nyström C. Studies on direct compression of tablets. XXIII. The importance of surface roughness for the compactibility of some directly compressible materials with different bonding and volume reduction properties. Int J Pharm 1990; 64:35–43.
82. El Gindy NA, Samaha MW. Tensile strength of some pharmaceutical compacts and their relation to surface free energy. Int J Pharm 1982; 13:35–46.
83. Li Q, Rudolph V, Weigl B, Earl A. Interparticle van der Waals force in powder flowability and compactibility. Int J Pharm 2004; 280:77–93.
84. Sakr FM, Pilpel N. The tensile strength and consolidation of lactose coated with non-ionic surfactants. II. Tablets. Int J Pharm 1982; 10:57–65.
85. Berggren J, Frenning G, Alderborn G. Compression behaviour and tablet forming ability of spray-dried amorphous composite particles. Eur J Pharm Sci 2004; 22:191–200.
86. Millili GP, Schwartz JB. The strength of microcrystalline cellulose pellets: the effect of granulating with water/ethanol mixtures. Drug Dev Ind Pharm 1990; 16:1411–26.
87. Selkirk AB, Ganderton D. An investigation of the pore structure of tablets of sucrose and lactose by mercury porosimetry. J Pharm Pharmacol Suppl 1970; 22:79S–85S.

88. Wikberg M, Alderborn G. Compression characteristics of granulated materials. VI. Pore size distributions, assessed by mercury penetration, of compacts of two lactose granulations with different fragmentation propensities. Int J Pharm 1992; 84:191–5.
89. van der Zwan J, Siskens CAM. The compaction and mechanical properties of agglomerated materials. Powder Technol 1982; 33:43–54.
90. Johansson B, Nicklasson F, Alderborn G. Effect of pellet size on degree of deformation and densification during compression and on compactibility of microcrystalline cellulose pellets. Int J Pharm 1998; 163:35–48.
91. Nordström J, Welch K, Frenning G, Alderborn G. On the physical interpretation of the Kawakita and Adams parameters derived from confined compression of granular solids. Powder Technol 2008; 182: 424–35.
92. Nicklasson F, Johansson B, Alderborn G. Tabletting behaviour of pellets of a series of porosities—a comparison between pellets of two different compositions. Eur J Pharm Sci 1999; 8:11–7.
93. Alderborn G, Nyström C. Radial and axial tensile strength and strength variability of paracetamol tablets. Acta Pharm Suec 1984; 21:1–8.
94. Nicklasson F, Alderborn G. Compression shear strength and tabletting behaviour of microcrystalline cellulose agglomerates modulated by incorporation of a solution binder (polyethylene glycol). Pharm Res 2001; 18:873–7.
95. Johansson B, Alderborn G. The effect of shape and porosity on the compression behaviour and tablet forming ability of granular materials formed from microcrystalline cellulose. Eur J Pharm Biopharm 2001; 52:347–57.
96. Alderborn G, Lång PO, Sågström A, Kristensen A. Compression characteristics of granulated materials. I. Fragmentation propensity and compactibility of some granulations of a high dosage drug. Int J Pharm 1987; 37:155–61.
97. Rue PJ, Seager H, Ryder J, Burt I. The relationship between granule structure, process of manufacture and the tabletting properties of a granulated product. Part II. Compression properties of the granules. Int J Pharm Technol Prod Manuf 1980; 1:2–6.
98. Ragnarsson G, Sjögren J. Influence of the granulation method on bulk properties and tabletability of a high dosage drug. Int J Pharm 1982; 12:163–71.
99. Wikberg M, Alderborn G. Compression characteristics of granulated materials. VII. The effect of intra-granular binder distribution on the compactibility of some lactose granulations. Pharm Res 1993; 10:88–94.
100. Cutt T, Fell JT, Rue PJ, Spring MS. Granulation and compaction of a model system. I. Granule properties. Int J Pharm 1986; 33:81–7.
101. Mullier MA, Seville JPK, Adams MJ. A fracture mechanics approach to the breakage of particle agglomerates. Chem Eng Sci 1987; 42:667–77.
102. Rowe RC. Correlation between predicted spreading coefficients and measured granule and tablet properties in the granulation of paracetamol. Int J Pharm 1990; 58:209–13.
103. Armstrong NA, Morton FSS. The effect of granulating agents on the elasticity and plasticity of powders. J Powder Bulk Solids Technol 1977; 1:32–35.
104. Doelker E, Shotton E. The effect of some binding agents on the mechanical properties of granules and their compression characteristics. J Pharm Pharmacol 1977; 29:193–8.
105. Reading SJ, Spring MS. The effects of binder film characteristics on granule and tablet properties. J Pharm Pharmacol 1984; 36:421–6.
106. Krycer I, Pope DG, Hersey JA. An evaluation of binding agents. Part I. Solution binders. Powder Technol 1983; 34:39–51.
107. Zuurman K, Bolhuis GK, Vromans H. Effect of binder on the relationship between bulk density and compactibility of lactose granulations. Int J Pharm 1995; 119:65–9.

8

cGMPs for the 21st Century and ICH Quality Initiatives

Moheb M. Nasr, Donghao (Robert) Lu, and Chi-wan Chen

Office of New Drug Quality Assessment Center for Drug Evaluation and Research, U.S. Food and Drug Administration, Silver Spring, Maryland, U.S.A.*

INTRODUCTION

Recently, the Food and Drug Administration (FDA) has begun to implement the current Good Manufacturing Practice (cGMPs) for the 21st Century Initiative to further ensure the availability of high quality pharmaceutical products in the Unites States market. The initiative was first announced in 2002 and became clearly-defined in its final report published in September 2004 (1). The centerpiece of this initiative is to rely on science-based and risk-based approaches to FDA regulatory decision-making throughout the entire lifecycle of a product. The guiding principles for implementing this cGMPs initiative are outlined in Figure 1. Based on these principles, the quality of pharmaceutical products is established through an efficient utilization of modern pharmaceutical development, quality risk management, and quality systems. With the advances in science and engineering in the 21st century, the modern knowledge and information can be readily applied to improve the efficiency and effectiveness of both manufacturing process and regulatory actions. The implementation of the cGMPs initiative is also coordinated with other international regulatory authorities through the development of harmonized guidelines and strategies. These science-based and risk-based efforts can lead to the global implementation of a more efficient quality-assurance system for pharmaceutical manufacturing and regulatory oversight and thus provide the most effective public health protection.

Pharmaceutical tablet is the most common dosage form of drug products. It provides patients with a convenient means of handling and administration of drugs. Thus, tablet dosage forms account for a large percentage of the drug products approved to date. According to the FDA's approved drug database (via www.fda.gov/cder/), the number of pharmaceutical tablet products make up 43.7% of all approved drug products that are listed in the orange book (2007). The development and manufacturing of pharmaceutical tablets, including the conventional and the more advanced controlled-release tablets, have become more sophisticated in recent years. The general scientific principles and specific technological advances are well presented and described in details in the other chapters of

*The views expressed in this article are those of the authors and do not reflect the official policy of the FDA. No official support or endorsement by the FDA is intended or should be inferred.

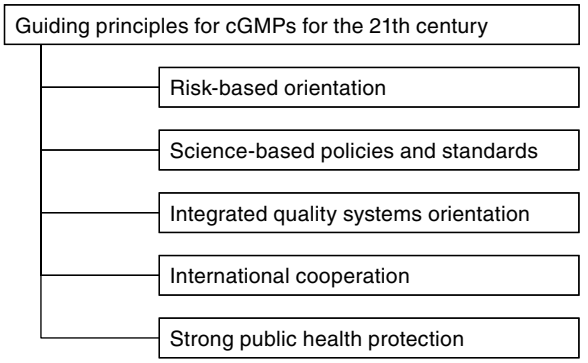


FIGURE 1 The guiding principles for implementing the cGMPs for the 21st century initiative.

this book. This chapter is intended (*i*) to provide an updated overview of regulatory implementation of the science-based and risk-based approaches to ensuring high quality drug products throughout product lifecycle, as laid out within the scope of cGMPs for the 21st Century Initiative, (*ii*) to present the newly established Pharmaceutical Quality Assessment System (PQAS) that manages the chemistry, manufacturing, and controls (CMC) review process of new drug products, including the tablet dosage forms, and (*iii*) to briefly describe the recent international harmonization efforts.

REGULATORY OBJECTIVES FOR CGMPS FOR THE 21st CENTURY

The cGMPs for the 21st Century Initiative has brought unprecedented challenges to both the pharmaceutical industry and the regulatory agency (FDA). To effectively develop and manufacture high quality drug products in the 21st century, pharmaceutical industry will need to move to the "desired state" (i.e., more efficient, agile, flexible operations that can reliably produce high quality drug products with less regulatory oversight) (2) for pharmaceutical manufacturing while FDA must utilize modern science-and risk-based approaches to regulatory decision-making. The cGMP initiative has clearly defined five regulatory objectives, as described in each of the following sections, respectively. These regulatory objectives, including innovation, quality system approaches, science-based and risk-based management, and consistent regulatory quality assessment, will guide both pharmaceutical industry and FDA in implementing necessary measures to assure the availability of high quality drug products in the United States market. To support these regulatory objectives, the Office of New Drug Quality Assessment (ONDQA) at FDA has developed a new PQAS to address the current regulatory challenges and to establish a modern regulatory system.

Encourage the Early Adoption of New Technological Advances by the Pharmaceutical Industry

Pharmaceutical development is rapidly evolving from an art to a science and engineering based endeavor. Drug delivery technology is advancing to a new era where innovative approaches are used in a significant number of drug products. The new drug delivery applications, including such areas as precisely-timed sustained release, self-regulated controlled-release, "intelligent" pharmaceutical polymers, cellular drug targeting, protein and gene delivery, and nanotechnology, will no doubt reshape the future pharmaceutical development and manufacturing. In fact, significant changes have already taken place in the currently marketed pharmaceutical products. For

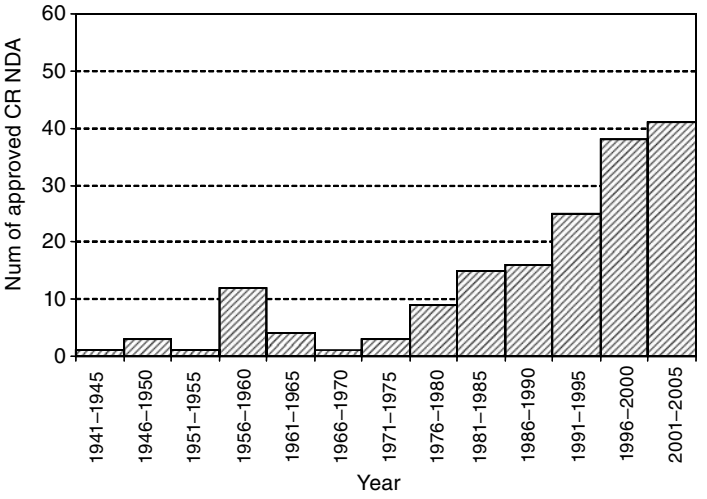


FIGURE 2 The number of approved controlled-release solid oral products in NDAs. *Abbreviation*: NDA, New Drug Application.

example, as indicated in the approved drug database (3), the number of approved controlled-release solid oral drug products has significantly increased in recent years, for both innovator drug products [submitted to FDA for evaluation as New Drug Application (NDA)] and generic drug products [submitted as Abbreviated New Drug Application (ANDA)]. Figure 2 shows the number of approved controlled-release solid oral products in NDAs and Figure 3 shows the number of approved controlled-release solid oral products together in NDAs and ANDAs, presented in a five-year increments. The data clearly illustrate the trend that a significant number of the new NDAs and ANDAs will have controlled-release solid oral dosage forms and the number will keep increasing as

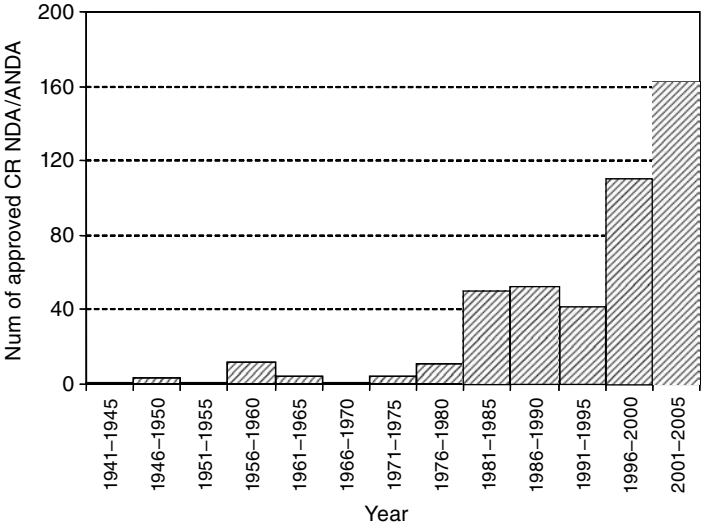


FIGURE 3 The number of approved controlled-release solid oral products in NDAs and ANDAs together. *Abbreviations*: NDA, New Drug Application; ANDA, Abbreviated New Drug Application.

new delivery technologies become more mature and more widely applied. Therefore, it is critical and timely for FDA to encourage pharmaceutical industry to become more innovative and to consider the early adoption of new technological and manufacturing platforms.

At present, the cGMPs for the 21st Century Initiative has already led to significant efforts at FDA to encourage innovation in the pharmaceutical industry. For example, Guidance for Industry Process Analytical Technology (PAT) —A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance (4) has presented a regulatory framework that encourages the voluntary development and implementation of innovative approaches to pharmaceutical development, manufacturing, and quality assurance. PAT is an innovative approach to pharmaceutical processing, defined as "a system for designing, analyzing, and controlling manufacturing through timely measurements (i.e., during processing) of critical quality and performance attributes of raw and in-process materials and processes, with the goal of ensuring final product quality". The PAT guidance provides a modern regulatory perspective and encourages the use of advanced technologies in pharmaceutical industry to improve efficiency and effectiveness of manufacturing process design, production, control, and quality assurance. The PAT regulatory framework covers two key components, the scientific principles and technology tools supporting manufacturing innovation as well as strategies for regulatory implementation hence, providing a proactive means to encourage innovation without perceived regulatory hurdles.

Facilitate Industry Application of Modern Quality Management Techniques, Including Implementation of Quality System Approaches, to all Aspects of Pharmaceutical Production and Quality Assurance

FDA has issued a Quality System Guidance in September, 2006 (5). The guidance states that "the overarching philosophy articulated in both the cGMP regulations and in robust modern quality systems is: quality should be built into the pharmaceutical product, and testing alone can not be relied on to ensure product quality". The concept of Quality by Design (QbD) is to design and develop a drug product and its manufacturing processes to ensure that the product consistently attains a predefined quality at the end of the manufacturing process. Based on the QbD concept, the implementation of modern and robust quality system approaches in pharmaceutical industry can ensure the production of high quality drug products and lead to the "desired state" of drug manufacturing.

The quality system model, described in the FDA guidance, lays out the operational framework that conforms to the cGMPs for the 21st Century Initiative and provides the necessary controls to consistently produce high quality drug products throughout the product lifecycle. There are four major components in the quality system model, as seen in Figure 4. Based on this model, the management responsibilities determine the overall success of the manufacturing operation. The responsibilities cover the entire operation, ranging from the planning, design, implementation, and overall management of the quality system, by providing active leadership and efficient organization structure, building a quality system suitable for the organization, establishing policies and objectives, and reviewing its adequacy and effectiveness. The proper allocation of resources, including personnel, facilities, equipment, and outsourced operations, plays a critical role in ensuring the robustness of the quality system. The manufacturing component in the quality system model effectively handles and controls

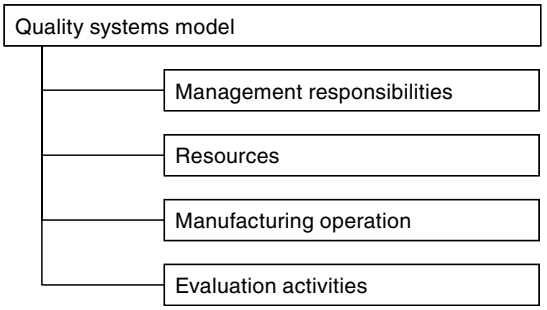


FIGURE 4 The quality systems model.

the product and process to meet the cGMP regulation requirements. The drug products should be well designed and developed. The corresponding manufacturing operations should be effectively performed and monitored. Any material that goes into a final product requires adequate qualification by thorough examination and its quality should be tested, audited, and controlled. If the system discovers nonconformities and deviations, appropriate modification capabilities should be established to handle the situation and to ensure the quality of the final product. The evaluation and correction capabilities, including data analysis for trends, internal audits, risk assessment, error correction, problem prevention, and system improvement, should be established within the quality system model. With the proper structural realization in above-mentioned management responsibility, resource, manufacturing operation, and evaluation activity, the quality system approaches can significantly enhance development and manufacturing processes in the pharmaceutical industry. It is expected that the implementation of quality systems, in combination with knowledge management from prior product design, manufacturing experience, and risk-based management practice, can deal with many types of changes and improvements to facilities, equipment, and processes without the need for prior approval regulatory submissions and can ensure consistency and high quality throughout the product lifecycle.

Encourage Implementation of Risk-Based Approaches that Focus both Industry and Agency Attention on Critical Areas

Quality risk management approaches to drug product consist of a systematic process for assessment, control, communication, and review of associated risks at various stages of the product lifecycle. For pharmaceutical industry, implementation of quality risk management approaches can ensure the consistent production of high quality products by providing a proactive means to identify, isolate, and eliminate potential risks to quality during product development and manufacturing. Risk-based management is an effective tool to identify critical process parameters and to facilitate the establishment of product specification and proposed design space, prior to the submission of drug applications to FDA. The cGMPs for the 21st Century Initiative emphasizes the maintenance of high product quality throughout the product lifecycle. The identification, scientific understanding, risk assessment, and subsequent control management of critical product quality attributes are the key to ensuring the long-term quality of the drug products. More detailed information on risk-based management approaches can be found in the International Conference on Harmonization (ICH) Q9 Guideline (6).

Risk-based management approaches to drug product quality are also important to the FDA regulatory decision-making process. In September 2004, the Office of New Drug Chemistry (ONDC) at FDA published a white paper on a new risk-based PQAS for the regulatory review of the CMC section of NDAs (7). The white paper and the subsequent reorganization and staff realignment of ONDC into the ONDQA established a new regulatory paradigm which uses the new PQAS approach and emphasizes risk-based CMC evaluation. The CMC review of an NDA will focus more on the critical quality attributes and their relevance to safety and efficacy. Based on the product knowledge and process understanding demonstrated during pharmaceutical development and submitted in the application, the regulatory assessment at ONDQA uses a risk-based approach, relying on the degree of the understanding of drug substance, drug product, pharmaceutical development, and manufacturing process. Risk-based CMC assessment is an integral component of the GMPs for the 21st Century Initiative and can greatly enhance the effectiveness of regulatory decisions.

Ensure that Regulatory Review, Compliance, and Inspection Policies are Based on State-of-the-Art Pharmaceutical Science

In the 21st century, pharmaceutical sciences have evolved into a multi-disciplinary field covering basic science principles as well as practical technology and engineering development. To ensure high drug product quality, the modern pharmaceutical sciences should be used as the foundation in establishing the regulatory review, compliance, and inspection policies, and conducting day-to-day regulatory business, both in the pharmaceutical industry and in the government agency. FDA has published a series of guidances (http://www.fda.gov/cder/guidance/index.htm) based on modern pharmaceutical science principles to establish the cGMP regulatory requirements and to provide recommendations on the CMC information for the drug substance and product that should be submitted in an NDA. The guidances and other regulatory review, compliance, and inspection policies also provide the necessary scientific justifications for the regulatory actions that are generated after the review process at FDA.

As stated in the PAT guidance (3), "Quality is built into pharmaceutical products through a comprehensive understanding of: (*i*) the intended therapeutic objectives; patient population; route of administration; and pharmacological, toxicological, and pharmacokinetic characteristics of a drug, (*ii*) the chemical, physical, and biopharmaceutic characteristics of a drug, (*iii*) design of a product and selection of product components and packaging based on drug attributes listed above, (*iv*) the design of manufacturing processes using principles of engineering, material science, and quality assurance to ensure acceptable and reproducible product quality and performance throughout a product's shelf life." For quality assurance in each of these areas, Guidance for Industry are provided by FDA, ranging from stability testing to specification establishment, for drug substances and drug products, including the tablet products. Examples include Q1A (R2) "Stability testing of new drug substances and products", Q3A(R)/Q3B(R) "Impurities in new drug substances/products", and Q6A "Specifications: test procedures and acceptance criteria for new drug substances and new drug products". Under the cGMPs for the 21st Century Initiative, ICH guidances Q8, Q9, and Q10 are intended to address the new directions in the regulatory review, compliance, and inspection policies, and they will be further discussed in the following sections. The complete list of the ICH Guidelines can be seen in Table 1.

TABLE 1 Currently Available ICH-Quality Guidances

Title and format	Type	Issue date
Q1A(R2) Stability Testing of New Drug Substances and Products	Final	11/2003
Q1B Photostability Testing of New Drug Substances and Products	Final	11/1996
Q1C Stability Testing for New Dosage Forms	Final	5/1997
Q1D Bracketing and Matrixing Designs for Stability Testing of New Drug Substances and Products	Final	1/2003
Q1E Evaluation of Stability Data	Final	6/2004
Q2A Text on Validation of Analytical Procedures	Final	3/1995
Q2B Validation of Analytical Procedures: Methodology	Final	5/1997
Q3A(R) Impurities in New Drug Substances	Final	2/2003
Q3B(R) Impurities in New Drug Products	Final	8/2006
Q3C Impurities: Residual Solvents	Final	12/1997
Q4B Regulatory Acceptance of Analytical Procedures and/or Acceptance Criteria (RAAPAC)	Draft	8/2006
Q5A Viral Safety Evaluation of Biotechnology Products Derived From Cell Lines of Human or Animal Origin	Final	9/1998
Q5B Quality of Biotechnological Products: Analysis of the Expression Construct in Cells Used for Production of *r*-DNA Derived Protein Products	Final	2/1996
Q5C Quality of Biotechnological Products: Stability Testing of Biotechnological/Biological Products	Final	7/1996
Q5D Quality of Biotechnological/Biological Products: Derivation and Characterization of Cell Substrates Used for Production of Biotechnological/Biological Products; Availability	Final	9/1998
Q5E Comparability of Biotechnological/Biological Products Subject to Changes in Their Manufacturing Process	Final	6/2005
Q6A Specifications: Test Procedures and Acceptance Criteria for New Drug Substances and New Drug Products: Chemical Substances	Final	12/2000
Q6B Specifications: Test Procedures and Acceptance Criteria for Biotechnological/Biological Products	Final	8/1999
Q7A Good Manufacturing Practice Guidance for Active Pharmaceutical Ingredients	Final	8/2001
Q8 Pharmaceutical Development	Final	5/2006
Q9 Quality Risk Management	Final	6/2006
Q10 Pharmaceutical Quality System	Draft	7/2007

Enhance the Consistency and Coordination of FDA's Drug Quality Regulatory Programs, in Part, by Further Integrating Enhanced Quality Systems Approaches into the Agency's Business Processes and Regulatory Policies Concerning Review and Inspection Activities

An important implementation of the cGMPs for the 21st Century Initiative is to establish consistent regulatory quality assessment of drug applications. To achieve this goal, a new PQAS was developed in September 2004 (7). PQAS supports science-based and risk-based regulatory approaches to pharmaceutical products in ensuring the quality throughout the product lifecycle. The new system promotes the following four regulatory assessment objectives: (*i*) to emphasize submissions rich in scientific information demonstrating product knowledge and process understanding, (*ii*) to focus on critical pharmaceutical quality attributes and their relevance to safety and effectiveness, (*iii*) to enable

FDA to provide regulatory flexibility for specification setting and post-approval changes based on demonstrated product and manufacturing process understanding, and (*iv*) to facilitate innovation and continual improvement throughout product lifecycle.

In coordination with the PQAS implementation, FDA's organizational structure for CMC review at the ONDC was rearranged into a new organization, the ONDQA, intended to be more efficient, effective and flexible in managing CMC review processes and internal workload. Significant changes were made in ONDQA, including (*i*) creation of a dedicated postmarketing division for CMC evaluation of NDA supplements; (*ii*) establishment of Pharmaceutical Assessment Lead positions to perform initial quality assessment and to serve as liaisons to FDA clinical divisions; (*iii*) development of assessment branches (including a new manufacturing branch), responsible for the quality evaluation of various therapeutic areas with specialized review expertise; (*iv*) integration of biopharmaceutics evaluation into the quality assessment process; and (*v*) addition of project management staff to streamline the assessment operation and to enhance the integration of CMC review with clinical review and pre-approval GMP inspection. The new ONDQA operational structure has proven to be effective in dealing with the rising number of NDA applications and supplements, as well as the increasing complexity of new drug products.

PQAS integrates enhanced quality system approaches into the CMC review processes and applies the risk-based management principles to regulatory decision-making. It focuses on critical pharmaceutical quality attributes and their relevance to safety and efficacy. The critical pharmaceutical quality attributes (chemistry, pharmaceutical formulation, manufacturing process, and product performance) are the product properties that can significantly influence the intended clinical outcomes if certain degree of variation is encountered. Risk-based assessment approaches are used in PQAS to identify these critical quality attributes and the potential sources for the variations and subsequently to ensure necessary controls being established in the manufacturing process. PQAS places more emphasis on the pharmaceutical development report, included in section 3.2.P.2 (Pharmaceutical Development) of an NDA based on the Common Technical Document (ICH topic M4) format, to achieve an overall scientific and technical understanding on product development and manufacturing process. The new system promotes active collaborations and shared responsibilities between ONDQA, Office of Regulatory Affairs and CDER's Office of Compliance in pre-approval and GMP inspections. Refinement of PQAS in conjunction with the full implementation of the QbD with a strong focus on manufacturing science, integration of review and inspection functions, and use of modern statistical methodologies, will ensure high quality throughout the product lifecycle.

INTERNATIONAL CONFERENCE ON HARMONIZATION

Establishment of a globally harmonized approach to drug development and regulatory assessment is an important task as the pharmaceutical sciences and drug manufacturing become more modernized in the 21st century. The ICH of Technical Requirements for Registration of Pharmaceuticals for Human Use has a long history in developing guidelines for pharmaceutical industry to consistently establish the quality of new drug substances and products in the European Union, Japan, and the United States. ICH has established guidelines Q8, Q9, and a draft Q10 to address the pharmaceutical development, quality risk management, and pharmaceutical quality systems, respectively.

Pharmaceutical Development (Q8)

ICH Guidance, Q8 Pharmaceutical Development, was officially published by FDA in the United Statea in May 2006 (8). Q8 specifically addresses the pharmaceutical development section (3.2.P.2, or the P2 section) in the NDAs. The guidance was developed based on the concept that quality cannot be tested into products and quality should be built in by design in the pharmaceutical products. The key aspect is the comprehensive understanding and enhanced knowledge established by applicants for the product development and manufacturing process. The general contents in the P2 section consist of (*i*) components of the drug product (physicochemical and biological properties of drug substance and formulation excipients), (*ii*) drug product (formulation development and identification of critical quality attributes), (*iii*) manufacturing process development (process development and validation, critical process parameters, and control strategies), and (*iv*) other components including container closure system, microbiological attributes, and compatibility of the drug product with reconstitution diluents. A design space can also be proposed that is established based on the scientific understanding and enhanced knowledge from the pharmaceutical development studies and manufacturing experience. Risk-based assessment can assist pharmaceutical development and the establishment of the design space. As defined in the guideline, the design space describes the multi-dimensional combination and interaction of input variables (e.g., material attributes) and process parameters that have been demonstrated to provide assurance of quality. The pharmaceutical development studies should be systemically designed to lead to an enhanced knowledge of product performance over a wider range of formulation attributes, material characteristics, process parameters, and control strategies. The information presented in the Pharmaceutical Development section provides an opportunity to demonstrate a higher degree of understanding of the product and process, and to facilitate regulatory decision-making through the quality risk management approaches.

One of the most significant aspects of Q8 is to lay out the principles in flexible regulatory approaches. Based on the knowledge gained from the comprehensive pharmaceutical development studies as well as the prior knowledge and enhanced understanding of product performance over a range of material attributes, manufacturing process options, and process parameters, flexible regulatory approaches will be available to facilitate regulatory risk-based decisions, continual manufacturing process improvements, reduction of post-approval submissions, and real-time manufacturing quality control.

Quality Risk Management (Q9)

ICH Guidance Q9 Quality Risk Management, was officially published by FDA in the United States in June 2006 (6). Q9 lays out the quality risk management principles for pharmaceutical industry and regulatory agency, and provides a systematic approach to quality risk management of pharmaceutical products. In consistence with the primary principles of quality risk management that include "(*i*) the evaluation of the risk to quality should be based on scientific knowledge and ultimately link to the protection of the patient; and (*ii*) the level of effort, formality, and documentation of the quality risk management process should be commensurate with the level of risk", the drug development, manufacturing and regulatory actions can be evaluated with a risk-based as well as science-based assessment to ensure high product quality. The quality risk management approach can provide the assurance of product quality, define the confidence on

industry's ability to deal with potential issues, and facilitate the regulatory decisions based on sufficient understanding of the product and process.

The general quality risk management process consists of (*i*) responsibilities, (*ii*) initiating a quality risk management process, (*iii*) risk assessment, (*iv*) risk control, (*v*) risk communication, and (*vi*) risk review. The overall relationship among all elements of the quality risk management process is illustrated in a diagram in Q9, as seen in Figure 5. It is important to point out that effective risk communication is a key element that links every stage of the risk management process. The risk management responsibilities are usually realized through a team of multi-disciplinary experts in different areas and at different stages of drug development and, therefore, requiring effective coordination among operational units. Risk identification, risk analysis, and risk evaluation are the components for the quality risk assessment element that usually focuses on a well-defined problem description or risk question. An adequate risk assessment can lead to an effective risk control (through either the risk reduction procedure or risk acceptance procedure) to maintain the quality of drug products. It is noted that risk review should be routinely conducted on the overall risk management process during manufacturing in order to incorporate the newly gained knowledge and experience. It is essential to recognize that the quality risk management is a process that supports science-based decisions as well as practical decisions during the regulatory evaluation. Drug applications rich in scientific knowledge and risk management information on manufacturing process can greatly facilitate the regulatory decision-making at FDA.

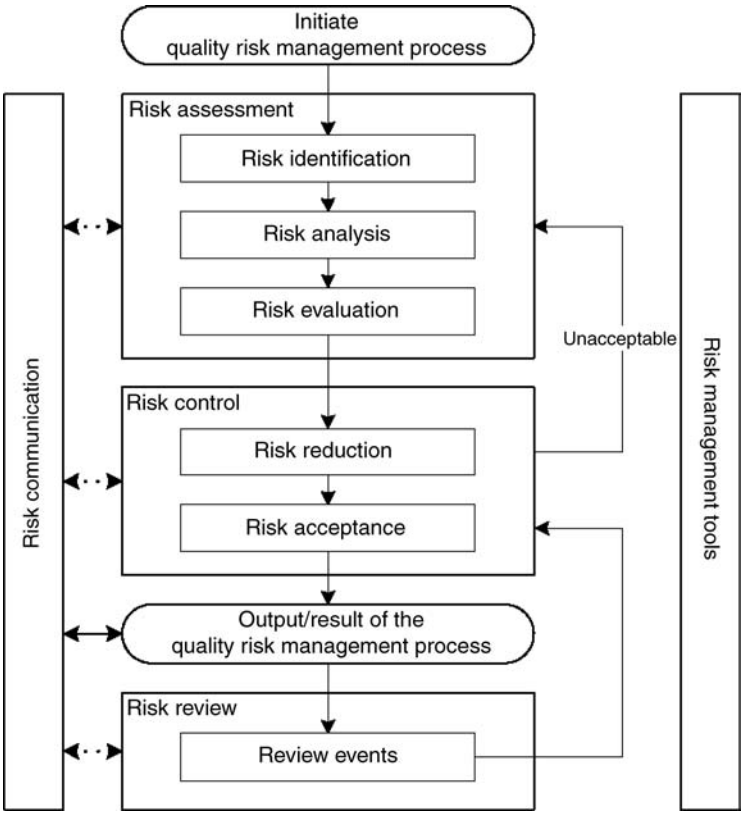


FIGURE 5 The overview of a typical quality risk management process. *Source*: From Ref. 6.

It is recognized that pharmaceutical industry and the regulatory agency can also assess and manage risk through the use of other risk management tools and internal procedures. A non-exhaustive list of some of the tools is shown in Table 2. In addition, informal risk management processes, such as empirical management tools, can be considered acceptable for use when adequate justifications are provided. However, the guidance has indicated that appropriate use of quality risk management can facilitate, but does not obviate industry's obligation to comply with regulatory requirements. Quality risk management does not replace appropriate communications between the applicant and regulator.

TABLE 2 Other Recognized Risk Management Tools

Tool name
Basic risk management facilitation methods (flowcharts, check sheets, etc.)
Failure Mode Effects Analysis (FMEA)
Failure Mode, Effects, and Criticality Analysis (FMECA)
Fault Tree Analysis (FTA)
Hazard Analysis and Critical Control Points (HACCP)
Hazard Operability Analysis (HAZOP)
Preliminary Hazard Analysis (PHA)
Risk ranking and filtering
Supporting statistical tools

Pharmaceutical Quality Systems (Q10)

ICH Guidance Q10 Pharmaceutical Quality System (draft), was published by FDA in the United States in July 2007 (9). Q10 presents a model for an effective quality management system for the pharmaceutical industry in order to achieve high quality throughout the product lifecycle. The overall objectives of Q10 are (*i*) to achieve product realization by establishing the well-defined product quality attributes, (*ii*) to establish and maintain a state of control by implementing effective process controls and quality assurance, and (*iii*) to facilitate continual improvement by promoting variability reduction, product innovations, and pharmaceutical quality system enhancements. The maintenance of high quality within a product lifecycle can be achieved on the basis of Q8 and Q9, i.e., from the pharmaceutical development knowledge and quality risk management. The regional GMP requirements, ICH Q7 Guidance and ISO Guidelines also serve as the foundation for Q10 pharmaceutical quality system.

The pharmaceutical product lifecycle involves many stages ranging from the product development to its discontinuation procedures. The general pharmaceutical product lifecycle can be summarized as shown in Figure 6. At pharmaceutical development stage, it is important to follow the ICH Q8 guidance and to adequately design and build the new drug products with desired quality attributes and intended clinical performance. At the technology transfer stage, the knowledge gained from the pharmaceutical development and from the subsequent manufacturing processes is properly shared among various operational units in the company to provide consistent understanding on the product and process. At the manufacturing stage, adequate controls and process improvement should be promoted to ensure high quality products. At the product discontinuation stage, appropriate documentation is critical to adequately managing the product termination procedures.

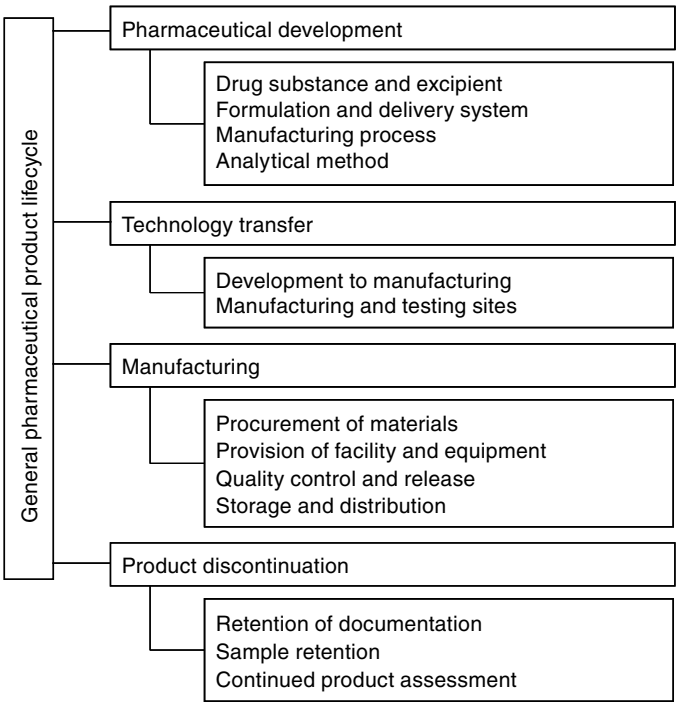


FIGURE 6 The general pharmaceutical product lifecycle.

Because it emphasizes the product quality lifecycle, Q10 defines the four pharmaceutical quality system elements for continual improvement of product and process: (*i*) process performance and product quality monitoring system, (*ii*) corrective action and preventive action system, (*iii*) change management system, and (*iv*) management review of process performance and product quality. The key components in the process performance element is the establishment of an effective monitoring and controlling procedure and the use of risk-based management approaches to maintaining high product quality within each stage of the product lifecycle. Subsequently, the ability for corrective actions and preventive actions in a timely manner is needed once product quality shows any defect during investigations. The continual improvement also requires an appropriate change management system for evaluation, approval, and implementation of any potential improvements. Finally, the management reviews of regulatory assessments, product quality controls, and overall effect of the continual improvements is another key element to ensuring the quality throughout the product lifecycle. Q10 emphasizes the importance of management leadership in implementation of the pharmaceutical quality system. The management commitment on quality, quality policy establishment within the organization, quality objectives and planning, resource management, internal communication, periodic system-wide review, and outsourcing oversight are critical management components within the quality system. The successful implementation of the pharmaceutical quality system, as outlined in Q10, step 2 document, can effectively maintain the product quality throughout its lifecycle by facilitating innovation, advancing new technology, and promoting continual process improvement.

REFERENCES

1. FDA Pharmaceutical cGMPs for the 21st century—a risk-based approach. Final report, 2004.
2. Woodcock J. Workshop on pharmaceutical quality assessment—A science and risk-based CMC approach in the 21st century. October 2005.
3. Drugs@FDA Data Files (May, 2007): http://www.fda.gov/cder/drugsatfda/datafiles/drugsatfda.zip (The zip file can also be found through http//www.fda.gov/cder/drugsatfda/datafiles/default.htm).
4. FDA Guidance for Industry. PAT—A framework for innovative pharmaceutical manufacturing and quality assurance. September 2004.
5. FDA Guidance for Industry. Quality systems approach to pharmaceutical cGMP regulations. September 2006.
6. FDA Guidance for Industry. Q9 Quality risk management. June 2006.
7. FDA White Paper. ONDC's new risk-based pharmaceutical quality assessment system. September 2004.
8. FDA Guidance for Industry. Q8 Pharmaceutical development. May 2006.
9. FDA Guidance for Industry. Q10 Pharmaceutical quality system. Draft, July 2007.

9

Intellectual Property, Patent, and Patenting Process in the Pharmaceutical Industry

Keith K. H. Chan
University of Maryland, Baltimore, Maryland, U.S.A.

Albert W. K. Chan
Law Offices of Albert Wai-Kit Chan, PLLC, New York, New York, U.S.A.

INTRODUCTION

The 21st century was termed as the century of knowledge. However, merely having the knowledge is not enough. It is the protection of that knowledge and conversion of that knowledge into profit which are important for the survival of any high-tech business and economy. The one who controls the knowledge and knows how to protect it is the winner in modern-day industry. The pharmaceutical industry, like any other high-tech industry, is no different. The company that has the upper hand will be the winner of the war. The stakes are high, and success or failure can make or break a company. The life blood of the pharmaceutical industry is innovative ideas and new products. It is clear that research productivity has gradually declined over the last few decades, and the cost to bring a new drug candidate to market has skyrocketed to an estimated whopping US$800 million or more (1). How one can create new ideas and products at the proper time and protect the life of current drug products has coined the term "Life Cycle Management (LCM)" in pharmaceutical industry (2). The whole objective of pharmaceutical drug product LCM is to maximize the profit of any drug product from start to market withdrawal and take full advantage of the intellectual rights and food and drug laws and regulations. This is extremely important for the survival of all pharmaceutical companies; no matter if it is a huge multinational company, a medium-size company, a one drug wonder company, a start-up company, or even a generic company. LCM is used as offensive or defensive tools to act and counteract against real or potential future competitors. The one who controls the knowledge and the know-how to develop and protect them is the sole qualified player in modern-day industry.

Intellectual property (IP) laws and the food and drug laws provide the pharmaceutical and biotechnology industry with unparalleled protection. For example, these laws provide exclusivity, patent term restoration, and patent extension under various conditions unmatched by any other industry. It is not the objective of this chapter to explain all facets of exclusivity and protection. The interested reader should conduct further research and seek appropriate professional advice. Rather, our assignment and

objective is to introduce patents and the patenting process commonly used in the pharmaceutical industry.

It is the authors' experience that most pharmaceutical LCM teams consist of three major types of professionals: (*i*) scientists of various disciplines, such as chemists, pharmacologists, formulation, and regulatory scientists, etc.; (*ii*) legal professionals, such corporate lawyers, patent lawyers, food and drug lawyers, litigation lawyers, etc.; and (*iii*) upper management, such as senior managers. The biotech and pharmaceutical business is really a "business of science." The success of business is totally dependent on the ability of upper management (i.e., leaders and managers or the management team) to convert an idea into a marketable product. The remaining essential elements and talents, such as scientific know-how, technology know-how, financial know-how, product development know-how, legal protection know-how, legal agreement construction know-how, management know-how, regulatory know-how, marketing and sales know-how, etc., can all be recruited or otherwise obtained. It is the authors' opinion that the biotech/pharmaceutical industry requires such skills in order to survive.

There are four major types of IP, namely, trade secrets, copyrights, trademarks, and patents (3). The pharmaceutical industry relies on all four types of IP protection, but patent protection is considered by far the most important and frequently used by pharmaceutical scientists.

It is the experience of the authors that most scientists are unfamiliar with the laws and the lawyers are unfamiliar with the cutting-edge of a specific technology. In order to function as a team and exert the maximum function, all team members must act in sync and at least have a working knowledge of each other's roles. Therefore, it is the objective of this chapter to provide the necessary working knowledge to deal with legal professionals. All patents start with science or, more specifically, an innovative scientific idea. However, the patent filing is a race against time, and balancing the perfection of science, which may take a long time to achieve, and the urge to file a patent application as soon as possible without substantial or definitive evidence due to fierce competition. Scientists are trained as perfectionists when it comes to generating new knowledge, but often are poor lawyers and businessmen. How to balance all concerns and accomplish the goals within the right time frame in the proper manner has made the patent filing process an art form. Hopefully the information provided in this chapter will reach beyond basic patent principle and normal patent practice in biotechnology and pharmaceutical industry. Specifically we would like to accomplish the following goals in this chapter:

1. IP fundamentals (trade secrets, trademarks, copyrights, and patents).
2. Fundamentals of patent concepts and the patenting process (patentability requirements, novelty and nonobviousness, enablement, written description, inventorship determination, different routes for filing and protection, i.e., provisional patent, patent cooperation treaty (PCT), direct national filings, cost and timing considerations, correct implementation and timeline, normal biotech/pharmaceutical patent practice, the right number of patents to pursue, etc.).
3. Patent due diligence process, patentability evaluation, concepts of freedom-to-operate, etc.
4. How to obtain local and international IP protection and how to protect your valuable technology/product.
5. The rationale for acquiring protection in specific countries, including when and how to seek protection and cost-and-benefit analysis.
6. Examples of pharmaceutical technology patents.

INTELLECTUAL PROPERTY FUNDAMENTALS (TRADE SECRETS, TRADEMARKS, COPYRIGHTS, AND PATENTS)

IP provides protection for ideas, designs and forms of expression which promote the advancement of science and technology. It is a form of intangible asset. IP includes trade secrets, trademarks, copyrights, patents, know-how and show-how. It requires lots of time. The protection starts with government process and is regulated by statutory laws. The following is a discussion of some of the specific areas of IP and their relationship to the pharmaceutical industry:

Trade Secrets

A trade secret is something that offers an advantage in business if kept as a secret (4). A trade secret can be a client list, the formula for a product, etc. A trade secret does not have to be patentable, but it must be capable of being maintained. For instance, a client list can be protected by a computer password, and a formula can be safeguarded by disclosing it only to a limited number of people.

Trade secrets are not registered with any government or any other agency. In fact, great pains are taken to prevent their disclosure. In contrast, patent protection requires disclosure.

Decisions are needed to be made for a patentable invention be held as a trade secret instead of a patent. Below are a few important questions to ask when making the decision to maintain an invention as a trade secret or disclose it as part of a patent application.

1. Can the patented invention be reasonably policed? If your invention is directed to products which are easily policed, a patent application may give you good protection. If your invention is a process which is difficult to police, a trade secret may be your only option.
2. Can the patented invention be easily circumvented? If yes, a patent will not give you the power to prevent others from entering the field, and you may not want to invest the time, effort, and money to obtain a patent.
3. Is the life of the patented invention relatively short? This is true for computer software, which is protected for only two-to-five years by a patent. Software developers might get better protection if they keep their inventions trade secrets rather than patenting them.
4. Does patent disclosure give competitors an edge? In other words, if a competitor knows the secret behind your invention, can the competitor generate the same product or a better one faster than you? This is sometimes true if the patentee is an independent inventor or has only a small company. Larger companies can easily upstage smaller ones using their plentiful personnel, expensive equipment, and broad resources.
5. Does the inventor want or need to publish the invention? Inventors who work in academia operate under the Publish-or-Perish Rule: If you don't publish papers, your career perishes. If this applies to you, a trade secret may be impractical. You may be pressured to disclose your invention because it is part of the work you are doing. Scientists who work in an active area of research, such as AIDS or Alzheimer's will find it especially difficult to maintain a trade secret. For these inventors, it is usually more advantageous to seek patent protection.
6. Will it be difficult to maintain the trade secret? Some inventions are created to be viewed publicly. A method for packaging, is an example of this. If this is the case, it will be impossible to keep such an invention a trade secret. As soon as it is on the market, it will lose its status as a secret. A patent would be advisable here. Alternatively, some inventions are easy to keep a secret. Coca-Cola has

maintained the formula of Coca-Cola as trade secret for a long time. Only two people on earth have access to the formula, which is locked in a safety box. If Coca-Cola files a patent application, it will disclose the formula and can only enjoy the legitimate patent term. After that, everyone would be able to copy it. That is why things like secret formulas and recipes are maintained as trade secrets and not as patents.

While many inventions must be patented in order to be protected, there are many inventions that do not require patenting to serve their inventors well. There are several distinct advantages to trade secret protection if your invention qualifies.

1. The expenses involved with obtaining patent protection and enforcing patent rights are not encountered when trade secret protection is used. The only costs involved in keeping a secret are administrative.
2. There is no time limit on trade secret protection.
3. Competitors are not apprised of the trade secret, compared to the full disclosure required for a patented invention.
4. Competitors are unable to practice the trade secret invention without a specific microbe or clone. Patent law in most countries mandates that patentees make available specific microbes or clones.
5. A trade secret does not have to be a patentable invention; it must be simply unique and secret.

In fact, in some countries, there is administrative protection for some "secret" formulas.

Trademarks

Trademark law protects symbols which are used on goods and on services (5). The symbol must be affixed onto the product or used with the service. Trademark law protects the trademark owner and prevents consumer confusion. Most consumers will rely on the labels attached to the product with a certain expectation of the quality of said product. There is no specific term for a trademark as long as it is in use. The notation ® may be used for the trademark only if it is federally registered. In the pharmaceutical arena, trade names for certain drug may be registered as a trademark.

Copyrights

Copyright protects forms of expression of original works. Copyright law protects the publications of the studies. Information provided by the drug companies may be protected by copyright law. Pharmaceutical companies routinely copyrighted their package insert yet the generic approval dictated that the package insert (including user guide and brochure) of generic drug to be the "same" as the reference listed drug. This apparent conflict of between drug approval under Federal Food Drug and Cosmetic Acts and the Copyright Law has been resolved in a court case [SKF versus Watson, 211 F.3d 21 (2d Cir. 2000)].

FUNDAMENTALS OF PATENT CONCEPTS AND THE PATENTING PROCESS

Patentability and Freedom-to-Operate

Patent protection is, perhaps, the most important IP protection in the pharmaceutical industry (6). Fundamentally, patent is a legal right to stop others from making, using, offering for sale or selling an invention, or importing a product made by a patented invention. Therefore, a patent is essentially preventing others from using or infringing the

invention. However, it did not guarantee the invention can be marketed especially when the product being marketed may require other technologies covered by other inventions. Patentability evaluation and freedom-to-operation evaluation are kind of separate concepts but complementary. Patentability is to determine whether the invention can qualify for patent application or not whereas freedom-to-operate is to determine if the possibility of the invention will infringe on other inventions. Both patentability and freedom-to-operate evaluation should be performed by qualified professionals.

What is Patentable?

An invention must fulfill four basic requirements before it can be deemed patentable. They are: novelty, utility, nonobviousness, and written disclosure. These four elements must be proven within the patent application.

Novelty

The invention seeking protection must be new. Usually the inventor already knows whether or not this is the case. Before investing in filing costs, attorney fees, and licensing efforts, it may be to your advantage to perform a complete patent search. The goal of performing a search is to ensure that the invention is original. A complete search includes both literature, patent and prior art (7) searches. Just like any results to be published in top tier journals, the data must be new. A thorough patent search would also be important to determine if the invention is new. A patent search includes world patents as well as U.S. patent applications. In most countries it is mandatory for patent applications to be published 18 months after filing. (e.g., http://www.uspto.gov). If it is an important invention, one may wish to hire search companies to perform the prior art searches. The cost of doing a search is dependent upon the level of certainty one wishes to attain. Searching will show you whether the invention fulfills the novelty requirement.

Utility

An invention must be useful for it to be patentable. Usefulness in the research sense, however, is insufficient; the invention must have some commercial application. For example, if one discovers a gene which is important for neurodevelopment, the assertion that this gene is then useful for studying neurodevelopment is insufficient for fulfilling the utility requirement. Using this example, the gene fulfills the utility requirement if its expression is indicative of a particular neurodisease.

Nonobviousness (Inventive Step)

The most common hurdle on the road to obtaining a biomedical patent is fulfilling the criteria for nonobviousness. The invention is judged for its obviousness in light of the level of skill in the art. In other words, obviousness is evaluated from the viewpoint of an ordinary person practicing in the same field as the inventor.

It is no secret that the standard for nonobviousness varies from patent examiner to patent examiner (those people at the Patent and Trademark Office (PTO) who are responsible for allowing or rejecting a patent). The level of ordinary skill in the art must be ascertained by a patent examiner. He/she then compares the claimed invention with the level of ordinary skill to judge whether your invention is obvious. In a patent application, "claims" define the legal rights which belong to the inventor (applicant).

Examiners review references to help them prove that an invention is obvious and, therefore, not patentable. References include any prior art, such as literature, scientific papers, advertised papers, oral presentations, public knowledge, etc., on an invention

released prior to the filing date of the application. Routinely, examiners cite a primary reference along with secondary references in order to prove that a claimed invention is obvious. These citations, ("office actions"), are then sent to inventors or their attorneys. The applicant then has a chance to review the examiner's comments and make a rebuttal, called a "response to an office action". In this response, the applicant's task is to indicate the differences between the cited reference(s) and the claimed invention and note the significance of such differences.

Written Disclosure

An applicant must provide a fully enabling written disclosure (8) (i.e., the patent application) in order to obtain patent rights. The written description has four components: (*i*) It must convince another ordinary scientist (an ordinary skilled artisan) at the time of the invention that the inventor (applicant) is in possession of the invention; (*ii*) The description teaches how to make the claimed invention; (*iii*) The description teaches how to use the claimed invention; and, finally, (*iv*) Specific to United States patent law, it needs to teach the best way to make or use the invention (best mode requirement).

Actual experiments do not necessarily have to be performed for a fully enabling written disclosure to be achieved. Prophetic examples (i.e., experiments which have not yet been carried out) are acceptable, as long as an ordinary skilled artisan would be able to perform the experiments and obtain the results claimed in the application. In writing the application, it is critical to use present tense for prophetic examples. If not, the application may be unenforceable (9).

The Enabling Idea

The basic rule is that the inventor is the person who has the first enabling idea which achieves the claimed invention. The day this inventor has the enabling idea is the day he conceives the invention. The inventor does not need to perform a single experiment if conception, i.e., the enabling idea, is complete. The key word here is "enabling," which means something which can be taught and repeated by a person who follows the instructions in the patent. For example: Principle Investigator X tells a postdoc: "Dr. Y, find me a cure for AIDS." After two years of research, Y discovers Invention A, a cure for AIDS. Even if X provides the space and salary for Y to make the discovery, and the patent application claims the use of Invention A to treat AIDS, Y is the inventor, not X.

The above example may have different result if Y reports to X every month about his/her progress after X establishes the original direction. Then X gives suggestions about future direction and comments on Y's experimental results. Finally, after working together two years, they come up with using the nucleotide analog for HIV inhibition and, in one experiment performed by Y, Invention A's activity against AIDS is discovered. In this case, even though X is not physically there when the discovery is made, he/she contributed enough to qualify as a co-inventor if the application claims the use of Invention A against AIDS.

Example

Now, let us say T is a technician who performed experiments for Y. Every day or so, Y instructs T to perform experiments, and T is the one who performs the Invention A experiment. T's contribution is insufficient for him/her to qualify as an inventor.

Sometimes, conception and reduction to practice occur simultaneously. For instance, if one is claiming a particular concentration of a reagent for an assay, the conception and reduction to practice may occur at the same time.

Further Example

A scientist may perform a titration assay (i.e., he/she tries different concentrations to determine the optimal concentration). After performing the experiment and examining the results, he she finds that 0.5 microgram per milliliter works best. When this particular concentration is claimed, the conception and reduction to practice occur at the same time.

Ownership and Inventorship

It is important to note that the determination of inventorship sometimes determines ownership of the invention. For example: A, who works at Institute X, makes Invention I. Later it is revealed that A has collaboration with B, who works at Institute Y. Without B's intellectual contribution, A could not have made the invention; therefore, A and B are joint inventors. If both A and B are obligated to assign their rights to their corresponding institutes, the institutes will co-own the invention. As shown in this example, it may be important to complete an institutional agreement before filing a patent application. This type of agreement defines the rights and duties of each party, i.e., who will be in charge of licensing the invention and how the profit will be divided. Similarly, if the invention is to be owned by the co-inventors, they should sign an inventors' agreement, which is like an institutional agreement, except that it includes only individuals.

Information Disclosure Statement

The inventor and her legal representatives are required to present to the PTO prior art which affects the granting of the patent by filing an Information Disclosure Statement (IDS). The literature can take the form of prior art references, invoices, brochures, models, demonstrations, press releases, news articles, etc.

The IDS should be filed within the first three months after the filing of the application. However, the PTO will not charge you fees if it is filed before the first office action has been issued, or three months after the filing, whichever is later. After the first office action, a late fee will be charged. It is highly recommended that an IDS be filed promptly. If a case receives a prompt Notice of Allowance, say, in the third month after filing, the submission of an IDS at that point will create many problems.

An IDS is important if the patent needs to be enforced. Usually when an infringer attacks the validity of the patent or patentee, his usual first argument is that the patentee did not present all pertinent prior art to the PTO and that this is why the patent was issued in the first place.

PATENT DUE DILIGENCE PROCESS, EVALUATION OF PATENT, ENABLING TECHNOLOGY AND CONCEPT OF FREEDOM OF OPERATION

Patent Due Diligence Process

Due diligence is the exercise of due care before a transaction occurs. Patent due diligence will be done during technology transfer and evaluation of the value of the technology. Only technology protected by a patent which survived the due diligence process may obtain high evaluation. Below is a typical checklist for patent due diligence:

1. Obtain technical description of products. In the pharmaceutical area, it should include formulations and manufacturing processes. Review FDA filings.

2. Assess the procedures for identifying patentable inventions and designs, and for ensuring applications are timely filed. Determine whether the procedures are followed and are appropriate and effective under the circumstances.
3. Obtain a complete list of the company's United States, international, and foreign patents and patent applications, both utility and design.
4. Obtain confirmation that the company has recorded assignments for all United States and foreign patents and patent applications.
5. Determine whether the company has assigned or granted security interests against any patents or patent applications.
6. Obtain patent maintenance and annuity fee records. Obtain confirmation from independent sources. Identify patents that are expired and/or no longer enforceable.
7. For patents of special interest, request all prior art in company's files. Determine whether there are any validity issues that would justify further investigation.
8. Obtain any correspondence from the company accusing others of infringing its patents and/or offering licenses under the company's patents. Consider whether any matters justify further negotiations and/or litigation.
9. Identify any actual or threatened litigation/claims against the company, such as cease and desist letters. Identify all license offers made to the company. Assess the merits of all such allegations against the company. Identify the current status of any ongoing proceedings or negotiations. Obtain copies of settlement agreements and releases.
10. Identify and review all license agreements, covenants not to sue, and indemnification agreements.
11. Review the results of patentability and right-to-use searches conducted or commissioned by the company. Consider whether to request corresponding legal opinions, keeping in mind that disclosure of such opinions may potentially waive the attorney-client privilege.
12. Review all records of audits conducted by or against the company pursuant to any type of IP license agreements and/or research and development agreements.
13. For U.S. patents of special interest, obtain assignment records from PTO and conduct UCC searches. Engage foreign counsel to confirm ownership and clear title to foreign patents of special interest.
14. Search for patents and patent applications in the names of key personnel, consultants, and principal investigators to ensure that they were assigned or licensed to the company.
15. For patents of special interest, where further investigation is justified, obtain prosecution histories from PTO.
16. Check employee, consultant, principle investigator, and officer agreements to confirm obligations to assign United States and foreign rights.
17. Conduct freedom-to-operate searches for company's products and processes, including contemplated future products and processes. Assess the results of the searches.

Reviews on Other Issues

Usually, it is not simply patents alone that should be of concern. When due diligence is performed, the investigation should perform the following as well:

1. Review Employment Agreements of all staff.
2. Review IP Policy if there is one.
3. Consider any potential improper anticompetitive effect or antitrust scrutiny under the circumstances.
4. Review press, reports from trade shows, SEC, and annual reports.

5. Determine whether key technologies and other IP rights have been transferred or licensed to one or more government agencies, e.g., via United States government purpose rights provisions.
6. Consider applicability of other types of IP, including semi-conductor chip protection, right of publicity, plant patents, domain name registrations, etc.
7. Assess adequacy of insurance coverage against IP infringement claims.
8. Consider the character of key licensed rights with respect to, e.g., exclusivity, field of use restrictions, geographic-restrictions, and royalty rate structures, etc.

Enabling Technology and Freedom of Operation

In order for products to be developed, sometimes, certain technology or materials may be required. Without said technology or material, one cannot manufacture the products. Accordingly, potential licensee for the product will need to consider if he wants to commercialize the product, he must be able to acquire rights for the enabling technology or material.

Similarly, patent rights only give the patentees rights to exclusive others from practicing the claimed invention but do not give positive rights to practice his own invention. The owner of the invention might not be "free" to operate the invention. See *supra* section "Fundamentals of Patent Concept and the Patenting Process", 1st paragraph. For example, the patent portfolio protects the new uses of an old compound. However patents covering the old compound have not expired. Therefore, the owner of the uses patent may not use the compound without infringing the rights of the compound patents (10). Therefore before the practice of an invention, owners should perform freedom of operation and product clearance analysis. Below is some basics:

1. Activities which leads to a product:
   a. process of how the product was made;
   b. what is the product; and
   c. how the product is used.
2. Searches of other entities' activities. These searches should be as complete and exhaustive as possible.
3. Analysis
   a. Are these activities protected by patent or other rights?
   b. such as IP rights?
   c. Could these rights be designed around?
   d. Side by side comparison: What others do versus what will be done on this product?

The above study and analysis should be done when plans are made for the development of any product.

LOCAL AND INTERNATIONAL IP PROTECTION AND HOW TO PROTECT YOUR VALUABLE TECHNOLOGY/PRODUCT CORRECTLY

As explain earlier, the owner of the technology might want to start with one locality for protection first, and then go for other jurisdictions. Patent rights are geographical rights and therefore, the protection needs to go from one country to another. Since patent protection is the most important form of protection in pharmaceutical technology, below we will focus more in this area. The applicant for a patent application will have one year to consider filing in other countries (11).

In the United States, applicants (inventors) are allowed to write prophetic examples, *supra*, and therefore, the applicants can design experiments to prove the concept before actual experimentation (reduction to practice). This is a great advantage as experimentation takes time and money. However, most countries do not accept prophetic examples. Hence, the first twelve months would be critical to perform the experiments if foreign rights are to be considered.

Patent Cooperation Treaty

Established in the eighties of last century, PCT has been administered by World Intellectual Property Organization. Now, there are more than 100+ countries which are members of PCT. Note that based on various reasons, there is still some countries or jurisdictions which are not (12). By filing one PCT application, copies of the application will be sent to all PCT members. The applicant will have either thirty or thirty-one months (13) from the first filing (priority) date. The deadline for filing the PCT is not extendable and the entry to each country (national stage) generally is not extendable (14). Therefore, if one is interested in filing a foreign patent application or considering doing so, marking of the anniversary date of the national filing is critical.

Protection of Specific Countries, When, How, Cost and Benefit Analysis

Generally, considerations should be given to market, technology, judiciary, and costs. When an application is ready to be filed internationally, the applicant should be cautious in compliance with different laws in different countries. We recommend:

1. review filed application carefully;
2. make sure that all experiments for proof of conception have been done correctly;
3. review the prophetic examples and reduce them to practice if possible; and
4. review the format of the application so that it can be used in multiple countries.

Direct or Via Treaty

We have noted the usage of PCT filing. There are other filings that can be done based on the Treaty. For example, European Patent Office (EPO) covers most Western countries, except Norway. The applicant has to decide whether to enter a country direct or indirectly. Generally, indirect entry is more economical if there are more than three countries which are covered by the Treaty. One shortcoming of entering indirectly is that it might slow down the process. Direct entry, though it may cost more, is the fastest way the applicant can get a patent in a certain country.

Which Countries?

Which country to file is really depending on the following factors:

1. Market: Is the market large enough and worth to pursue the protection.
2. Technology: Could the people in this country master the technology so that they might infringe if there is no protection filed.
3. Judiciary system: Does the judiciary system of this country protect the issued patent. If the system is corrupted, it simply does not matter who is right or wrong.
4. Cost: Generally, budget ten thousand U.S. dollars per country: some more, some less.
5. *Difficult decision yet should be decided early*. Which Countries to pick? For example, for Pacific Rim protection, one may want to cover Australia, China (P.R.C.),

Japan, Hong Kong, India, Korea, New Zealand Singapore, and Taiwan, How about Macau since Hong Kong is protected? Macau is just a neighbor. Those questions can be readily extended: How about North Korea as Korea is protected? How about Mongolia, Malaysia, Indonesia, and Vietnam?

In general, the United States, EPO, Japan, and India probably cover most of the market shares in the pharmaceutical industry. Depending on the situation, one may want to seek protection in Canada, Australia, and Pacific Rim (15).

Early Planning

After knowing that the process is complicated, it is then easy to appreciate the importance of planning in the first twelve months after the first filing. Work needs to be done during this time and should be carefully mapped out. In the laboratory, more experiments should be done to substantiate the invention claimed in the patent application. More importantly, the commercial side of the invention needs to be exploited:

1. Identification of the commercially viable products which are covered by the patent(s);
2. Licensing Potential;
3. Partnership for sponsored research;
4. Counseling—find people who can help commercialization of the product; and
5. Need to know who the players are.

Decisions need to be made early to reduce costs and avoid making mistakes that will require last minute rush decisions.

EXAMPLES OF PATENT IN PHARMACEUTICAL INDUSTRY

Example 1

The first example exemplifies the true advancement of science and innovative idea in pharmaceutical industry. A novel oral controlled release drug delivery system using osmotic pressure and a laser drilled hole to obtain a zero-order drug release for oral administration. The first patent, an elementary osmotic pump, was filed by Alza Corporation (US Patent No. 3,916,899, granted November 4, 1975). Figure 1 illustrates such an oral osmotic drug delivery tablet for osmotically administering a physiologically or pharmacologically-effective amount in the gastro-intestinal tract of animals including veterinary animals and humans. Subsequently, a flourish of patents moved the original patent into an advancement of science and many drug products. Figure 2 illustrates an apparatus for drilling holes with a laser beams for those tablets (US Patent No. 4,063,064 and related US Patent No. 4,088,864). The simple osmotic delivery device also advanced into several modifying forms. Figure 3 illustrates a modified osmotic device with a separate layer or compartment of a fluid swellable hydrogel to force or push the content of another compartment of drug that is insoluble to very soluble in aqueous and biological fluids (the so-called "push–pull" tablet, US Patent No. 4,327,725). Figure 4 illustrates yet another modified osmotic device that inside the tablet comprises of two separate drug compartments separated by a swellable hydrogel partition. When the hydrogel partition swells and pushes both drug compartments to deliver two drugs simultaneously in a controlled manner. Such a tablet was termed "pull–pull" tablet (US Patent No. 4,449,983). This example demonstrates the change of technology and advancement of scientific sophistication from a simple elementary pump to various osmotic tablets.

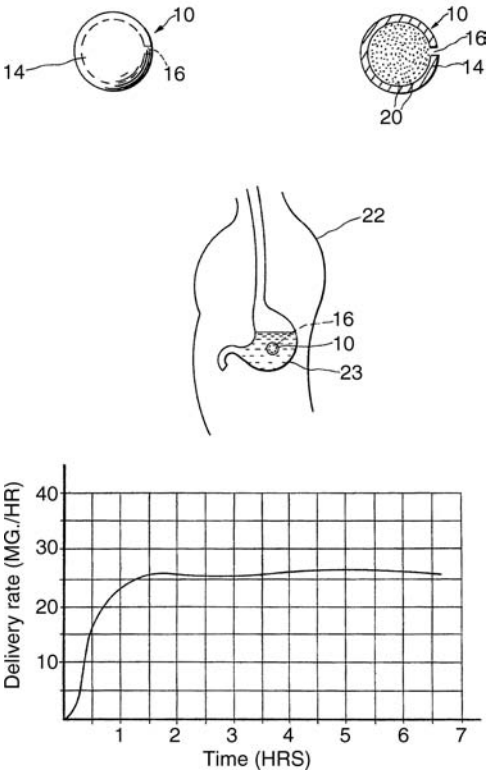


FIGURE 1 An oral osmotic drug delivery tablet for osmotically administering a physiologically or pharmacologically-effective amount in the gastro-intestinal tract of animals including veterinary animals and humans.

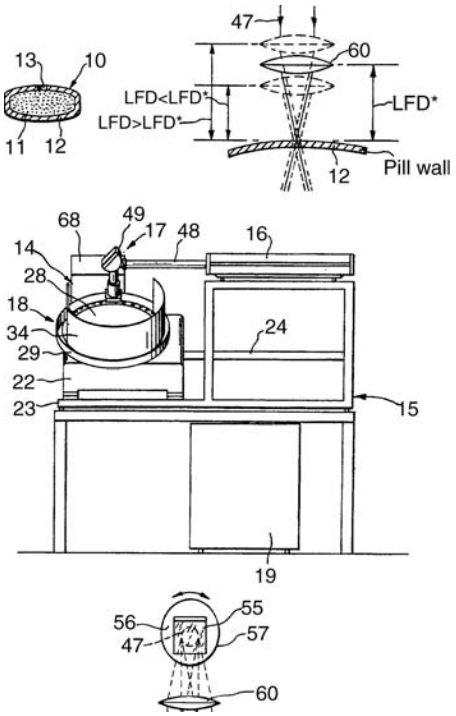


FIGURE 2 An apparatus for drilling holes with a laser beams for those osmotic tablets.

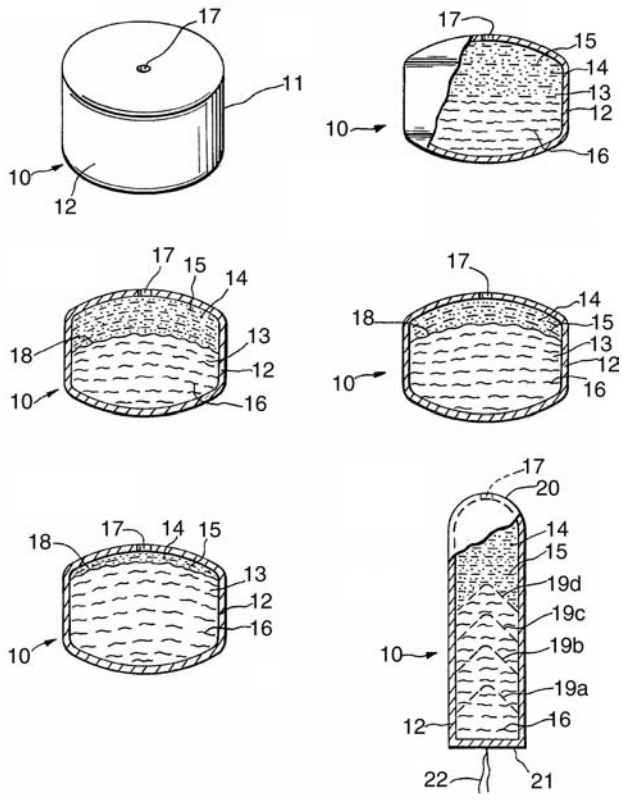


FIGURE 3 A modified osmotic device with a separate layer or compartment of a fluid swellable hydrogel to force or push the content of another compartment of drug that is insoluble to very soluble in aqueous and biological fluids (the so-called "push–pull" tablet).

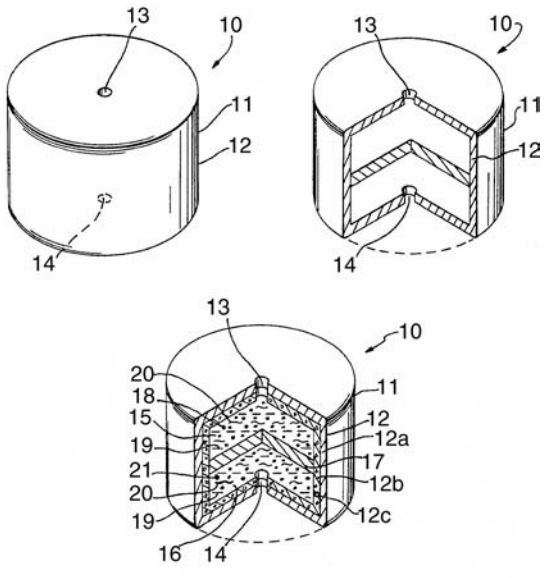


FIGURE 4 Another modified osmotic device that inside the tablet comprises of two separate drug compartments separated by a swellable hydrogel partition. When the hydrogel partition swells and pushes both drug compartments to deliver two drugs simultaneously in a controlled manner (the "pull–pull" tablet).

Example 2

The first example exemplified the advancement of science and improvement of technology. However, there are some examples that demonstrate innovative idea can delay generic drug entry (but unfortunately has nothing to do with advancement of science). One of the examples is Desyrel® (trazodone hydrochloride) 150- and 300-mg oral tablets are designed to be split into three equal parts (the so-called Dividose® design). The design is covered by US Patents No. 4,215,104 and 4,258,027. Figures 5 (rectangular) and 6 (oval and round) illustrate some examples with various shapes of those so-called

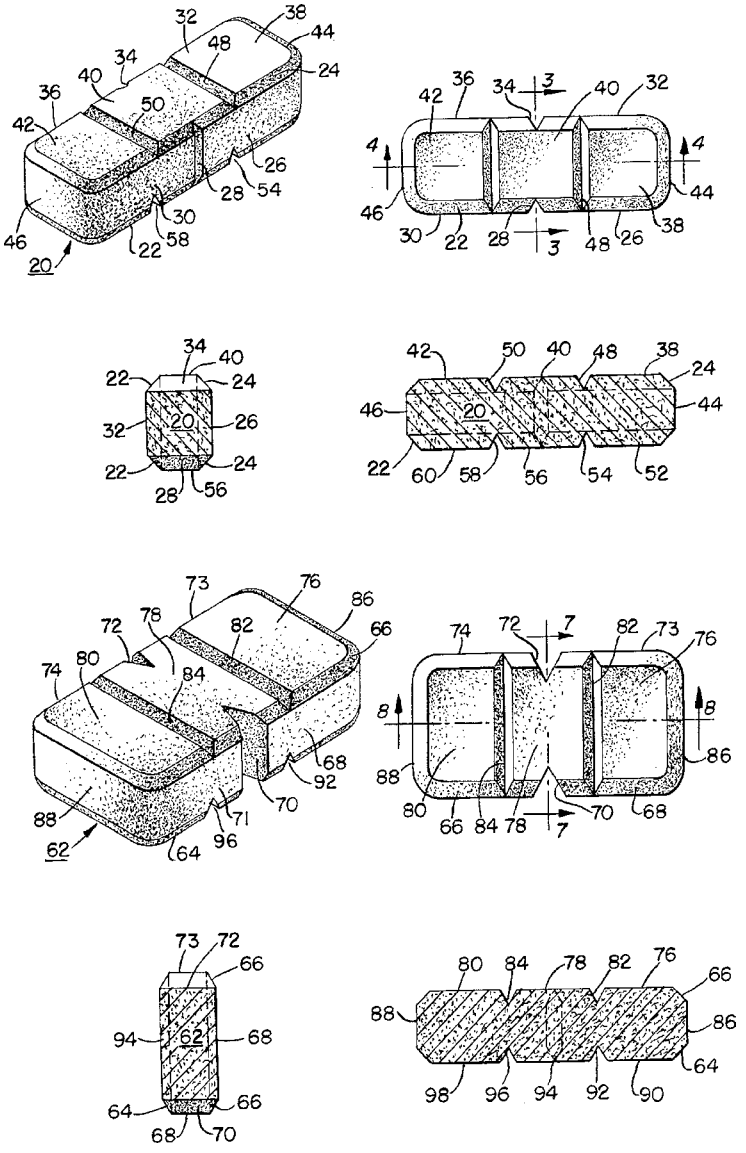


FIGURE 5 An example of the so-called multi-fractionable pharmaceutical tablets that can be separated into three equal parts (rectangular tablet).

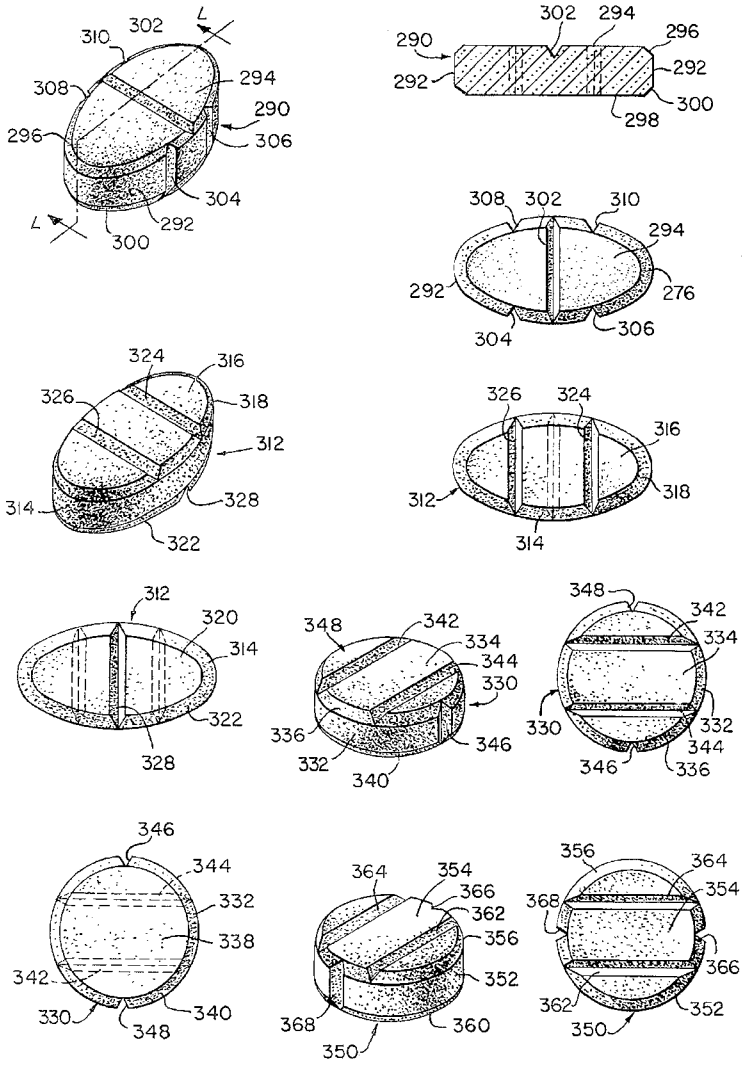


FIGURE 6 An example of the so-called multi-fractionable pharmaceutical tablets that can be separated into three equal parts (oval and round tablets).

multi-fractionable pharmaceutical tablets that can be separated into three equal parts. The patent holder is able to keep a generic version of the drug off the market claiming that the generic tablets infringe on the form of the pill since the generic drug product, like the brand-name medicine, also has two grooves on it to split the tablet into three equal parts. This example demonstrates the importance of patents as offensive and defensive tools to defend its product.

CONCLUSION

This chapter attempted to discuss the importance of IP in biotechnology as well as the pharmaceutical industry. Due to the ever escalating high cost of new drug development,

the drought of new drug pipeline and fierce competition of generic drug industry, it is extremely important for all pharmaceutical scientists working in the industry to understand the protecting mechanism for their invention.

REFERENCES

1. DiMasi JA, Hansen RW, Grabowski HG. The price of innovation: New estimates of drug development costs. J Health Econ 2003; 22(2):151–85; Kaitin KI, eds. Cost to develop new biotech products is estimated to average $1.2 billion. Tufts Center for the Study of Drug Development Impact Report, 2006; Nov/Dec; 8(6).
2. Life Cycle Management is an integrated concept for managing the total life cycle of goods and services towards more sustainable production and consumption. http://www.fivewinds.com/uploadedfiles_shared/LifeCycleManagement040127.pdf.
3. Albert W-KC. Inventor's Guide for Patent Protection. 1992; www.kitchanlaw.com.
4. The tort of trade secret misappropriation protects only information that is properly classified as a trade secret. A trade secret is information (*i*) that is used in a business, (*ii*) that is secret, and (*iii*) that gives a competitive advantage to the person with knowledge of it. (Citation omitted) by Perritt HH, Jr. Trade Secrets A Practitioner's Guide published by Practicing Law Institute, New York City, 1995:3–4.
5. If on goods, it is called trademark, while on services, it is called a service mark, e.g., In the airline industry, "Fly the Friendly Skies^{SM}" is the service mark for United Airlines. Similar "Work Hard, Fly Right^{SM}" is Continental Airlines' service mark.
6. It has been claimed that the biotechnology industry was created by patent protection. See e.g., a recent article in The New York Times which commented that there are many biotechnology or pharmaceutical companies which do not have any product yet but maintain a strong patent portfolio. Andrew Pollack, It's Alive! Meet One of Biotech's Zombies, Sunday, New York Time, February 11, 2007.
7. Prior art is patent jargon. Prior art means what is known or published at the time of the invention. Generally, it includes not only literature and patents but also certain activities, such as exhibits in trade show; public speeches. See 35 U.S.C. §102.
8. 35 U.S.C section 112 recites: "The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention."
9. Roche H-L. Inc. v. Promega Corp., 323 F.3d 1354, 2003; Reviewed by Kevin Mack, Intellectual Property: Patent: Note: Reforming Inequitable Conduct to Improve Patent Quality: Cleansing Unclean Hands 21 Berkeley Tech. L.J. 147, 2006.
10. Said compound patents are called "blocking" patents, which block the practice of other patents. http://www.aicpa.org/pubs/jofa/nov2004/cromley.htm.
11. Most of the countries are signatories of the Paris Convention, which will give one year grace period for filing in countries who are also member of the Paris Convention. E.g., Algeria, Austria, Belgium. See Patent Corporation Treaty, Article 4. http://www.wipo.int/pct/en/seminar/basic_1/priority.pdf.
12. For example, Taiwan, Republic of China, is not a member of PCT based on political reasons. http://www.wipo.int/pct/en/texts/pdf/pct_paris_wto.pdf.
13. More and more countries now turn to a thirty-one month country. However, United States maintain to be a thirty month country. http://www.wipo.int/pct/en/texts/pdf/time_limits.pdf.

14. There are exceptions e.g., For People's Republic of China, extension of additional two month is possible upon payment of a fee. See Implementing Regulations of the Patent Law of the People's Republic of China, Rule 101.
15. An invention such as compounds again Severe Acute Respiratory Symdrome virus should be better protected in the pacific rim.

10

Near-infrared Chemical Imaging for Characterizing Pharmaceutical Dosage Forms

Gerald M. Sando, Linda H. Kidder, and E. Neil Lewis
Malvern Instruments, Columbia, Maryland, U.S.A.

INTRODUCTION TO NEAR-INFRARED CHEMICAL IMAGING

Near-infrared chemical imaging (NIRCI) characterizes pharmaceutical solid oral dosage forms by measuring molecular absorption properties in the near-infrared region in a spatially resolved manner. Molecular absorptions in the near-infrared are primarily due to overtones and combination bands of fundamental molecular vibrational frequencies of C–H, N–H and O–H bonds. This spectral information can be used to characterize the chemical composition of organic material. Single point near-infrared techniques, which result in a single spectrum that is averaged over the entire sample, provide information about the identity and abundance of the chemical components of a sample. In addition to this information, NIRCI characterizes spatial distribution by generating tens of thousands of spatially resolved spectra. NIRCI in essence provides a chemical picture of the sample. The technique combines chemical and image analyses, allowing for the characterization of chemical distributions (level of heterogeneity) and also for morphological analysis of the sample. The size and shape of single component domains, granules, or other particles within the sample can be measured.

The measurement time of a near-infrared imaging experiment depends on the type of imaging instrument used. In general, there are three typical implementations that generate imaging data, namely global imaging, and two types of mapping instruments based on interferometers or monochromators. In global imaging, the entire image is measured at once, and spectral information is built up through wavelength scanning. A mapping instrument measures only a portion of the ultimate image area at any given time, and the sample must be moved in order to map the entire desired image area. This can increase the measurement time required to image the same area for a mapping system over that of a global imaging system. However, there are monochromator based systems that acquire data rapidly, in which the sample movement during a process is used for scanning. A full range scan on a global imaging instrument can take anywhere from less than 1 minute up to 4 minutes, depending on the amount of signal averaging. As with most spectroscopic techniques, increased signal averaging requires more time, but will result in an increased signal-to-noise ratio. For interferometer based mapping, a typical full range scan takes 7–30 minutes. In addition, in a global imaging experiment, the time can be shortened down to a few seconds per sample if only a few wavelengths are needed. This is generally not possible with mapping systems.

NIRCI measurements are typically performed using a diffuse reflectance configuration, where the illuminating radiation can penetrate from ~50–100 μm into the sample. For a global imaging system, virtually no sample preparation is needed, the sample is simply placed on the instrument and focused. For a typical mapping system, a flat surface is needed to maintain system focus throughout a scan. This poses difficulties when measuring non-flat samples, such as tablets with domed surfaces, or granules or powders. In addition, global imaging shows more promise than interferometer-based mapping for in-, on-, and at-line applications because of the data acquisition speed, lack of sample preparation needed, and the fact that global imaging systems have no moving parts. Monochromator based scanning systems are also ideally suited for on-line applications because of data acquisition speed and the fact that they have no moving parts.

The general result of a near infrared chemical imaging measurement is what is called a data cube. It is called a cube because it consists of three data dimensions, two spatial and one spectral, representing many spatially resolved spectra. The cube can either be viewed as individual spatially resolved spectra, or as images of absorption intensity at a single wavelength. There are usually tens of thousands of spectra, far too many to manually analyze. Absorption spectra in the near-infrared usually contain features that are broad and overlapping, resulting in less chemical specificity than Raman or mid-infrared spectroscopy. For these reasons, there are specialized data analysis packages that use multivariate chemometric algorithms to sort and classify data (1,2). Analyses can be grouped into two general categories: Supervised, and unsupervised. Supervised analysis, as the name implies, requires some input from the analyst, and is useful if the number and identity of chemical components in a sample is known ahead of time. This is generally the case in pharmaceuticals, where the ingredients are known, but the distribution of these known ingredients is of interest. These methods, such as partial least squares (PLS), use a library of the known components to quantitatively and reproducibly predict the abundance and distribution of each component. If not all of the components are known, an unsupervised method with no analyst input, such as principal component analysis, can be used. One disadvantage of unsupervised methods is that quantitative information about the abundance may not be as readily available.

INSTRUMENTATION TYPES

As mentioned earlier, there are three typical implementations that generate imaging data, namely global imaging and two types of mapping instruments based on interferometers or monochromators. These approaches differ in the method used to build up the image. A global imaging system uses a focal plane array camera to image the entire sample at once. An interferometer based mapping system uses either a single detector or a linear array to measure spectra in one area of the sample and then translates the sample in order to build up an image of the entire sample. Instrumentation that uses an interferometer and a two dimensional (2D) detector also exists, but these have been mostly limited to mid-infrared imaging applications. A monochromator system also uses a 2D detector, where the wavelengths are dispersed along one axis, and the other axis is used to record spatial information. There are several approaches to wavelength resolution. Global imaging uses an image quality, high resolution liquid crystal tunable filter (LCTF) with 6 nm resolution at 1600 nm. The monochromator based approach has similar spectral resolution, generally 5–8 nm. Interferometer-based mapping systems utilize an interferometer for wavelength selection, and are therefore capable of producing much higher spectral

resolution. However since most NIR spectral features are broad, increased resolution does not necessarily add capability.

When measuring spectral features at wavelengths longer than 2000 nm, a cooled detector is generally necessary. One approach is to use a liquid nitrogen cooled detector, such as mercury cadmium telluride detectors. The use of liquid nitrogen can be problematic if unattended operation is desired, since periodic dewar refilling is necessary. Another approach is to use a Stirling cooled Indium Antimonide (InSb) or for wavelengths shorter than 1720 nm, a temperature stabilized Indium Gallium Arsenide (InGaAs) detector, both of which run unattended, and do not require liquid nitrogen.

There are also two types of optics that are typically employed in near-infrared imaging, all reflective Cassegrainian optics, or refractive optics. The use of refractive optics results in a larger working distance and a larger depth of focus, allowing for greater flexibility in samples and sample preparation. For example, imaging of rounded or non-flat samples is easily accommodated by this type of optical arrangement. In addition, there is more flexibility in the available fields of view, or magnifications when using refractive optics compared to Cassegrainian optics. This is particularly true when moving to larger fields of view. Despite the general lack of flexibility of reflective optics, they introduce no chromatic aberration over large wavelength ranges, whereas refractive optics are optimized over narrower wavelength ranges.

APPLICATIONS

Experimental Details

The following applications examples were all taken using a global imaging instrument, specifically a Spectral Dimensions SyNIRgi™ (Malvern Instruments, Inc, Columbia, MD). The samples are illuminated with broadband NIR light. After interaction with the sample, some of the light is diffusely reflected and collected and focused through the instrument optical train. The resulting collected light is wavelength selected using a high resolution LCTF with 6 nm resolution at 1600 nm. The wavelength selected radiation is then focused into an image of the sample onto a Stirling cooled InSb focal plane array with 320×256 pixels. Data are collected over an area ranging from 3.2×2.6 to 40×32 mm depending on the particular system magnification. Unless otherwise noted, images shown in this chapter were recorded with a 10 nm increment over a spectral range of 1200–2400 nm. The images are combined to form a data cube and result in 81,920 NIR spectra. The full range data cubes were collected in less than three minutes.

The resulting image data cubes are processed using the ISys® chemical imaging software (Malvern Instruments, Inc, Columbia, MD). The data undergoes basic pre-processing steps to remove the instrument response function by subtracting the dark current and by taking a ratio with a background consisting of reflected light from a highly scattering white ceramic. The data is then converted to absorbance, mean centered, and normalized to unit variance. Normalization is performed in order to remove effects due to physical differences, such as hardness, density, or scattering, the goal being to isolate chemical from physical differences in the sample.

Chemical Distribution in Tablets

The heterogeneity of an Over-the-Counter (OTC) analgesic was characterized using NIRCI. A PLS model was developed to determine the distribution of the three main components, acetaminophen, aspirin, and caffeine. Each pixel in the image contains a

complete NIR spectrum and the PLS model is applied to each of these 81,920 NIR spectra. A score value of 0 means that the component is not present at that pixel, while a score value of 1 means that the component is 100% pure at that pixel. Most pixel scores vary across the range from 0–1, representative of component mixtures. The images of the PLS scores provide a visual and qualitative representation of the spatial distribution of the material in the sample. The resulting chemical distribution of the tablet is shown in Figure 1. In the composite image, high score pixels for each component are assigned a single color, with acetaminophen in black, aspirin in grey, and caffeine in white. This composite image provides a visual representation of the spatial distribution of all three components in a single image.

The PLS results can be quantitatively analyzed to characterize the component distribution. Figure 2 shows histograms of the PLS results showing the number of pixels at a given PLS score. This is a different way to represent the same information presented in the image, but it enables quantitative and therefore objective analysis of the same information. Images are intuitive, and therefore a powerful way to present data, but for any real quantitative and reproducible analysis, the histogram is a much more useful analytical tool. The primary parameters of interest in the histogram distribution are the mean, standard deviation, skew, and kurtosis. The mean corresponds to the bulk abundance and is equivalent to HPLC or a bulk NIR concentration measurement. The standard deviation measures the width of the distribution. A heterogeneous sample will show a greater pixel-to-pixel variation across the sample and will have a larger standard deviation, whereas a homogeneous sample will have a narrow distribution and a small standard deviation. The skew measures the asymmetry in the distribution. A positive skew shows "hot spots" or areas of localized high abundance, whereas negative skew indicates "holes" or localized areas of low or no abundance. The kurtosis is a measure of the peakedness of the distribution and larger values indicates greater localized sample heterogeneity.

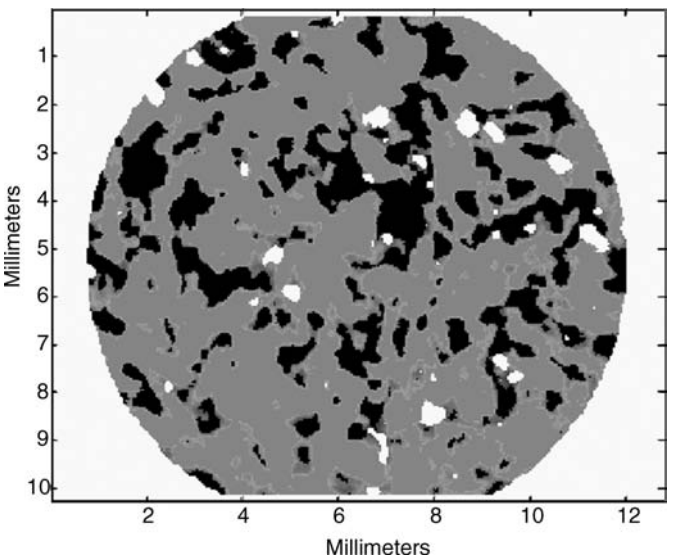


FIGURE 1 Composite image of PLS scores for an OTC analgesic table. The colors correspond to acetaminophen (black), aspirin (gray), and caffeine (white). *Abbreviations*: PLS, partial least squares; OTC, Over-the-Counter.

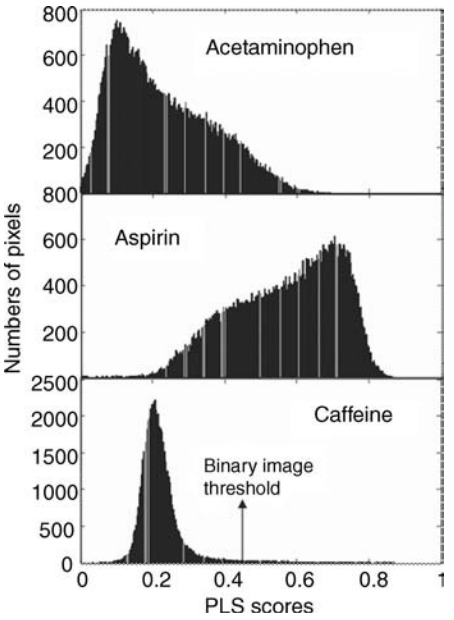


FIGURE 2 Histograms of PLS scores for corresponding to the image in Figure 1. *Abbreviation*: PLS, partial least squares.

The resulting statistics are shown in Table 1. The asymmetry in the distributions is revealed in the skew values. Caffeine, which appears only in relatively small domains of very high concentration, has a very high positive skew value. This is reflected in the tail toward high PLS scores in the histogram distribution. The skew allows for a quantitative and reproducible measure of the extent to which the component aggregates into domains of much higher than average concentration. It can also be seen from the distributions that acetaminophen tends to have "hot spots" that fill in "holes" in the aspirin distribution. This is reflected in the positive and negative skew values for acetaminophen and aspirin, respectively.

Now that the sample has been chemically segmented, morphological image analysis is possible. For this sample, caffeine is the best candidate since it appears to form well defined domains. In order to perform this analysis, a binary image is created. This is done by choosing a threshold and setting all of the pixels above this threshold to 1, and all those below to 0. In this case, the threshold is the mean plus 3 standard deviations. Setting the threshold using this type of statistical parameter is an effective way to ensure reproducibility and to remove the often subjective nature of image threshold determination. The threshold is shown in Figure 2. The PLS scores image and the resulting binary image are shown in Figure 3.

Analysis of the domain size is now possible. There are 33 caffeine domains that cover 2.2% of the area of the tablet. The domain sizes are converted to a circular equivalent diameter, which is the diameter of a circle with the same area. The resulting mean and standard deviation for the diameters are 0.25 and 0.12 mm, respectively.

TABLE 1 Summary of the Statistics of the Histograms in Figure 2

	Acetaminophen	Aspirin	Caffeine
Mean	0.23	0.57	0.23
STD	0.14	0.15	0.07
Skew	0.60	− 0.43	3.43

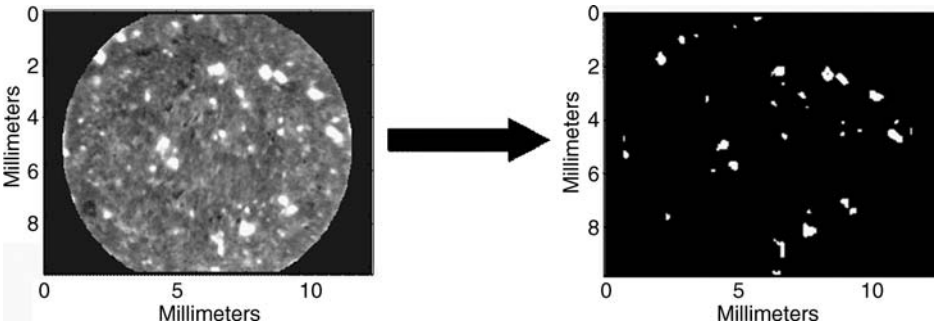


FIGURE 3 Image of the caffeine PLS scores (left) and the resulting binary image (right) created from setting all pixels above a threshold to 1 and those below the threshold to 0. *Abbreviation*: PLS, partial least squares.

Various shape parameters are also available to characterize the various domains. In addition, size and shape parameters are available to characterize each individual domain.

This information can be very useful for product development. Controlling the distribution of components in solid dosage forms can be extremely important in controlling the performance of a product. For example, dissolution rates can be directly affected by the size of domains of active pharmaceutical ingredients (API), or by the co-location of the API with a particular excipient (2,3). Changing a product formulation changes its behavior, however, the various mechanisms by which this occurs are not well understood. There is a need to go beyond empirical observation to understand the impact of changes in the blending process, such as change in size distribution or shape of raw materials, or even the order in which a blender is loaded.

Understanding these processes is the drive behind the Quality by Design initiative. The basic concept is a commonsense approach where quality is designed into, rather than tested into the product (4). A better understanding of the blending process will also make it easier to identify problems before manufacture of the final solid dosage form, where it is most likely too late to prevent a costly loss of product. The information available using NIRCI provides valuable information for correlating the changes in the blending process to chemical distribution, and then correlating chemical distribution to performance. Therefore, near-infrared imaging provides a connection between the blending process and product performance.

High Throughput

An imaging system used in conjunction with a computer controlled translation stage can be used to change samples in an automated manner and to perform repetitive measurements. In addition, the flexible wavelength selection available in a tunable filter-based imaging system can allow for further speed increases. For example, if only a few wavelengths are needed, it is not necessary to collect data over the entire spectral range and this can reduce data collection time to a few seconds per sample. Although near-infrared spectral features are broad and not well separated, this selected wavelength approach can often be applied to many systems.

Shown in Figure 4 is a comparison of results from a PLS prediction on full range spectral data with a five wavelength scan. The sample is the same OTC analgesic tablet as presented in the previous application example. On the left are the PLS predictions for acetaminophen (A) and caffeine (B). On the right are results from the five wavelength

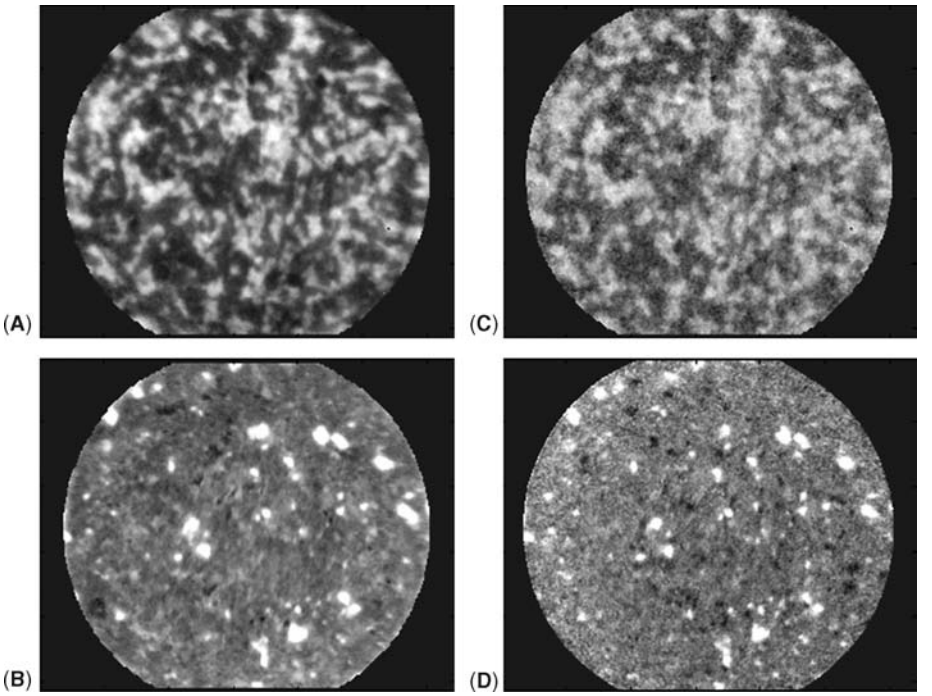


FIGURE 4 PLS predictions for acetaminophen (**A**) and caffeine (**B**) and results from the five wavelength scan for acetaminophen (**C**) and caffeine (**D**). *Abbreviation*: PLS, partial least squares.

scan for acetaminophen (C) and caffeine (D). For each image one wavelength is used for baseline correction, one for normalization, and one to represent a unique spectral feature of the component, a so-called marker band. The normalization wavelength for acetaminophen was used as the baseline correction wavelength for caffeine. The resulting images are very similar to those using PLS on full range data.

To illustrate the usefulness of this approach, fifteen samples were measured using a five wavelength scan. Each measurement took approximately 5 seconds. The analysis of the data was also automated through the use of software macros (ISys®, Malvern Instruments Ltd.) and took less than 1 minute to complete. The statistical results are shown in Table 2. For acetaminophen, all the samples appear to be statistically similar when looking at the mean values, but sample 3 has much larger values for the standard deviation and the skew. Sample 3 is a notable outlier in terms of the caffeine distribution, with a lower mean and larger standard deviation. By doing a statistical comparison of the values between the samples for the caffeine component, sample 3 differs from the mean by at least three standard deviations for these parameters, while the remaining samples fall within one standard deviation of the mean. This procedure, the rapid acquisition of limited wavelength data, followed by automated data processing quickly identified an outlier, in this case a tablet from a different manufacturer.

The combination of high-speed near-infrared imaging with automated data collection and analysis allows for the possibility of high throughput analysis. The use of an automated stage to change samples allows for unattended operation and the measurement of a statistically relevant number of samples with little operator input. This can open up near-infrared imaging for quality control/quality assurance (QA/QC) purposes.

TABLE 2 Statistical Results of the Five Wavelength Scan on a Series of 15 OTC Analgesic Tablets

	Acetaminophen			Caffeine		
Sample	Mean	STD	Skew	Mean	STD	Skew
1	0.64	0.20	0.18	1.26	0.07	1.06
2	0.71	0.23	0.18	1.28	0.09	1.69
3	0.71	0.33	0.47	1.11	0.19	1.60
4	0.69	0.22	0.20	1.29	0.09	1.92
5	0.75	0.23	0.11	1.28	0.09	1.51
6	0.67	0.21	0.08	1.27	0.08	1.43
7	0.67	0.24	0.11	1.27	0.09	2.05
8	0.67	0.22	0.19	1.28	0.08	1.22
9	0.56	0.21	0.18	1.28	0.09	1.81
10	0.64	0.22	0.15	1.27	0.08	1.32
11	0.60	0.22	0.19	1.28	0.07	1.25
12	0.69	0.22	0.09	1.27	0.07	1.28
13	0.64	0.24	0.12	1.27	0.10	1.84
14	0.59	0.21	0.23	1.28	0.08	1.42
15	0.60	0.21	0.13	1.28	0.08	1.32
Average	0.66	0.23	0.17	1.26	0.09	1.51
STD	0.05	0.03	0.09	0.04	0.03	0.29

Abbreviation: OTC, Over-the-Counter.

CONCLUSIONS

Information available through NIRCI such as data on component agglomeration, preferential association of components, and the distribution of free and bound water, provides a significant tool for optimizing formulation development, and global imaging and interferometer based mapping systems are powerful R&D tools in this environment. Global imaging is the best option for a QA/QC lab, where rapid data collection is needed. Global imaging implementations and monochromator based mapping systems which have no moving parts are both ideally suited for manufacturing environments. The ability to acquire data that includes both chemical and spatial information makes NIRCI systems significant analytical tools.

REFERENCES

1. Gendrin C, Roggoa Y, Collet C. Content uniformity of pharmaceutical solid dosage forms by near infrared hyperspectral imaging: A feasibility study. Talanta 2007; in press.
2. Luypaert J, Massart DL, Vander Heyden Y. Near-infrared spectroscopy applications in pharmaceutical analysis. Talanta 2007; 72(3):865–83.
3. Koehler IV FW, Lee E, Kidder LH, Lewis EN. Near infrared spectroscopy: the practical chemical imaging solution. Spectroscopy Eur 2002; 14(3):12–9.
4. ICH Harmonised Tripartite Guideline Pharmaceutical Development Q8, 2005:1–7.

11

Surface Area, Porosity, and Related Physical Characteristics

Paul A. Webb
Micromeritics Instrument Corp., Norcross, Georgia, U.S.A.

INTRODUCTION

The surface area and porosity characteristics of materials are related to the physical arrangement of the molecules rather than their chemical makeup. However, these physical characteristics can be just as important as the chemical constituents in regard to how a chemical reaction proceeds and, thus, is an example of a *physicochemical process*.

Before two or more molecules of the requisite energy can react or interact, they must converge; the probability of such an encounter dependents on several variables. One of the most obvious of these is population—increases the number of qualified participants and the rate of reaction increases. In a solid–gas system, the availability of fluid phase reactant typically is much greater than that of the solid phase. Increasing the number of solid molecules per unit mass available to react is achieved by increasing the area of the solid surface.

The two most common methods of manipulating surface area are by control of particle size (the smaller the particles, the more surface area per unit mass) and by control of the open porosity of the material. In the former case, a material with high surface area would be in the form of a fine powder; in the latter, the material may be granular or even a single solid piece. Almost any solid material can be reduced in size to achieve high surface area, but reforming a material into a highly porous form requires considerably more technology. However, pores not only have surface area, but also volume and the utilization of that volume provides an additional dimension of applicability of a porous material. Porosity also affects the volume and, therefore, the density of materials.

In addition to influencing the rates of reactions, surface area, and porosity can be utilized to store a chemical component permanently (e.g., collection of toxins by activated carbon to prevent stomach and intestinal absorption) or for subsequent release under the appropriate conditions or at an appropriate rate (e.g., osmotic flow through controlled porosity coatings).

Surface Area and Porosity

A simple way to illustrate the concepts of surface area and porosity on a macroscopic scale is to imagine a 300-page, 500 g paperback book as being a particle. Let its dimensions be as

follows: Width (W) = 16 cm, Height (H) = 24 cm, Thickness (T) = 2.5 cm. Closed tightly, the book has a volume (WHT) of 960 cm^3, total surface area ($2WH + 2TH + 2WT$) of 0.0968 m^2 and a *specific surface area* (surface area per unit mass) of 0.000194 m^2/g. Therefore, this single "particle" has a calculated *particle density* of 0.521 g/cm^3.

To increase the available surface area of this example particle by the size reduction method, remove each page and spread them out. This would result in 300 individual pieces, each having 0.038 m^2 of surface area on each side plus the surface area contribution of the edges (thickness = $T/300$), yielding a *total surface area* of 23.05 m^2 and a specific surface area of 0.0461 m^2/g. Of course, the density of each piece is the same as the original "particle" and the total mass remains 500 g.

Increasing surface area by including porosity may be illustrated using the same imaginary particle as above, opening it until the front and back covers just touch, and then carefully fanning out each page so that no two pages touch except at the binding. Effectively, this produces a right circular cylinder of 16 cm radius and a height of 24 cm. In this example, it remains a single "particle," but now has within it an array of slit-shaped pores, represented by the volume between adjacent pages, each page representing a pore wall. This newly formed porous "particle" has the same exposed total surface area (23.05 m^2) and specific surface area (0.0461 m^2/g) as the 300 small "particles" resulting from size reduction described in the paragraph above. The notable difference between the two examples is that the latter case begins and ends with a single particle rather than a collection of smaller particles. The total surface area of the example particle is increased by the total surface area of the pore walls. Actual particles that can be expanded in a similar manner to the example particle are those in a group referred to as vermiculites. They occur naturally in laminar structures resembling mica. The particles expand in a process called exfoliation in which they unfold in an accordion-line manner.

It is important to note that the calculated specific surface area of the example "particle," 0.000194 m^2/g, is extremely small. Expanding the surface by the illustrated methods resulted only in 0.0461 m^2/g of specific surface area, which would be considered very small for an actual material. Now compare the surface area of the example particle to real particles. The specific surface area of a typical pharmaceutical ingredient ranges from about 0.1 to 300 m^2/g. The specific surface area of various carbon structures extend from <1 m^2/g for some graphites, to 500 m^2/g for powdered carbon, to 1000 m^2/g for activated carbon and up to 2000 m^2/g for advanced activated carbons. Synthesized and activated isoreticular metal organic framework structures have specific surface areas reported to extend from 500 to 4500 m^2/g (1). The differences are attributed to microscopic surface features.

The area calculated for a page from the book assumed a perfectly flat surface with no surface features. Purely geometrical calculations of surface area may serve adequately when working at the centimeter and meter scale, but, chemical reactions occur at the molecular level, so surface features of micrometer dimensions and smaller must be taken into account. With such considerations, the specific surface area of a piece of paper typically is found to be a few hundred square meters per gram, perhaps ten thousand times that calculated from linear dimensions.

The Effect of Porosity on Density

There is another important physical attribute associated with the second example "particle." This cylindrical, porous "particle," although maintaining the same mass as when in the cubic rectangular form, now occupies more space. If only the outer dimensions of the cylinder are considered and applying the formula $V = \pi r^2 h$ for

determining the volume of a cylinder, it is found to occupy 19,292 cm^3 while the original "particle" only 960 cm^3. When the cylindrical volume is used to calculate density, the newly formed "particle" has a density of 0.026 g/cm^3 compared to the original "particle" density of 0.521 g/cm^3.

When the volume of an object includes pore volume, as does the cylindrical object just described, it is termed *envelope volume*. Following the "book-particle" example, if all of the 300 separated pages from the size reduction example were to be collected and restacked, it is unlikely that the height of the stack would be the sum of the thickness of each page as it originally was, but considerably greater since there would be voids between the pages since they no longer are flat as when neatly bound between two covers, but now are bent, curved, creased, and wrinkled. In a collection of actual particles, these voids are called *interpartical voids* or *interstitial voids* and they contribute to the volume of the loosely reassembled mass. When the dimensions of the loosely stacked collection of individual "particles" are measured and volume calculated, the value represents the *bulk volume* and includes interparticle void volume.

When total mass is divided by either bulk or envelope volume, the result is *bulk density* or *envelope density*, respectively, both being less than *particle density (skeletal density)*, the density of the material calculated with a volume value that excludes the volume of pores and voids. These definitions provide a way to determine total pore volume. Using the case of the example cylindrical "particle," both the envelope and skeletal volumes were calculated from physical measurements. The difference between these, 18,332 cm^3, is the total pore volume. The same type calculation using skeletal volume and bulk volume yields the interparticle void volume.

In drug development, understanding the relationship between a desired effect and the extent of surface area (or degree of porosity) requires measurements of these physical characteristics. The production and quality assurance process also depends upon the same analyses from inspection of incoming raw materials, control of production and quality control of the finished product. However, as has been illustrated, simple linear measurements, even on a microscopic scale, are inadequate for the determination of surface area and the same applies to the characterization of porosity. What is required is a technique by which the surface features and pore space are investigated with a probe of a size no larger than the smallest feature to be measured. Although several automated analytical techniques currently are in use, the most widely used for accuracy and precision are the physical adsorption of gas molecules for both surface area and porosity determinations, high-pressure mercury intrusion for porosimetry, and gas displacement pycnometry for volume determinations.

The following sections provide overviews of these analytical techniques and the physical characteristics for which they provide information. Prior to the discussion of each instrumental technique, the physical theory utilized by the instrument is presented. Each section concludes with data reduction methods and theoretical models used to extract information about the sample material from the raw data.

PHYSICAL ADSORPTION AS AN ANALYTICAL TECHNIQUE

Physical adsorption is a surface phenomenon by which gas molecules (the *adsorptive*) are weakly bound (*adsorbed*) to the surface of the solid (the *adsorbent*) by van der Waals forces. Physical adsorption takes place on all surfaces provided temperature and pressure conditions are favorable. Stated more precisely, physical adsorption results in a higher concentration of the fluid molecules at the fluid–solid interface than exists in the fluid

bulk. Physical adsorption does not affect the structure or texture of the adsorbent, and *desorption* takes place readily when conditions are reversed.

The definition above applies to the bulk process. At the atomic or molecular level, the time a specific, individual molecule remains on the surface is extremely small and an adsorbed molecule quickly breaks the surface bond (*desorbs*) and rejoins the bulk gas phase surrounding the solid. Although the time an individual molecule spends on the surface is small, others quickly replace those liberated.

An adsorbed molecule escapes the surface by acquiring more energy than that of the adsorption site to which it is bound. The liberating energy is of thermal origin and is passed from one molecule to another (solid–solid, solid–gas, and gas–gas) by collision and is manifested in vibratory motion of the adsorbed molecules and those of the solid surface. It, then, is understandable that lowering the temperature of the system reduces the probability of escape from the surface, thus increasing the number of molecules on the surface at a given instant.

A Physical Adsorption Experiment

Imagine a solid material with no pre-adsorbed contaminants on its surface and enclosed in a perfectly evacuated sample tube (Fig. 1). The open end of the tube is sealed from atmosphere by a valve system (manifold) and the temperature of the tube and its contents is maintained at T degrees Kelvin by a cold bath. Assume that a valve is momentarily opened to allow n moles of gas to enter the tube. The gas will expand to fill the free volume (V) of the tube and the pressure, P_1, within will equilibrate at nRT/V, where R is the universal gas constant and n is the quantity of molecules expressed in moles. (In subsequent discussion, the general quantity of molecules will be symbolized by q unless a specific quantity unit is more conventional in the context of the subject.)

The gas molecules are in random motion, colliding with each other, the walls of the sample tube and the surface of the solid. As previously described, some molecules will temporarily adsorb onto the solid surface. At some time (t) after opening the valve, the number of molecules (q_1) residing on the surface at any instant thereafter will assume a

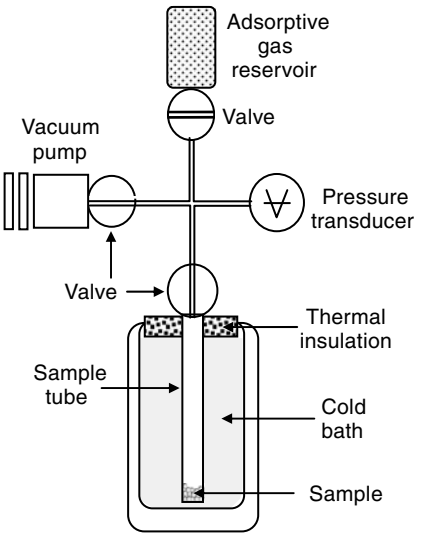


FIGURE 1 A physical adsorption experiment. A simple apparatus is illustrated in which a physical adsorption experiment could be conducted. In the valve configuration shown, the sample tube is being evacuated.

constant value indicating that the rate of adsorption equals the rate of desorption. This condition is called *adsorption equilibrium* and t is the *equilibration time*.

Assume the valve is again opened to allow another dose of n moles of gas to enter the tube; the same process ensues as described above, but, with additional molecules contained in the same volume, the frequency of collision with the surface increases. After all processes have equilibrated, pressure (P_2) within the tube will be higher than (P_1) and the number of molecules on the surface at any instant will have increased to q_2.

If this stepwise process is continued until pressure within the tube achieves that of the atmosphere, over the course of the experiment there will have been observed a set of n pressure versus quantity adsorbed ordered pairs that, when plotted over the range 1 to n, produce a graph called an *adsorption isotherm*, the name indicating that each ordered pair (P_i,q_i) was measured at the same temperature.

Physical adsorption is a reversible process. Imagine that the vacuum valve in Figure 1 is manipulated to remove small quantities of gas at each step and the above experiment continued. In this phase of the experiment, each momentary opening of the valve withdraws n moles of gas from the tube. The values of P and q would decrease after each step; a plot of all (P_i,q_i) data is called the *desorption isotherm*.

Contrary to what may seem intuitive from the simple explanation above, the plotted data points from actual adsorption experiments will not produce a straight line. Instead, variations of one of six types of isotherms will be produced; examples are presented in Figure 2. The first five originally were assigned type numbers by Brunauer (2). The sixth is a recent addition. Type 1 is characteristic of adsorbents having extremely small pores (micropores). Types 2 and 4 are indicative of either nonporous adsorbents or adsorbents having relatively large pores, and Types 3 and 5 arise under conditions where adsorptive molecules have greater affinity for one another than they do for the solid. The Type 6 isotherm, indicative of a nonporous solid with an almost completely uniform surface, is quite rare.

A plot of desorption data is unlikely to retrace the adsorption path until pressure has been considerably reduced. This produces a *hysteresis loop* as illustrated in Figure 2 for the Types IV and V isotherms. The shape of the isotherm contains information about the surface of the solid—its surface area, surface energy distribution, pore volume, the sizes of the pore openings at the surface and, to some extent, the shape of the pore cavity.

Applications of the Ideal Gas Law to Determine the Number of Molecules Involved in Surface Coverage and Pore Filling

The following information is essential not only to understanding the adsorption process on the solid surface and within pores, but also in understanding the instrument's measurement process. Awareness of exactly what the instrument measures provides the necessary insight to develop efficient and accurate analytical methods to characterize the surface of various solid materials.

A fundamental relationship when working with gases is the ideal gas law

$$PV = nRT \tag{1}$$

where n is the number of moles of gas, P the absolute pressure, V the physical volume of the vessel containing the gas, R the universal gas constant, and T is the absolute temperature. For a specific number of moles of gas subjected to various combinations of pressure, temperature, and container volume, it is apparent from Equation (1) that if no gas escapes the system and no additional gas is allowed to enter the system, the only simultaneous values of P, V, and T that are possible are those that satisfy the condition

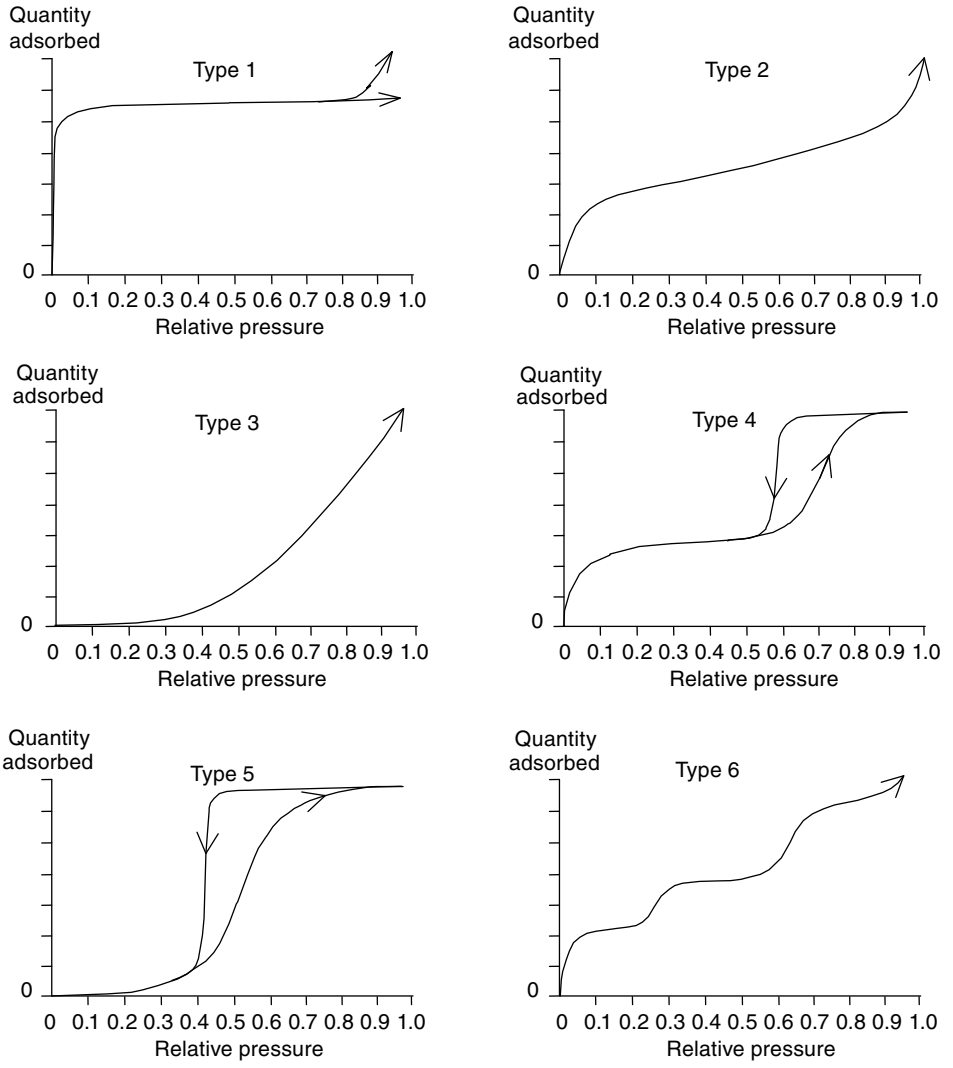


FIGURE 2 The six types of physical adsorption isotherms. A visual inspection of the isotherm can provide information about the surface features of the material under test. Considerably more information is available through the application of one or more data reduction methods.

$$PV/RT = n \tag{2}$$

In terms of any initial and final values of P, V, and T that are associated with a change of configuration,

$$P_i V_i / T_i = P_f V_f / T_f \tag{3}$$

is the controlling relationship between configuration 1 and configuration 2. Equations (2) and (3) in various rearrangements are applied throughout the following sections to determine the quantity of gas in a container of constant volume by measurements of pressure and temperature.

Standard Volume

A convention which is used at times and which employs the $PV/T = constant$ relationship is the expression of gas quantities in terms of standard volumes. Consider a sample holder of physical volume V_i that contains n moles of gas at pressure P_i and temperature T_i. The same quantity of gas, if at standard temperature T_{std} (273.15 K) and standard pressure P_{std} (760 torr), will have a volume V_{STP} that the relationship

$$P_i V_i / T_i = nR = P_{std} V_{STP} / T_{std}$$

or

$$V_{STP} = V_i (P_i / P_{std})(T_{std} / T_i) \quad (4)$$

where V_{STP} has units of cm^3 STP. It is accepted that one mole of ideal gas at standard temperature and pressure occupies a volume of 22,414 cm^3. The number of moles n contained in any standard volume of ideal gas can be determined by dividing the volume expressed in cm^3 STP units by 22,414 cm^3/mole. So, a quantity of gas expressed in units of standard volume express the molar quantity of gas and, by use of Avagadro's number, also conveys the number of gas molecules.

Determinations of Surface Area and Porosity from the Adsorption and Desorption Isotherms

As has been stated, measuring surface area and porosity is of primary importance in controlling and gaining maximum advantage of various phenomena associated with these two physical attributes. A single analytical technique that is capable of determining both characteristics takes advantage of the physical adsorption phenomenon. This technique allows the specific and total surface area of a sample to be determined as well as the total pore volume and the distribution of pore volume by pore diameter. It also can reveal information about the surface energy of the material.

The generic instrument type is a gas sorption analyzer, "sorption" implying either adsorption or desorption. Gas sorption analyzers that are used to determine surface area and porosity can be divided into two types: (*i*) volumetric and (*ii*) dynamic physical adsorption analyzers. A volumetric physical adsorption analyzer was described in the physical adsorption experiment at the beginning of this section and is illustrated in Figure 1. Dynamic physical adsorption analyzers, also called "flowing gas" analyzers typically operate at about atmospheric pressure and expose the sample to various concentrations of the analysis gas mixed with an inert carrier gas. Adsorption equilibrium is established at the partial pressure of the analysis gas at the prevailing concentration.

Because of the requirement to blend gases or to have a supply of pre-mixed gases, analyses are more tedious, particularly if more than a very few equilibrium points are desired. These instruments, however, are useful for obtaining very fast single point Brunaure, Emmett, and Teller (BET) surface area determinations. But, because of their limitations, only volumetric analyses are discussed further in this work.

Sample Preparation and Analysis

Elevating temperature is the primary method of cleaning the surface of a specimen in preparation for an adsorption experiment. The liberated molecules are carried away from the sample by either vacuum or by flowing inert gas over the sample, neither of these methods having significant advantages over the other in the majority of applications. The

importance of beginning a test with a sample free of adsorbed molecules cannot be overstressed.

Atmospheric contaminants are the most common and adsorbed water vapor is of particular concern. If the temperature of the sample is elevated too rapidly during preparation, steam can form within the pores of the sample and result in physical alteration of the material. To avoid this, the temperature should be raised to just below 100°C and held at that temperature for some time before proceeding with the temperature ramp.

Attempting to analyze a sample with adsorbed contaminant molecules on the surface will result in anomalies in the adsorption isotherm as the contaminate competes with the analysis gas for adsorption sites or is liberated to join the bulk gas above the sample. Another consideration is the purity of the source of analysis gas. As will be seen, precisely determining surface area and porosity by the physical adsorption technique requires that a single gas of analytical purity be dosed into the sample space. Even if the recommended 99.99% purity gas is received from the supplier, the regulators and gas lines can introduce impurities.

Data Reduction Theories

It will be noted in the literature that most data reduction methods express pressure in relative terms, P/P^0, where P^0 is the saturation pressure of the adsorptive gas. A benefit of this choice of units is to more easily allow isotherms to be compared since all isotherms are then bound to a range between zero and one, the point at which the adsorptive condenses to a liquid. It also "normalizes" the saturation point for all gases to be when $P/P^0 = 1$. It will be noted, also, that the quantity of adsorbed molecules (y-axis) is expressed in conventional units of standard volume; a more recent preference is to express this quantity in moles.

There are numerous theories or models of the adsorption and desorption processes that account for the shape of the isotherm. The two most widely used in the determination of surface area are the Langmuir theory (3) and the BET theory (4), the latter being applied most widely in physical adsorption. Both theories describe the progression of surface coverage by gas molecules and both theories describe a point in the process at which the surface is covered with a single layer of molecules. This point in the adsorption process is termed *monolayer coverage* and the quantity of molecules required to form the monolayer is called the *monolayer capacity*, symbolized by q_m.

If the number of molecules required to form a monolayer can be determined and it is known how much surface area is occupied by each molecule at the experimental temperature, then the surface area of the solid is revealed simply by multiplying these two numbers. The first challenge, however, it to develop a method that will yield the monolayer capacity from the experimental data set.

Langmuir Theory

The Langmuir model assumes that only a single layer of molecules can adsorb on the solid surface. When the adsorptive gas is first introduced, the surface is bare and many adsorption sites are available on which to adsorb, therefore, adsorption proceeds rapidly. As the surface becomes more densely covered, fewer surface sites are available, and the rate of adsorption decreases since the probability of a molecule randomly colliding with an available site is greatly diminished.

An equation describing the Langmuir isotherm can be derived, as follows, from information about the adsorption process previously presented.

Let θ represent the fraction of the monolayer that has been formed; $(1 - \theta)$ then represents the fraction of the surface remaining available for adsorption. The rate at which adsorption occurs is proportional to the number of molecules in the volume of the container (i.e., pressure, P), and the fraction of bare surface. Therefore,

$$\text{Rate of adsorption} = k_1(1 - \theta)P \tag{5}$$

where k_1 is the proportionality constant.

As already descried, a molecule resides on the surface for only a short time, so for a unit area of coverage, molecules will be liberated at some rate of desorption, k_2. Thus, for a given fraction of monolayer coverage, θ,

$$\text{Rate of desorption} = k_2\theta. \tag{6}$$

When adsorption equilibrium is achieved, the rate of adsorption and desorption are the same and can be expressed as,

$$\begin{aligned} k_1(1 - \theta)P &= k_2\theta \\ \theta &= k_1P/(k_2 + k_1P) \\ &= bP/(1 + bP) \end{aligned} \tag{7}$$

where b equals k_1/k_2.

Clearly, the quantity of gas that has adsorbed on the surface after the ith dose of adsorptive is proportional to the fraction of surface coverage. Likewise, the same is true at the completion of monolayer coverage, where the quantity of gas adsorbed is symbolized by q_m.

$$\begin{aligned} q &= k_3\theta \\ &= k_3bP/(1 + bP) \\ &= k_3P/(1 + bP) \end{aligned} \tag{8}$$

where $a = k_3b$.

This equation describes the Type 1 (Langmuir) isotherm. The equation can be rearranged into linear form by first dividing both sides by P, then taking the reciprocal. This yields,

$$P/q = (1/k_3) + (b/k_3)P. \tag{9}$$

The values of k_3 and b are constants related to the gas–solid system and the experimental temperature. A plot of experimental values of P/q vs. P will yield a straight line if the adsorption mechanism conforms to the Langmuir theory. One may find that linearity is evident only over a specific pressure range. The linear region allows the evaluation of b/k_3 and k_3, the slope and y-intercept, respectively, leading to a numerical value for q even if the experimental pressure was not extended sufficiently to achieve monolayer coverage. However, when monolayer coverage is achieved in the Langmuir model, it is apparent from the isotherm, which parallels the x-axis because no further buildup of layers is permissible. Extending the flat region back to the y-axis directly yields the monolayer capacity.

Multi-Point BET Theory

Brunauer, Emmett, and Teller proposed a theory that accounts for the isotherms of Types II and III. In their theory, the forces responsible for condensation of gas are also responsible for binding gas molecules in multimolecular layers. Furthermore, BET theory, as it has come to be known, permits the second and greater layers to begin formation prior to the completion of preceding layers. As prescribed by basic theory of physical adsorption, molecules adsorb and desorb at various rates until equilibrium is established. The same holds for each layer in multilayer adsorption, so the quantity of molecules adsorbed when the system is equilibrated must be obtained by summing for an infinite number of layers. This leads to

$$V = \frac{v_m CP}{(P^0 - P)\left[1 + (C-1)\frac{P}{P^0}\right]} \tag{10}$$

where V is quantity of gas adsorbed at P/P^0 and expressed as a gas volume at STP, v_m the monolayer capacity also expressed in *standard volume* terms, and C is a constant related to the *heat of adsorption*, which is the energy liberated when a molecule adsorbs.

Rearranging Equation (6) into linear form gives

$$\frac{P}{V(P^0 - P)} = \frac{1}{v_m C} + \left(\frac{C-1}{v_m C}\right)\frac{P}{P^0} \tag{11}$$

If the adsorption process conforms to the BET model, a plot of

$$\frac{P}{V(P^0 - P)} \text{ vs. } \frac{P}{P^0} \tag{12}$$

will yield a straight line, particularly in the "BET range" of approximately $0.05 - 0.30$ P/P^0. The slope of the line will be

$$\left(\frac{C-1}{v_m C}\right) \tag{13}$$

and the intercept

$$\frac{1}{v_m C} \tag{14}$$

permitting the values of v_m and C to be determined.

Single-Point BET Theory

In some instrumental applications it is difficult or inconvenient to collect a series of V_a vs. P/P^0 data points. In such cases a single point near the upper limit of the linear range is collected and Equation (11) is modified to accommodate a single point in the following manner.

Recognizing that the intercept term of Equation (11) is generally small compared to the slope, it may be approximated as insignificant, thereby forcing the linear plot of Equation (11) through the origin, but changing the slope very little. This is equivalent to assuming that $1/V_0 C = 0$, or that $C >> 1$. If $C >> 1$, then $C - 1 \approx$ C. Making these substitutions into the right side of Equation (11) yields the BET single point relationship

$$P/[V_a(P^0 - P)] = (1/V_0)(P/P^0) \tag{15}$$

which is an approximation of the BET model. The speed and often the convenience at which a single data point is collected (as opposed to collecting several data points) is achieved at the cost of the inherent error introduced by the single-point method.

Determining Surface Area from the Monolayer Quantity

The volume of the monolayer having been determined allows the surface area of the sample to be determined simply by multiplying the area occupied by a single adsorbate molecule by the number of molecules in the monolayer, or

$$\sigma = (4)(0.866)[M/4(2N_A\rho)^{0.5}]^{0.666} \tag{16}$$

where σ is the mean area per molecule, M the molecular weight, N_A Avogadro's number, and ρ the density of the liquid adsorbate. There is not consensus on the surface area of a solid occupied by a single adsorbed molecule of a specific species at a specific temperature primarily because the area depends on the structure of the solid surface itself. In the absence of specific contrary information, typical values of 16.2 $Å^2$ for the area occupied by a nitrogen molecule and 21.0 $Å^2$ for krypton at LN_2 temperature, 14.2 $Å^2$ for argon at liquid argon temperature, and 17.0 $Å^2$ for carbon dioxide at ice water temperature suffice. For a compendium of values for various gases at various temperatures, the reader is referred to McClellan and Harnsberger (5).

Data Reduction Theories Pertaining to Porosity

Micropores are those having openings less than 20 Å (2 nm) in diameter. Currently, porosity in this size range is rarely encountered in pharmaceutical materials, however, nomaterial research may change that. Due to the current rarity of microporous pharmaceutical ingredients, analytical methods of quantifying microporosity is covered very briefly at the end of this section.

Most materials used in drug development and finished pharmaceutical products contain meospores and macropores. Mesopores generally are defined as those having widths between 20 and 500 Å (2 and 50 nm) and macropores those with widths greater than 500 Å. Analyzing mesoporous and macroporous materials is the main topic of this section.

Methods of Characterizing Mesoporous and Macroporous Materials

It is well established that the pore space of a mesoporous solid fills with condensed adsorbate at pressures somewhat below the prevailing saturated vapor pressure of the adsorptive. When combined with a correlating function that relates pore size with critical condensation pressure, this knowledge can be used to characterize the mesopore size distribution of the adsorbent. The correlating function most commonly used is the Kelvin equation. Refinements make allowances for the reduction of the physical pore size by the thickness of the adsorbed film preexisting when the critical condensation pressure is achieved. Still further refinements adjust the film thickness for the curvature of the pore wall.

This section explores both the classical application of the Kelvin equation and more modern computational approaches.

Kelvin equation: Kelvin (6) derived an expression describing the spontaneous filling of a cylindrical capillary with condensed liquid (capillary condensation) at a pressure below the bulk saturation pressure P^o of the gas phase, this critical pressure P^* being dependent on the radius of the meniscus formed by the condensate. The derivation

assumes an ideal gas and incompressible liquid phase and a well-defined separation between liquid and gas phases.

The Kelvin equation usually is written

$$\ln(P^*/P^0) = -(2\gamma v \cos\theta)/RTr_m \tag{17}$$

where P^* is the critical condensation pressure, γ the liquid surface tension, v the molar volume of the condensed adsorptive, θ the contact angle between the solid and condensed phase (taken to be zero when the adsorptive is nitrogen, hence $\cos\theta = 1$), r_m the mean radius of curvature of the surface of the liquid meniscus, and P^*/P^0, R, and T as used previously. The value of r_m is determined by the equation

$$\frac{2}{r_m} = \frac{1}{r_1} + \frac{1}{r_2} \tag{18}$$

where r_1 and r_2 are the radii of the curvature of the three-dimensional surface of the meniscus in two perpendicular planes. For a meniscus in a right circular cylinder or radius r, $r_1 = r_2 = r$ and Equation (18) becomes

$$r_m = r \tag{19}$$

Therefore, the relationship between the pressure and capillary radius determines if capillary condensation will or will not occur, P^* being dependent upon r_m.

BJH method (and variations) employing Kelvin's equation: The calculation method for determining pore size distribution using the Kelvin equation follows generally that described by Barrett et al. (7), hence, it is called the Barrett, Joyner, and Halenda (BJH) method. The mathematics of the technique is equally applicable whether following the adsorption branch of the isotherm downward from high to low pressure or following the desorption branch. In either case the condition is set arbitrarily that all pores are considered to be filled. Therefore, experimental data up to at least 99.5% relative pressure ($P/P^0 = 0.995$) must be available.

The general procedure for calculating pore size distributions using the Kelvin equation was elucidated by Gregg and Sing (8). It can be illustrated by imagining a stepwise emptying of condensed adsorbate from pores as the relative pressure is likewise decreased. It is apparent from previous discussions of adsorption theory that all pores, whether emptying or filling with condensate, have some degree of adsorbate coverage on their walls. These molecules form a film of statistical thickness t on the surface. The value of t is derived from thickness equations or from reference isotherms, and is a function of P. Therefore, at the molecular level, it is important to recognize that when pressure is decreased by a step ΔP, evaporation from some pores will occur, from exactly which pores depends on the curvature of the meniscus of the condensate as described by Kelvin. However, after evaporation, there will remain a film of condensate on the pore walls as described by the thickness equations. Thus, only the core of the pore evaporates at the critical pressure and not the entire pore volume. This varies from the macroscopic view of the Kelvin equation in which the radius of the core condensate and the radius of the capillary are considered equal (Equation 19). When working with small pores, r_m in the Kelvin equation relates the core radius r_k and not the pore radius r. The pore radius is equal to the core radius plus the adsorbed layer thickness, t.

To simplify the following discussion of the BJH method, Equation (17) is rearranged and regrouped, yielding

$$r_k = -\mathrm{K}/\ln(P_i/P^0) \tag{20}$$

where K is a constant factor representing $(2\gamma v \cos \theta)/RT$, and P_i is the experimental pressure after step i.

For the first step only, the amount of adsorptive evaporated, V_1, represents the total volume of the cores of pores that emptied during the pressure step from $P_{\max}$ to P_1. The thickness of the adsorbed layer remaining on the pore walls, t_1, is calculated from the thickness equation at P_1/P^0. With the substitution of $r_k = r_1 - t_1$ in Equation (18), a value for pore radius r_1 is calculated ($r_1 = r_{k1} + t_1$).

The first pressure reduction step opened the core of some larger pores leaving a film of condensate on the pore walls. Subsequent pressure reduction steps cause both the emptying of smaller pore cores and a reduction in the thickness of the film on the walls of pores from which cores previously were evaporated. For example, the liquid volume V_2 of adsorptive evaporating and rejoining the bulk gas as the result of pressure reduction step 2 represents the sum of core volumes V_{k2} emptied plus the volume V_{f2} of condensed film that evaporated when the thickness of the adsorbed film is reduced from t_1 to t_2.

A distribution of pore volume or area over pore width is obtained after the above-described process is completed for all steps $i = 1$ to n, concluding at minimum pressure P_n. Performing such a long series of calculations was a tedious and time-consuming task when the procedure first was developed, but today it is accomplished quickly by computer. Now, any of a number of thickness expressions can be surveyed readily, as well as working with pore shapes other than cylindrical. Among the more popular alternate pore models are those of slits for plate-like material, and of cavities formed by packed spheres such as the case with sintered objects.

The Kelvin equation (Equation 17) is enlightening with regard to hysteresis as noted previously in the Types IV and V isotherms. In a straight capillary open at both ends, the mean radius is related to the two primary radii r_1 and r_2, by

$$\frac{1}{r_m} = \frac{1}{2r_1} + \frac{1}{2r_2} \tag{21}$$

Only radius r_1 is finite when pores are filling ($r_2 = \infty$), hence r_m in Equation (21) equals $2r_1$ during filling. However, when cores are evaporating, $r_m = r_1 = r_2$. Consequently, the Kelvin equation has different values for the parameter r_m during the adsorption and desorption processes for the same pore size. Thus, when all pores are indeed open-ended and cylindrical, and when Equation (21) is incorporated, Equation (17) can be rewritten

$$\ln(P/P^0) = -\gamma v/RT(r - t) \tag{22}$$

for the adsorption branch and

$$\ln(P/P^0) = -2\gamma v/RT(r - t) \tag{23}$$

for the desorption branch. These two expressions differing by a factor of 2 have been shown by Orr (9) to be appropriate based on experimental data for the rare case of a membrane with many nearly uniform but quite small round holes through it. A distinction between the two equations is neither possible nor justified in the much more common occurrence of pores created chaotically that turn, branch, intersect, and come in all manner of sizes and shapes.

The BJH method provides the most reliable data for pore size distribution when the shape of the pore is cylindrical. However, the BJH method and capillary condensation theory do not apply when the pore size is smaller than about 20 Å, that is, in the

micropore size range. With these small pores, a completely different filling mechanism prevails. However, since pharmaceuticals seldom are microporous, classical theories and models of micropore filling will not be covered.

Density functional theory: In addition to the BET and BJH methods described above, a number of different data reduction methods are in use for extracting information from the physical adsorption isotherm. Each is applicable only to particular types of isotherms and, more specifically, to limited pressure regions of these isotherms. Traditional adsorption theories attempt to describe experimental adsorption isotherms with an isotherm equation containing a small number of parameters. At a minimum, these parameters include the extent of the surface, such as the monolayer capacity (V_m), and the intensity of the gas-surface interaction, such as the BET C constant.

A more modern approach to describing the isotherm is to use a molecular-based statistical thermodynamic theory that allows relating the adsorption isotherm to the microscopic properties of the system: the fluid–fluid and fluid–solid interaction energy parameters, the pore size, the pore geometry, and the temperature.

The stepwise dosing and subsequent adsorption of a gas was described at the beginning of this chapter as a means to explain the analytical process involved in collecting a set of data that describes an isotherm. As presented, the gas molecules randomly approach the solid surface where they come under the influence of an external attractive force (dispersion forces or van der Waal's forces) and this force causes the gas molecules, on average, to spend more time near the surface than in the bulk. As a result, at equilibrium the space near the surface has acquired a greater average density of gas molecules than regions farther removed.

If the equilibrium distribution of the gas molecules near the surface can be described as a function of system pressure and the molecular properties of the components of the system, then a model can be constructed for the adsorption isotherm for the system. Modern physical chemistry provides several ways to calculate this distribution. All these methods are based on the fundamental thermodynamic law that such a system will adopt a configuration of minimum free energy at equilibrium. In addition, a description is needed of the pair-wise interaction energy between atoms, $U(s)$, usually given by a Lennard–Jones potential:

$$U(s) = 4\varepsilon[(\sigma/s)^{12} - (\sigma/s)^{6}] \tag{24}$$

where ε is the characteristic energy of the adsorptive, σ the diameter of the adsorptive molecule, and s is the separation distance.

Two calculation methods are commonly used to determine the distribution of gas molecules in a system in equilibrium: the molecular dynamics method and the Monte Carlo method. Both of these are used as reference methods because their results are considered exact for the modeled conditions. The position and velocity of individual gas molecules (typically referred to as particles in statistical thermodynamics) are calculated in the molecular dynamics method over very short time intervals, typically 10^{-14} seconds. Although the mathematics are simple, the number of calculations required for a system of even a modest number of particles is immense and challenges even the fastest computers. Monte Carlo simulations require considerably less computation time than molecular dynamic simulations and can yield the same results; however, neither method provides a practical way to calculate complete isotherms. Density functional theory (DFT) offers a practical alternative to both molecular dynamic and Monte Carlo simulations. When compared to reference methods based on molecular simulation, this theory provides an accurate method of describing inhomogeneous systems yet requires fewer calculations.

Because the theory provides accuracy and a reduced number of calculations (thereby being practical for typical desktop computers), it is the basis of the technique embodied in DFT data reduction algorithms.

Background on the application of DFT to the adsorption process is described by Tarazona and Evans (10); Seaton et al. (11); and Peterson et al. (12). Solution of the equation of state allows a prediction of the adsorption isotherm for porous solids and leads to a method of characterization.

Ultimately, the mathematical process yields the equilibrium density profile. The quantity adsorbed per unit area of surface is obtained by integrating the equilibrium density profile over the spatial coordinate*s* and subtracting the quantity of adsorptive that would be present in the absence of surface forces (i.e., the contribution of the bulk gas). Since analytic solutions are not possible, the problem must be solved using iterative numerical methods. Although calculation using these methods still requires exceptional computing speed, the calculation of many isotherm pressure points for a wide range of materials with various surface features is a feasible task.

Applying the above process to find the equilibrium density profile over an analytical pressure range from ultra low to saturation pressure while maintaining constant surface features is required to generate a single model isotherm for a specific material with specific surface features. Generating a set of model isotherms for a range of pore sizes requires incrementing pore size from about the size of the gas molecule (a few angstroms) up to a free surface (essentially, non-porous), and repeating the series of calculations for each pore size over the pressure range.

For specific bath temperatures, adsorptive molecules, substrate material, and pore shapes, Olivier and Conklin (13,14) and Olivier et al. (15) have generated sets of model isotherms. Examples are nitrogen on carbon at 77 K, argon on carbon at 87 K, CO_2 on carbon at 273 K, all these examples being slit pore models.

It should be noted that, unlike some classical methods for micropore and mesopore analysis, the Olivier–Conklin method is neither calibrated for nor biased in any way toward a pore of a particular size or a size distribution of a particular type. A significant feature is that the DFT method applies over the complete range of the isotherm and is not restricted to a confined range of relative pressures or pore sizes as are the classical models.

Methods for the Analysis of Micropores

The Type I isotherm shown in Figure 2 is associated with microporosity. Note that the uptake of the adsorptive gas is initiated and completed in the low pressure range of the isotherm. This is because micropores fill spontaneously rather than building up layers of adsorbent over a wide range of pressures.

To detect the nuances of the isotherm in the pressure range in which micropores fill requires specialized adsorption equipment that is capable of achieving very low pressures, maintaining these pressures over extended lengths of time and detecting minute changes in pressures. Additionally, the equipment must be able to deliver small doses of adsorptive to the sample.

The Kelvin model does not apply to micropores, therefore neither does the BJH method. The DFT method, previously discussed, is applicable and is rapidly becoming the preferred method for probing micropores. Other data reduction methods include those of Dubinin–Radushkevich (16), Dubinin–Astakhov (17), and Horvath and Kawazoe (18).

DETERMINATIONS OF POROSITY AND DENSITY BY MERCURY INTRUSION

Mercury intrusion porosimetry is one of only a few analytical techniques that is applicable over such a broad dynamic range using a single theoretical model. Mercury porosimetry routinely is applied over a pore diameter range from 0.003 to 360 micrometers—five orders of magnitude.

The dynamic range of the mercury intrusion technique is only one of many advantages of this measurement technique. The fundamental data it produces, volume of mercury intruded into the pores space as a function of applied pressure, is indicative of various characteristics of the pore netword and also is used to reveal a variety of physical properties of the solid material itself.

As with physical adsorption, understanding how the fluid behaves under specific conditions provides insight into how a mercury porosimeter probes the surface of a material and moves within the pore structure. This allows one to better understand what mercury intrusion and extrusion data mean in relation to the sample under test and allows one to understand the data outside of the bounds of the theoretical model. It also allows one to make an educated comparison between data obtained for the same sample using other measurement techniques such as physical adsorption.

The Intrusion Phenomenon

A drop of liquid placed on a solid surface either will contract into a bead, or will flatten out over the surface. In the first case, the liquid is considered to be a *non-wetting liquid* for the solid and in the second, a *wetting liquid*. Examples are mercury beading on a glass surface and water spreading over the same surface.

If one end of a capillary tube (a solid) if forced to penetrate the surface of a liquid, one of two things will happen. If the liquid is a wetting liquid, it will spontaneously enter the capillary and rise to a level above the surface of the bulk liquid. If a non-wetting liquid, it will resist entering the capillary. Only when the end of the capillary is submerged sufficiently deep to experience the necessary head pressure will a non-wetting liquid enter the capillary and it will rise to a level always below the surface of the bulk liquid. The relevant observation is that a force must be applied to a non-wetting liquid to influence it to enter a capillary.

If the above experiment with the non-wetting liquid is repeated with capillaries of various diameters, it will be found that it is necessary to push the smaller capillary tubes deeper into the liquid (increase head pressure) before the liquid enters the capillary. The results suggest that there is an inverse relationship between the applied force and the size of the capillary that the non-wetting liquid will enter.

A Mercury Intrusion Experiment

Imagine the following experiment. A porous solid (essentially a matrix of capillaries of different diameters and lengths) is placed into a vessel and the vessel sealed. By way of a valve, air in the remaining void space of the vessel is removed and the vacuum valve is closed. By way of another valve connected to a mercury reservoir, mercury is allowed to enter the vessel and fill the accessible voids. Under the described conditions, mercury will bridge the opening of all pores smaller than about 12 micrometers diameter and completely fill those larger since there is no resisting atmospheric pressure within the pores.

As was learned from previous experiments, for mercury to enter the smaller pores, an external pressure must be applied; increasing pressure in the mercury reservoir accomplishes this. Assume that pressure on the reservoir is monitored as well as the volume of mercury in the reservoir.

Upon the first increasing pressure step, mercury will be forced into any pores of the appropriate size, which will be somewhat smaller than those already filled. As mercury enters this set of pores, mercury from the reservoir replaces it so that the sample vessel remains full of mercury. The current pressure, P_1, is recorded as well as the volume (V_1) of mercury that was removed from the reservoir. This provides the first ordered pair of experimental data points, (P_1,V_1), where V_1 is the intrusion volume and also the volume of the pores that were filled.

The pressure is again increased and the intrusion volume determined. This process continues until there is clearly no more intrusion occurring as pressure is increased. A plot of these points is called an *intrusion curve*. If the pressure is decreased in a stepwise manner and measurement made, it will be observed that mercury leaves the pores in the same order they were filled and the mercury is returned to the reservoir. A plot of those data produce an *extrusion curve*. When examining the two curves, it will be noted that the extrusion curve did not retrace the intrusion curve.

Repeating the experiment with several different porous materials yields a wide variety of shapes for the intrusion and extrusion curves. Clearly, within these data is information about the pore structure of the sample. Before that information can be extracted, considerably more must be known about the intrusion and extrusion processes.

Intrusion Theory

Inside a capillary, the liquid–solid interface assumes an angle that results in equilibrium between the relative magnitude of the forces of cohesion between the liquid molecules and the forces of adhesion between the liquid molecules and the walls of the capillary. This is known as the *contact angle* and is characteristic of the specific solid–liquid interface. The liquid–vapor interface in the capillary (the meniscus) will be concave for a wetting liquid and convex for a non-wetting liquid.

Washburn (19) in 1921 derived an equation describing the equilibrium of the internal and external forces in terms of the surface tension of the liquid, the contact angle between the liquid and solid, and the cross-sectional shape of the capillary. For simplicity, the latter is usually assumed to be a circle. The equation states simply that the pressure required to force a non-wetting liquid to enter a capillary of circular cross-section is inversely proportional to the diameter of the capillary and directly proportional to the surface tension of the liquid and the angle of contact with the solid surface.

Mercury is used almost exclusively as the analytical liquid in porosimetry and there are several good reasons. The primary one is that mercury does not wet the majority of substances, thus will not penetrate pores by capillary action—it must be forced to do so. Another attribute of liquid mercury is its high surface tension, usually taken to be 485 dyne/cm. Mercury also exhibits a high contact angle at the interface with most solids, in most cases ranging from 112° to 142°, with 130° being the most widely accepted. Mercury is a metal and, therefore, conducts electricity. Although this is not important in regard to intrusion, it is very significant in regard to metering the quantity of mercury moving into and out of the pores.

When mercury is in contact with a pore opening of circular cross-section and diameter D, the surface tension of the mercury acts along the circle of contact over a length equal to the perimeter of the circle, which is πD. Thus the force opposing the entry

of mercury into the pore equals $-\pi D\,\gamma\cos\theta$, where γ is the surface tension of mercury and θ the contact angle between the mercury and solid. An external pressure is applied to overcome the resistive force and cause intrusion of the mercury into the pore. Since pressure is defined as force per unit area ($P = F/A$), it follows that the total force produced by a pressure is pressure multiplied by the area upon which the pressure is applied. The pressure promoting intrusion acts over the area of the circular pore opening ($\pi D^2/4$), which the mercury bridges; the intrusion force, then, is $(\pi D^2/4)P$. At equilibrium the intrusion force and the force opposing entry are equal; thus

$$-\pi D\gamma\cos\theta = \frac{\pi D^2 P}{4} \tag{25}$$

or, simplified

$$D = \frac{-4\gamma\cos\theta}{P} \tag{26}$$

which is the Washburn equation.

The minimum size pore that can be probed with a porosimeter depends upon the capability of the porosimeter to generate high pressures. Assuming the surface tension of mercury is 485 dyne/cm and the contact angle is 130° and the maximum applied pressure is 414 MPa (60,000 psia), the upper limit of pressure for most commercial mercury porosimeters, Equation (26) reveals that mercury will enter pores down to 0.003 micrometers (30 Å or 3 nm) diameter At ambient pressure, pores of about 12 micrometer and larger are already filled, so to work with pores above this size, the system must be evacuated. At 0.0034 MPa (0.5 psia), only pores larger than 360 micrometers in diameter are filled.

The general assumption that pores are cylinders of different diameters is a simplification that produced a readily known equation by which to express the perimeter of the pore opening Another pore shape for which there is a simple equation is that of a slit. Slit pores arises from materials composed of stacked, thin sheets. For slits of unlimited dimensions in all but their width, the same derivation that led to Equation (26) would lead to

$$W = \frac{-2\gamma\cos\theta}{P} \tag{27}$$

where W is the width between the plates. In subsequent discussions, cylindrical pores are assumed.

Extracting Information about the Sample Material from Intrusion and Extrusion Curves

Envelope, Bulk Volume, and Density

The first category of information that can be extracted from mercury intrusion porosimetry data does not depend on the shape of the intrusion curve nor Washburn's equation, but are derived simply from measurements of masses and volumes.

In the section, Fundamental Measurements, an experiment was imagined in which a porous solid was placed in a sample vessel (called a *penetrometer*; Fig. 3), the penetrometer evacuated, and mercury introduced to fill the accessible voids. Mercury enveloped the solid, but only filled the largest pores. This is the beginning point of a mercury intrusion analysis and this starting point provides an opportunity to determine the *envelope volume* of the sample. With the sample mass being known, *envelope density* also can be determined.

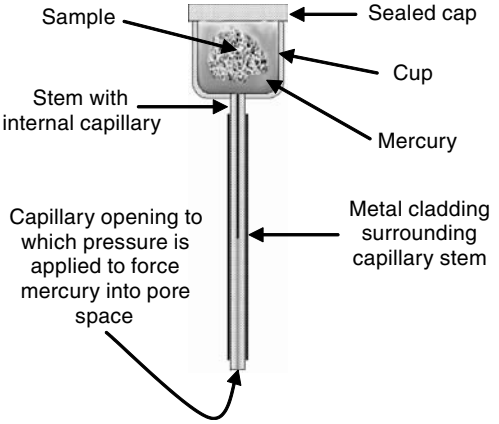


FIGURE 3 A penetrometer used in the measurement of mercury intrusion. The penetrometer is not only a sample holder, but also a measuring device. When initially filled with mercury, not only is the sample cup filled to surround the sample, but the capillary in the stem is filled. This acts as a reservoir for mercury that is forced into pores during the analysis. The combination of the mercury and the metal cladding surrounding the stem creates a capacitor. Any change in the volume of mercury in the stem results in a proportional change in capacitance. Therefore, measuring the change in capacitance is analogous to measuring the volume of mercury moving out of the stem and into the pore space of the sample.

Had the sample material been a fine powder, essentially the same conditions would prevail in the penetrometer. Mercury would surround the sample bulk, but would not penetrate into the interparticle voids because the initial pressure is too low to force mercury into them. In this instance, the conditions allow determination of *bulk volume* and *bulk density*.

Envelope and bulk density determinations by mercury porosimetry require finding the total volume of the sample before pores or interstitial voids are filled. The volume of the sample material is the volume of the empty sample penetrometer minus the volume of mercury required to fill the penetrometer when the sample is included. Dividing the sample weight by this volume difference provides either the envelope or bulk density, depending on the form of the sample material.

Determining sample volume and bulk or envelope density by this method requires measurements of the weight of the empty penetrometer Wv, the weight of the sample Ws, and the total weight of the penetrometer W with the sample loaded and filled with mercury. The weight of the mercury W_{Hg} contained in the penetrometer is the total weight minus the sample and empty penetrometer weights. Dividing by mercury density ρ_{Hg} gives the volume of mercury V_{Hg}, the mathematical expression being,

$$V_{Hg} = \frac{W_{Hg}}{\rho_{Hg}} = \frac{W - W_p - W_s}{\rho_{Hg}} \tag{28}$$

If V_p is the volume of the empty penetrometer, the envelope volume of the sample V_{se} is the volume of the penetrometer minus the volume of the mercury. The envelope density of the sample ρ_{se} is then

$$\rho_{se} = \frac{W_s}{V_p - V_{Hg}} \tag{29}$$

Skeletal Volume and Density

Described thus far are sample characteristics that can be obtained at the lowest pressure before development of the intrusion curve begins. Other volume and density characteristics can be determined at the highest pressure value after the intrusion curve is completed (all pores are filled with mercury).

The first determination at high pressure is the *skeletal volume* of the sample, V_S. This can be determined by subtracting the total pore volume from either the envelope or bulk volume of the sample, depending on which was obtained initially. The total pore volume is the total volume of mercury, V_{Hg}, injected into the sample material between the first low pressure data point on the intrusion curve and the last point collected at the maximum attainable pressure. Dividing the weight of the sample by skeletal volume gives the *skeletal density* ρ_s of the sample, expressed in a general equation by

$$\rho_s = \frac{W_S}{V_S - V_{Hg}} \tag{30}$$

Percent Porosity

After data at the highest pressure has been collected, the percent porosity of the sample material can be determined as follows

$$\text{Porosity}\,(\%) = \left(1 - \frac{\rho_s}{\rho_{se}}\right) \times 100 \tag{31}$$

Pore Volume and Pore Area Distributions by Pore Diameter

The next category of information that is available from mercury porosimetry pertains to pore sizes and volumes based on characteristics of the intrusion curve. The raw experimental data are reduced by application of the Washburn Equation. Plots of mercury porosimetry data are presented in Figure 4 with explanations for characteristics in their shapes.

Cumulative pore volume vs. pore diameter is immediately obtainable from application of Equation (26). Likewise incremental pore volumes are obtained by differentiation. Pore wall area A is related to pore volume V by $A = 4V/D$ when the pores are taken to be right cylinders. This model is used to calculate cumulative and incremental pore wall areas. Since pore area is related to pore length L by $L = A/\pi D$, total cumulative and incremental pore lengths can be obtained. The pore areas and lengths for each interval are summed over all pores in the interval.

In some instances, when the sample is a film or sheet, for example, the length of pores in a sample may be estimated with some degree of certainty. In these cases, the number of pores N in an interval can be calculated by $N = V_T/V$, where V_T is the total volume of all pores in the interval, and V the volume of one pore calculated using a diameter representative of the size interval (average diameter, for example) and the estimated length.

Total pore volume per weight of sample—the specific pore volume—is the maximum volume of mercury penetrated into the sample at the highest pressure. Likewise, total pore area and length are the accumulated wall areas and lengths at the highest pressure as calculated from the assumed pore model, typically a right cylinder. Median pore diameter is that at the 50 percentile point on any volume, area, or length distribution

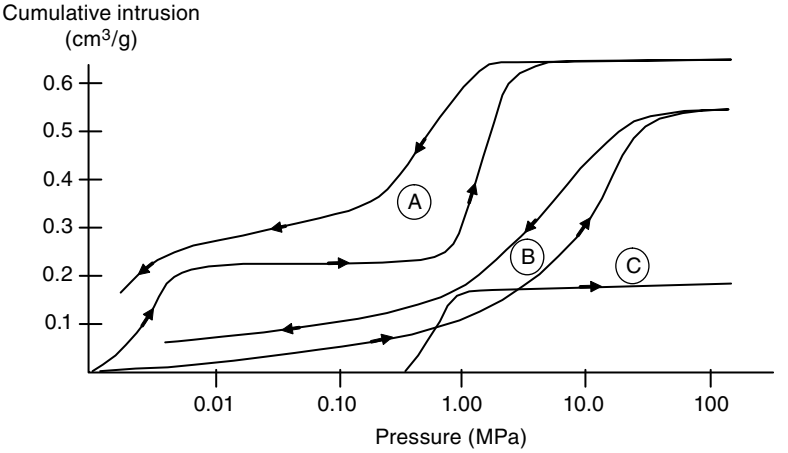


FIGURE 4 Examples of intrusion and extrusion curves. Curve A is typical of a coarse grained sample bed. The relatively steep initial rise at low pressure is due to intrusion into inter-particle voids, and the second rise is due to filling of the pores within the individual grains. Curve B is a single piece of material in which there is a wide distribution of pore sizes. Curve C is a fine powder essentially without pores and the volume indicated is due entirely to filling of interparticle voids. The extrusion curve is indicated by the arrows pointed in the direction of lower pressure. That the mercury is not fully expelled is primarily due to entrapment within bottle-necked pores.

curve. The average pore diameter depends on the model, but, when the model is assumed to be a cylinder, it is equal to $4V/A$.

Particle Size Distribution and Other Characteristics of the Sample

Over time, new theories have emerged for extracting from the intrusion and extrusion curves various types of information beyond that described above. Examples include fractal dimensions of the pore volume distribution, pore *tortuosity* and *tortuosity factor*, pore shape and material *permeability*. Because of the high pressures available (up to 60,000 psi) and the sensitivity of the instrument to small changes in mercury volume, the mercury intrusion porosimeter also can be used to study the compressibility and restitution of materials.

An interesting application of mercury intrusion and one that analyzes the low pressure region of the intrusion curve to extract information about particle size distribution. The method was developed by Mayer and Stowe (20,21), extending the works of Frevel and Kressley (22) and Pospech and Schneider (23). The model is based on the penetration of fluids into the interstitial voids in a bed of uniform nonporous spheres. The model accommodates a range of three-dimensional packing from close packing to simple cubic packing. The pressure required to force mercury into the interparticle spaces of the bed (the "breakthrough" pressure) is expressed as a function of the packing geometry. Their model defines the geometry in terms of a single acute angle σ which describes the rhombohedron produced when connecting the centers of the spheres that cluster to form the interstitial cavity.

Mayer and Stowe were able to derive an equation that relates the "breakthrough pressure" not only to the size of the access opening, but also to the radii of the spheres forming the cavity. Using the same physical parameters as in porosity determinations and including density, the Mayer–Stowe method reveals the percent mass distribution by size for the sample material. Although mercury porosimetry is not a common technique for determining particle size distributions, it may be the only technique that can provide particle size information on strongly agglomerated materials.

For the determination of bulk and envelope volumes, a mercury porosimeter is used in the manner of a simple displacement device, applying Archimedes displacement method. The same method is applied to determine absolute volume, but more sophistication is required of the instrument to fill the pores and to determine how much fluid entered the pore space. Once volumes are determined, the associated densities follow. Total porosity is determined from the difference between bulk or envelope volume and absolute volume, the assumption being that all pores in the sample material communicate with the surface and no or negligible "*blind*" *pores* exist.

VOLUME, DENSITY, AND POROSITY DETERMINATIONS BY OTHER ANALYTICAL TECHNIQUES

There are two additional displacement type automated analytical instruments that can determine the same volume dimensions as a mercury porosimeter when used either separately or in conjunction; both are classified as pycnometers since they primarily determine volume.

The Gas Pycnometer

The most popular pycnometer for determining the skeletal volume of solids is the gas pycnometer. Helium is the most common gas used as the displacement fluid because of its capability to invade extremely small pores at low pressure (approximately 20 psia). Since the volume it determines excludes all open pores, it determines skeletal volume and, when the sample mass is included, it also provides skeletal density values.

The primary measurement is that of pressure change. As advised in the section on physical adsorption isotherm measurements, which also depends on pressure measurements, the sample material must be properly prepared before reliable data can be obtained. Sample preparation requirements for analyses by gas pycnometry is not as rigorous as that when gathering gas adsorption data, but it is important none the less. The most important preparation steps are to assure that all moisture is removed and that no volatile components are associated with the sample. In either case, pressure measurements will be affected by the outgassing of these vapors and, particularly in the case of water vapor, sample weight will be affected. Although best suited for solid samples, pastes, slurries, and liquids having low vapor pressures can be analyzed. In the case of a slurry, the instrument is capable of determining the percent solid concentration. Also, by a series of measurements, the ratio of open- to closed-cells can be determined for rigid foams.

There are two volumes associated with a gas pycnometer, an analysis chamber of volume V_A, and an expansion chamber of volume V_E. The precise volumes of these

chambers is determined by use of a calibration volume, traceable to an ISO, NIST or other standard organization. Very basically, an analysis is performed as follows.

A dry sample is placed into the analysis chamber and the chamber sealed; the free volume in the analysis chamber has been reduced by the volume of the sample, or to $V_A - V_S$. A valve connecting the expansion and analysis chambers is opened and the equilibrium pressure, P_1, determined. Next, the interconnecting valve is closed and the expansion chamber is charged to an elevated pressure, P_2, after which the interconnecting valve is again opened. Pressure in the analysis chamber increases and pressure in the expansion decreases and both equilibrate at P_2.

If no gas is lost and the temperature is constant, then, according to Boyle's law,

$$P_2(V_A - V_S + V_E) = P_1(V_A + V_E) \tag{32}$$

Expanding the left side gives,

$$P_2V_A - P_2V_S + P_2V_E = P_1(V_A + V_E) \tag{33}$$

Move the known terms to the right side,

$$P_2V_S = P_2(V_A - V_E) + P_1(V_A + V_E) \tag{34}$$

and divide both sides by P_2, yielding

$$V_S = (V_A - V_E) + (P_1/P_2)(V_A + V_E) \tag{35}$$

which expresses the volume of the sample in terms of known variables.

Solid Medium Displacement

Another automated analytical technique used to determine volume utilizes a dry, free-flowing solid medium as the displacement "fluid." All particles of the medium are small, hard spheres. They are too large to enter pores, but sufficiently small to envelop an object in a closely conforming "skin." The apparatus consists of a cylinder in which the sample and medium are placed, and a piston that applies a selectable and reproducible force to the medium to form a compacted bed as the cylinder vibrates to augment packing.

Prior to an analysis, a compacted bed of medium is created and its baseline volume determined. The piston is withdrawn, the sample is placed in the same medium and again a compacted bed is created which encompasses the sample. The difference in the first and second bed volumes is the volume of the sample plus its pores, which is the envelope volume. The analysis technique is not sensitive to the presence atmospheric contaminants on the sample, so no special preparation is required.

With the skeletal volume known from gas pycnometry measurements and the envelope volume known from the solid displacement method, the total pore volume is derived simply by taking the difference in these two values.

The instrument also produces a bulk density determination that is, in principle, equivalent to tap density. In this application, the dry medium is not used and only the finely divided sample material is placed in the cylinder. However, rather that tapping the container to achieve compaction, the instrument is set to drive the piston forward, compacting the bed as the cylinder vibrates, until a user defined resistive force as produced by the bed. This provides a very repeatable, reproducible, and controllable way to obtain automated determinations of bulk density.

GLOSSARY

Adsorbate Gas molecules that have adsorbed on the surface of the solid
Adsorbed The condition of being retained (detained) on the surface
Adsorbent The solid material on which adsorption occurs
Adsorption An increase in the concentration of the gaseous phase at the gas–solid interface due to the influence of surface forces
Adsorption equilibrium The condition at which the rate of adsorption and desorption are equal; when the quantity of adsorbed gas no longer changes with time after a change in environmental conditions
Adsorption isotherm A plot or function which relates, at constant temperature, the quantity of gas adsorbed after pressure with the gas phase has equilibrated
Adsorptive The material in the gas phase which is in the bulk and capable of being adsorbed
BET surface area Surface area determined using the surface coverage model of Brunauer, Emmett, and Teller
Contact angle The angle between the line tangent to the liquid surface at the liquid–solid contact point and a tangent to the solid
Density Defined as mass per unit volume, however there are several definitions of "volume," each resulting a different values
Density functional theory (DFT) In the present case, DFT is a formally exact theory based on the density of a system of gas molecules surrounding a solid for which there is some degree of affinity of the gas for the solid surface
Density, bulk The mass of a collection of particles divided by the volume of collection including inter-particle voids and particle pores
Density, envelope The mass of an object divided by its envelope volume (see **volume, envelope**)
Density, particle See *density*, *envelope*
Density, skeletal The mass per unit volume of a material for which the volume excludes open porosity, i.e., the skeletal volume
Desorb To escape from the adsorption site on the solid surface
Desorption isotherm A graphical representation of a set of data points (pressure versus quantity adsorbed) measured at constant temperature as pressure is decreased monotonically
Equilibration time The time required for a system to achieve balance and cease to change in response to opposing actions. In the current context, either: (*i*) the time required for the rate of adsorption to equal the rate of desorption after a pressure change, or (*ii*) the time required for mercury to intruded into all voids that are accessible at the prevailing pressure after a positive change in pressure or to extrude from voids after a negative step in pressure
Extrusion curve A graphical representation of the cumulative or incremental volume of mercury exiting the pores of a sample as pressure is decreased monotonically
Heat of adsorption The energy liberated when a molecule adsorbs
Interpartical (interstitial) voids Void space between particles
Intrusion curve A graphical representation of the cumulative or incremental volume of mercury entering the pore space of the sample as pressure is decreased monotonically
Macropore A pore of diameter greater than about 50 nm
Mesopore A pore of diameter from about 2 nm to 50 nm
Micorpore A pore of diameter less than about 2 nm
Monolayer capacity The quantity of gas required to form a single layer of molecules on the surface of a material
Monolayer coverage When a single layer of gas molecules covers the exposed surface of a sample material; often can be identified by a particular inflection point on an adsorption isotherm
Particle density The mass per unit volume of the particle, where the volume excludes that of open pores, but includes that of closed pores

Penetrometer, mercury In the current context, a device for determining the quantity of mercury that penetrates the voids of a sample material

Permeability The rate a liquid or gas flows through a porous material

Physical adsorption A condition in which a gas (the adsorbate) is held by weak physical forces to a solid surface (the adsorbent). A increase in the concentration of a fluid near the solid surface more so that in the bulk fluid surrounding the solid

Physicochemical process Processes involving changes in both the physical properties and the chemical structure of a material

Pore diameter The diameter of a pore derived from data obtained by a specified procedure using a specific model (typically cylindrical)

Pore volume The volume of open pores unless otherwise stated

Pore volume, specific Pore volume per unit mass of material

Pore, blind (closed) A pore with no access to an external surface (also called "closed pore")

Porosity (a) The ratio of open pores and voids to the envelope volume (BSI) (b) The ratio, usually expressed as a percentage, of the total volume of voids of a given porous medium to the total volume of the porous medium (ASTM)

Porosity, interparticle Void space between particles

Porosity, intraparticle All porosity within the envelopes of the individual particles

Porosity, particle The ratio of the volume of open pore to the total volume of the particle

Porosity, powder The ratio of the volume of voids plus the volume of open pores to the total volume occupied by the powder

Specific surface area The surface area per unit mass of a material, usually expressed in square meters per gram

Standard volume The volume of gas converted under standard conditions of temperature and pressure; expressed in units of cm^3 STP

Tortuosity The ratio of the actual distance traversed between two points to the minimum distance between the same two points

Tortuosity factor The ratio of tortuosity to constriction (used in the area of heterogeneous catalysis); the distance a fluid must travel to get through a film, divided by the thickness of the film

Total surface area The total measured surface area of a material as opposed to the specific surface area which is the surface area per unit mass of the material

Volume, bulk The space occupied by an assemblage of divided particles including the solid and void components

Volume, envelope The space within a closely conforming "skin" that envelops a solid object and which includes the superficial and internal voids of the object

Volume, specific The volume of a material divided by it's mass; reciprocal of density

REFERENCES

1. Rowsell JLC, Spencer EC, Eckert J, et al. Gas adsorption sites in a large-pore metal–organic framework. Science 2005; 309:1350–4.
2. Brunauer S. The Adsorption of Gases and Vapors. Vol. I. Physical Adsorption. Princeton, NJ: Princeton University Press, 1943.
3. Langmuir IJ. The adsorption of gases on plane surfaces of. glass, mica, and platinum. Am Chem Soc 1918; 40:1361–403.
4. Brunauer S, Emmett PH, Teller E. Adsorption of gases in multimolectulr layers. J Am Chem Soc 1938; 60:309–19.
5. McClellan AL, Harnsberger HF. Cross-sectional areas of molecules adsorbed on solid surface. J Colloid Interface Sci 1967; 23:577.
6. Thomson W (Lord Kelvin). On the equilibrium of vapour at a curved surface of liquid. Philos Mag 1871; 42:448.
7. Barrett EP, Joyner LG, Halenda PP. The determination of pore volume and area distributions in porous substances. I. Computations from nitrogen isotherms. J Am Chem Soc 1951; 73:373.

8. Gregg SJ, Sing KSW. Adsorption, Surface Area and Porosity, 2nd ed., NY, 1982.
9. Orr C. Surface Area Measurement—The Present Status. Dechema–Monographien NR 1976; 79(B):1589–615.
10. Tarazona P, Marconi UMB, Evans R. Phase equilibria of fluid interfaces and confined fluids. Non-local versus local density functionals. Mol Phys 1987; 60:543.
11. Seaton NA, Walton JPRB, Quirke N. A new analysis method for the determination of the pore size distribution of porous carbons from nitrogen adsorption measurements. Carbon 1989; 27:853.
12. Peterson BK, Walton JPRB, Gubbins KE. Fluid behaviour in narrow pores. J Chem Soc 1986; 82:1789.
13. Olivier JP, Conklin WB. Presented at the 7th International Conference on Surface and Colloidal Science, Campiegne, France, 1991.
14. Olivier JP, Conklin WB. Determination of pore size distribution from density functional theoretic models of adsorption and condensation within porous solids. Presented at International Symposium on Effects of Surface Heterogeneity in Adsorption and Catalysis on Solids, Kazimier Dolny, Poland, 1992.
15. Olivier JP, Conklin WB, Szombathely M. Determination of pore size distribution from density functional theory: A comparison of nitrogen and argon results. Presented at the COPS III, 1993.
16. Dubinin MM, Radushkevich LV. The equations of the characterisitc curve of activated carbon. Proc Acad Sci USSR 1947; 55:331.
17. Dubinin MM, Astakhov VA. Description of adsorption equilibria of vapors on zeolites over wide ranges of temperature and pressure. Adv Chem Soc 1971; 102:69.
18. Horvath G, Kawazoe K. Method for the calculation of effective pore size distribution in molecular sieve carbon. Chem Eng Jpn 1983; 16:470.
19. Washburn EW. Proc Natl Acad Sci 1921; 7:115.
20. Mayer RP, Stowe RA. Mercury pososimetry—breakthrough pressure for penetration between packed spheres. J Colloid Interface Sci 1965; 20:893.
21. Mayer RP, Stowe RA. Mercury porosimetry: Filling of toroidal vopid volume following breakthrough between packed spheres. J Phys Chem 1966; 70:3867.
22. Frevel LK, Kressley L. Modifications in mercury porosimetry. Anal Chem 1963; 35:1492.
23. Pospech R, Schneider P. Powder particle sizes from mercury porosimetry. Powder Technol 1989; 59:163.

Index

f = location of figures.

t = location of tables.