Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Pharmaceutical and Biomedical Analysis 40 (2006) 3–15
JOURNAL OF PHARMACEUTICAL AND BIOMEDICAL ANALYSIS
www.elsevier.com/locate/jpba

Review

Official USP Reference Standards: Metrology concepts, overview, and scientific issues and opportunities

Roger L. Williams [*,1]
Project Team 4[2]
the 2000–2005 Reference Standards Committee of the
USP Council of Experts and Its Advisory Panel[3]
USP Staff and Consultant[4]

United States Pharmacopeia, 12601 Twinbrook Parkway, Rockville, MD 20852-1790, USA

Received 11 May 2005; received in revised form 26 June 2005; accepted 6 July 2005
Available online 21 November 2005

Abstract

The United States Pharmacopeia (USP) is a private standards-setting body created in 1820 by practitioners who wished to promote the quality of therapeutic products in commerce. The principal product of USP, then and now, is the *United States Pharmacopeia* (*USP*), to which was added the *National Formulary* (NF) in 1975. The two compendia are published as a combined text annually (*USP–NF*). Originally a book of process standards, USP–NF evolved over time into compendia containing primarily product standards that are expressed in monographs for therapeutic ingredients, products, and excipients. As a public health service, USP supplies official USP Reference Standards to manufacturers and others who wish to test an article according to selected procedures of a monograph or *General Chapter*. During the past decade, understanding of USP monographs and official USP Reference Standards as a means of controlling the quality of a therapeutic article has evolved, based on advances in metrology, on activities in the International Conference on Harmonization of Technical Requirements for the Registration of Pharmaceuticals for Human Use (ICH), and on considerations by the USP Council of Experts and its Expert Committees and USP staff. This article discusses the evolution of this understanding, focusing on drug substances and excipients for well-characterized small molecules and their corresponding dosage forms.

Keywords: United States Pharmacopeia; Reference Standards; Specifications; *USP–NF*; Metrology; Monographs

Contents

* Tel.: +1 301 816 8300; fax: +1 301 816 8373.
E-mail address: rlw@usp.org.
[1] Council of Experts Chair.,
[2] Lise Alaimo, Mark Bailey, Matthew Borer, Raymond Cox, David Fay, Gregory Kaster (Chair), Reenie Parris, and Jeffrey Singer.
[3] Reference Standard Committee: Judy P. Boehlert, Raymond A. Cox, Thomas S. Foster, Samir A. Hanna, Krishan Kumar, Pauline M. Lacroix, Guirag K. Poochikian, Susan J. Schniepp, Eric Sheinin (Chair), Manfred E. Wolff, and Wesley E. Workman; Advisory Panel: David A. Fay, Ralph Gomez, Vivian A. Gray, Ruth Homan, Maria Ines, Gregory T. Kaster, Judy A. Lee, Michael Ribick, and R.M. Santoro.
[4] Todd Cecil, Shawn Dressman, John Esker, L.Valentin Feyns , Walter Hauck (Consultant), Ronald Manning, Stefan Schuber, and Eric Sheinin.

0731-7085/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpba.2005.07.017

1. Introduction

The United States Pharmacopeia (USP) is a private standards-setting body created in 1820 by practitioners who wished to promote the quality of therapeutic products in commerce [1]. The principal product of USP, then and now, is the *United States Pharmacopeia*, to which was added the *National Formulary* (*NF*) in 1975. Together the two compendia are published as a combined text annually (*USP–NF*) with two *Supplements*. Originally a book of process standards (recipes for preparations), *USP–NF* evolved over time into compendia containing primarily product standards [2]. These standards are expressed in monographs for drug substances, excipients, dosage forms and other articles, and in *General Chapters*, which are dedicated to procedures, general information, and general requirements widely used throughout the compendium. A monograph contains introductory statements, packaging, storage, and other labeling statements, and the article's public specification, which consists of tests, procedures, and acceptance criteria. As a public health service, USP supplies official USP Reference Standards to manufacturers, regulatory agencies, and other interested parties who wish to test an article according to selected procedures of a monograph or of a *General Chapter*.

During approximately the past 10 years, advances in metrology [3], activities in International Conference on Harmonization of Technical Requirements for the Registration of Pharmaceuticals for Human Use (ICH), and considerations by the USP Council of Experts and its Expert Committees with USP staff have advanced understanding of the role of USP monographs and official USP Reference Standards as a means of controlling the quality of a therapeutic article [4]. This article discusses the evolution of this understanding, focusing on drug substances and excipients for well-characterized small molecules and their corresponding dosage forms. An Appendix A considers special statistical issues with regard to collaborative testing.

2. Nomenclature and legal implications

2.1. Nomenclature

Modern metrology concepts have developed during the past decade and longer through research and standards-setting activities in government, academia, industry, the pharmacopeias, and elsewhere. Harmonizing nomenclature for these concepts has been presented in publications of International Standards Organization (ISO) guides and also in publications of the National Institute of Standards and Technology (NIST), the International Union of Pure and Applied Chemistry (IUPAC), EURACHEM, Co-Operation on International Traceability in Analytical Chemistry (CITAC), and International Laboratory Accreditation Cooperation (ILAC). Successful application of modern metrologic approaches, when used in ratio-method measurements in the pharmacopeia, are associated with a reference material (RM), defined as "a material or substance one or more of whose property values are sufficiently homogeneous and well established to be used for the calibration of an apparatus, the assessment of a measurement method, or for assigning values to materials" [5]. A certified reference material (CRM) is defined as a "reference material, accompanied by a certificate, one or more of whose property values are certified by a procedure which establishes its traceability to an accurate realization of the unit in which the property values are expressed, and for which each certified value is accompanied by an uncertainty at a stated level of confidence" [6]. A general objective of modern metrology allows the content (purity) of an RM to be expressed in the Système International d'Unités (SI) units of mass (e.g., kilogram) and/or amount (e.g., mole) [7]. When used appropriately, RMs and CRMs allow value assignment for a measurand in SI units. Although USP offers RMs (official USP Reference Standards) for calibration and assessment of a measurement method (procedural standard), the bulk of the USP collection is composed of neat materials used in quality control and similar laboratories to assign a value—expressed in terms of mass—to measurands. These

are pharmaceutical ingredients (drug substances, excipients, and other ingredients, e.g., antimicrobial agents, preservatives, and anti-oxidants) as well as dosage forms. The use of official USP reference standards has risen over the past several decades as a result of the increasing use of primary ratio methods in instrumental techniques.

2.2. Legal implications

A complex relationship exists between USP's official publications, *USP–NF,* and the *Federal Food, Drug, and Cosmetic Act* (FDCA) [8]. For enforcement purposes, the FDCA states that a drug shall be deemed to be adulterated if it purports to be or is represented as a drug, the name of which is recognized in an official compendium and its strength differs from, or its quality or purity falls below, the standards set forth in such compendium [9]. Such determination regarding strength, quality, or purity shall be made in accordance with the tests or methods of assay set forth in such compendium. USP is not itself a regulatory body, but it provides Reference Standards that support the regulatory framework for therapeutic agents.

Reference Standards are chemicals—not drugs or devices. They are not intended for use in the diagnosis, cure, mitigation, or treatment of diseases but are offered for use as comparison standards in monograph procedures. The use of a molecular formula for the active ingredient named in defining the required strength is intended to designate the chemical or chemicals having absolute (100%) potency, e.g., cortisone acetate tablets contain 90–110% of the labeled amount of cortisone acetate, not 90–110% of cortisone acetate ($C_{23}H_{30}O_6$) itself, which, as the drug substance, must separately meet an acceptance criterion of 97–102% of $C_{23}H_{30}O_6$. The official USP Reference Standard is $C_{23}H_{30}O_6$, labeled to indicate the estimated content that is in fact $C_{23}H_{30}O_6$. Procedures in new *USP* or *NF* monographs requiring the use of Reference Standards are not in effect until the Reference Standard is available [10]. USP's official lot system presupposes all users are testing to procedures traceable to the official USP Reference Standard. Using validated analytical procedures, users should obtain the same result in testing in accordance with sound metrologic principles. This centrally controlled system with a single lot in commerce is dissimilar to that for a therapeutic article, for which many different lots may be in commerce at a specified point in time.

3. History

The USP Reference Standard collection began with an announcement of the availability of standards in *USP X* (1926). Early Reference Standards frequently were complex materials for biological assays, moving over time with the rise in modern pharmaceutical chemistry and manufacturing, toward well-characterized articles, and returning more recently to less-well-characterized articles (botanical dietary supplements and natural-source or recombinant DNA biologicals). This evolution is evident through the years: pepsin (1936); cod liver oil (1942); sulfanilamide, insulin, posterior pituitary (1942); melting point standards (1946); penicillin G sodium, heparin sodium (1950); negative control plastic (1966); dissolution calibrators (1978); endotoxin (1981); insulin human (1985); particle count set (1990); oxytocin (1996); and powdered ginger, Alliin (1999).

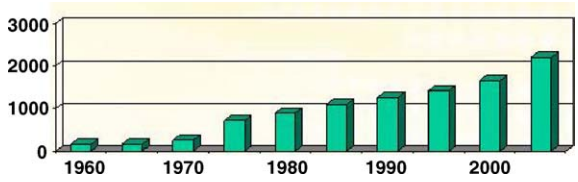

Fig. 1. Growth of USP Reference Standards Collection (number of Reference Standards per year).

Advances in analytical procedures during the past several decades have moved from those relying on direct measurement to those that rely on instrumental techniques (e.g., spectroscopic or chromatographic procedures) that frequently rely on physical standards for comparisons. With this transition, the USP catalog of Reference Standards has grown substantially (Fig. 1). FDA generally has welcomed USP's efforts in making Reference Standards available to facilitate analytical testing, with transfer of selected biologic standards in the 1930s and, more recently, the shift of antibiotic Reference Standards from FDA to USP in the 1970s [11].

4. Collection overview

The collection of official USP Reference Standards usually consists of highly characterized specimens of articles (drug substances, excipients, and impurities, including degradation products), and procedural standards, e.g., melting point, dissolution, particle size, and other calibrators. They range widely in *appearance* (crystalline or amorphous powders, volatile or viscous liquids, solutions or suspensions, gels or pastes, plastic sheets, and photomicrographs (in development)), *structure* (from simple inorganic salts to proteins produced by recombinant technology and cells (in development)), and *composition* (from single components to complex mixtures drawn from plant or animal sources). Their qualitative and quantitative uses also range widely according to the tests and procedures of the *USP–NF* monograph, including: *Identification* (qualitative); Assay (quantitative); *Impurities* (quantitative as feasible); *System Suitability* tests; and blanks and controls. Despite their varying applications, the primary uses of official USP Reference Standards are for spectroscopic and chromatographic procedures. They are used less often in other situations, including microbial assays (antibiotics), enzymatic reactions, animal tests, in vitro biochemical tests, titrations, and thermal analysis. Compendial uses of official USP Reference standards and detailed information about their storage are provided in *USP–NF* General Chapter ⟨11⟩ USP *Reference Standards*.

4.1. Sources of candidate material

Until the early 1980s, USP acquired most of its candidate Reference Standard materials from innovator companies that manufactured both active ingredients and dosage forms. As pharmaceutical manufacturing has moved offshore, so has USP's Reference Standards candidate acquisition. For the five-year period between 1999 and 2004, one-third of the USP Reference Standards candidate materials were obtained from non-US companies that had a manufacturing site within the US, and another third came from US companies that have manufacturing plants overseas as well as in the US. The remaining third comes from international companies without manufacturing sites in the US. As long as a Reference Standard candidate represents an article that is legally approved for marketing in the US, USP's Reference Standards Laboratory is able to seek it in both domestic and global commerce [4].

At times, USP will resort to custom synthesis of a reference standard. The custom-synthesis laboratories are qualified under USP's vendor qualification program (which is governed by SOPs in accordance with USP's certification to ISO 9001 and 17025 standards), and the material is tested according to the same laboratory process as are candidate materials obtained from innovator and generic manufacturers. As part of its ISO 9001–compliant quality management program, USP has implemented a supplier evaluation program. This program ensures that USP selects suppliers based on their ability to provide products or services that are in accordance with predefined internal requirements. A specific audit checklist, developed to ensure the quality of bulks provided by companies other than the innovator and primary generic manufacturer, is completed prior to the approval of any of these bulks for use as Reference Standards. Completed checklists are audited and reviewed by USP Quality Assurance staff to ensure that the appropriate quality systems are in place at these facilities. In addition to this checklist, an on-site verification of quality systems may be conducted, as needed. In addition, the selection process includes a review of all FDA and ISO compliance audits that have occurred at the manufacturing site. Review of these audits adds to USP's confidence in the supplier's ability to deliver a quality product.

4.2. Evaluation of candidate materials: ingredients

All candidate ingredient materials received for consideration as an official USP Reference Standard are evaluated by a minimum of three laboratories. These include USP's Reference Standard Laboratory, an FDA laboratory, and an independent, third-party laboratory that may belong to the primary manufacturer. Since September 2004 USP has been involved in a collaborative research agreement with Health Canada's Health Products and Food Branch Inspectorate. This agreement formalizes long-standing cooperation between the two nations and specifies that Health Canada laboratories will assist in the testing and characterization of candidate Reference Standards materials. USP, FDA, Health Canada, and any commissioned contract laboratories tend to follow an identical protocol, but prior testing conducted in a primary manufacturer's laboratory may have followed a different protocol. For a drug substance, the evaluation consists of at least one laboratory's performing a full set of tests according to the monograph specification (and additional tests as required), and all laboratories carry out additional studies to estimate content. USP relies on comprehensive testing and careful review of generated data by USP staff and by the USP Reference Standard Committee to ensure the quality of an official USP Reference Standard. Additional continued suitability for use testing is done to ensure that this quality is maintained.

The processes by which a candidate ingredient material becomes an official USP Reference Standard is at times complex and necessarily varies according to the article, its compendial uses, the analytical procedures utilized, whether or not the article has undergone a regulatory approval, and other factors. A general path for an article that has undergone a regulatory approval is shown in Fig. 2. In this most common approach, candidate material accompanies a manufacturer's Request for Revision [12], using information reflective of the regulatory filing and approval (e.g., characterization studies, methods validations, and the private specification concluded between the sponsor and the regulatory body). In many instances, the *USP–NF* monograph procedures are standardized approaches that for convenience and

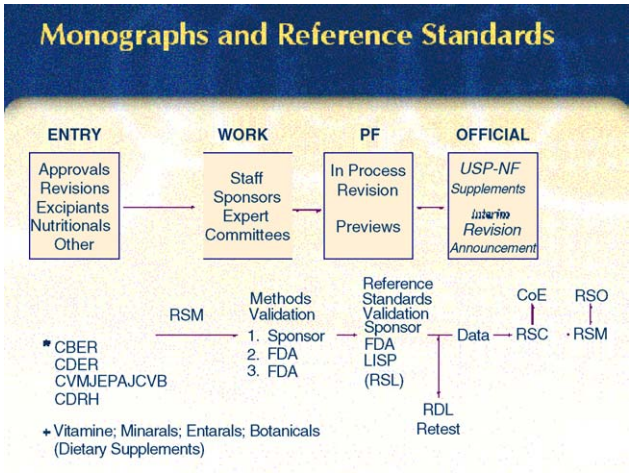

Fig. 2. Process by which a candidate Reference Standard Material (RSM) becomes an official USP Reference Standard (PF, *Pharmacopeial Forum*; RSL, Reference Standards Laboratory; RDL, Research and Development Laboratory; RSC, Reference Standards Committee; CoE, Council of Experts; RSO, Reference Standards Operations; CDER, Center for Biologics Evaluation and Research; CDER, Center for Drug Evaluation and Research; CVM, Center for Veterinary Medicine; EPA, Environmental Protection Agency; CVB, Center for Veterinary Biologics; CDRH, Center for Devices and Radiological Health). Note that more than one FDA laboratory may participate in methods validation, and Reference Standards Validation Sponsors may include other organizations. For example, USP has a collaborative research agreement with the Health Products and Food Branch Inspectorate of Canada to evaluate candidate Reference Standards.

Table 1
Four general approaches to assess the purity of a candidate reference material

Approach	Pros	Cons
Functional group analysis (e.g., titration)	Accurate and precise; Absolute determination; does not require a standard	Nonselective Requires standardized titrant, calibrated volumetric glassware
Assay against another standard	The standard is evaluated in the conditions in which it will be used; Selective; Ensures continuity between different lots	Not very precise; requires large number of experiments; Assumes that existing standard has not changed and that original purity determination was valid
Mass balance (subtracting from 100 the sum of volatiles, residue on ignition, and chromatographic impurities)	Potentially low error	In many cases the chromatographic impurities are unknown, or standards for them are not available. The impurities are then determined by the very risky assumption that the impurities have the same response factor as the main component.
All (or some) of the above	Flexibility in making the decision	Flexibility in making the decision

general applicability are expressed in General Chapters (e.g., ⟨281⟩ *Residue on Ignition* or ⟨921⟩ *Water Determination*) that can be adapted and verified as suitable for a particular article. In other instances, a specific compendial test must be developed both for the private and public specification, particularly for a monograph's *Identification*, *Impurities*, and *Assay* tests—and other article-specific tests as appropriate. This generally occurs after an FDA review and acceptance of the private specification, coupled with methods validation testing in FDA laboratories [13].

USP uses at least four general approaches to assess the purity of a candidate RM (Table 1). The most commonly employed is mass balance, which is based on measurement of impurities, including water, and subtracting from 100. For noncomplex active drug substances and excipients, usually results from three separate laboratories are combined to yield an estimate of purity (which may be expressed as a calculation value, for example as mg/mg).

USP General Chapter ⟨11⟩ *USP Reference Standards* notes that official USP Reference Standards are established and released under the authority of the Board of Trustees upon recommendation of the USP Reference Standards Committee (RSC),[5] which evaluates the selection and suitability of each lot [4]. Based on review of data from characterization and collaborative testing studies, RSC balloting occurs and must be unanimous for a positive decision.[6] With a positive decision, the material is subdivided and labeled, quality control checks are performed, and the material, now an official USP Reference Standard, is listed in the USP catalog and becomes available for distribution. Statements provided include safety warnings, required information for controlled substances, content or potency for quantitative-use standards, acceptance ranges as needed, and other useful information (Fig. 3 shows a proposed Reference Standard certificate). USP does not provide information from characterization studies and collaborative testing, which might be included for a certified reference material, because all the information that the user needs for the official applications of the standard is provided in the label text and, as necessary, in the additional documentation provided.

4.3. Evaluation of candidate materials: impurities

In the mid-1980s, based on various advances USP reconsidered its approaches to the measurement and control of impurities in drug substances and other articles. In 1986 USP held an open conference (its first) on impurities. Most participants supported a need for improved impurities approaches. The 1986 open conference set the stage for *USP–NF*'s General Chapters ⟨461⟩ *Ordinary Impurities* and ⟨1086⟩ *Impurities in Official Articles*. General Chapter ⟨466⟩ was created to provide detailed instructions for carrying out thin-layer chromatography (TLC) to measure impurities. General Chapter ⟨1086⟩ defined various impurities (ordinary, toxic, signal, concomitant, related substances, foreign substances, and process contaminants). Via a *Pharmacopeial Forum* (PF) Headquarters Column [14], USP invited manufacturers to submit TLC procedures for drug substance impurities. In the absence of a strong response, USP laboratories tested candidate Reference Standard material for impurities using TLC. Based on data from these analyses, the Subcommittee on Impurities proposed impurity tests for several drug substance monographs via *PF*. As many as 40–50 monographs in a single year were revised to include TLC testing for ordinary impurities using this approach, with approximately 200 concluded during several years.

These early efforts form the basis for an approach, present in many *USP* monographs, to control impurities via the *Related Substances* and *Chromatographic Purity* tests. These approaches did not always control specific impurities and generally stated that all impurities taken together should not exceed 2% unless otherwise indicated in the monograph. This approach was less than optimal, not only because of

[5] In the 2000–2005 cycle and in prior cycles, the RSC at USP traditionally was an appointed rather than elected committee of the Council of Experts, chaired by staff. In the 2005–2010 cycle, it is an elected Expert Committee.

[6] From 1999 to 2004, 13 donated candidate materials were rejected out of 1399 received. Twelve items did not successfully complete the collaborative testing process and were not presented to the RSC for review. The RSC rejected one item. The success rate for custom synthesis is similarly high. Of the 10 items synthesized, only one was synthesized improperly (Succinylmonocholine Chloride was synthesized with the wrong salt. During this period, 37 items from India, China, and Japan completed the full evaluation process, and none was rejected by either Standards Operations at USP or by the Reference Standards Committee.

USP
U.S. Pharmacopeia
The Standard of Quality℠

USP Certificate

Avobenzone
LOT G0B280

Molecular Formula
$C_{20}H_{22}O_3$

Molecular Weight
310.40

CAS Number
70356-09-1

LABEL TEXT

Avobenzone
500 mg
Dry portion in vacuum at 70° for 4 hours before using. For quantitative applications, use a value of 0.996 mg of avobenzone per mg on the dried basis. Keep container tightly closed.
FOR USE WITH SPECIFIED *USP–NF* TESTS • NOT FOR USE AS A DRUG • READ MSDS BEFORE USING
CAT. NO. 1045337 **USP Rockville, MD** LOT G0B280

USP certifies that the USP Reference Standards Committee, in accordance with their rules and procedures, determined that this USP Reference Standard lot is suitable to assess compliance with the monograph standards for which it is specified. The critical characteristics of this lot are usually determined independently in three or more laboratories, including USP, FDA, and academic or industrial collaborators.

QA Director

Page 1 of 2 09-Aug-2005

Calculation Value
Unless otherwise stated on the Reference Standard label, a value of 100.0% should be used in USP or NF compendial applications for which the use of this Reference Standard is intended. Please refer to the specific Reference Standard label for further information.

Expiration
Current lots are identified in the Official USP Reference Standards catalog. In some cases, the previous lot may still be considered official. If so, it is identified in the column marked "Previous Lot/Valid Use Date." Ordinarily, the previous lot is carried in official status for about one year after the current lot enters distribution.

It is the responsibility of each user to determine that this lot is current when used. To ensure up-to-date information, USP publishes the Official USP Reference Standards Catalog, which contains official lot designations. This information is also available on the USP web site, at www.usp.org, as well as in the bimonthly subscription publication, *Pharmacopeial Forum*

Instructions for Use
Follow the instructions in the appropriate USP or NF Monographs and General Requirements for Tests and Assays of the current *USP–NF*. In the event that instructions on the label of this lot differ from those found in the current *USP–NF*, those on the label supersede any instructions listed in Chapter <11>.

Non-Monograph Use
The suitability of this Reference Standard for use in non-compendial applications is solely the responsibility of the user.

LEGAL NOTICE
USP MAKES NO REPRESENTATION OR WARRANTY WITH RESPECT TO THE ACCURACY, COMPLETENESS, OR CURRENTNESS OF THIS CERTIFICATE; AND USP SPECIFICALLY DISCLAIMS ANY OTHER WARRANTY, EXPRESS, IMPLIED, OR STATUTORY, INCLUDING BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. USP DOES NOT WARRANT THAT THE INFORMATION CONTAINED HEREIN MEETS THE CUSTOMER'S REQUIREMENTS. USP SHALL NOT BE LIABLE ON ACCOUNT OF ANY SUCH ERRORS OR OMISSIONS.

USP Reference Standards are not intended for use as drugs, dietary supplements, or as medical devices.
This document is not a Material Safety Data Sheet.

This certificate may not be reproduced without the express written permission of USP.

Page 2 of 2 09-Aug-2005

Fig. 3. Example of a proposed Reference Standard certificate.

its *de minimus* character but also because it was based on a one-size-fits-all approach without regard for different impurities arising from different routes of synthesis. The approach has generally been superceded by approaches in the Quality documents of ICH. For noncomplex drug substances and dosage forms, these documents focus on stability testing (Q1 documents), analytical validation (Q2 documents), and impurities (Q3 documents). The ICH Q6A document defines characterization studies that lead to private specifications concluded between an applicant and a regulatory agency [15]. The private specification—and subsequently the public one—consists of universal tests (*Description*, *Identification*, *Assay*, and *Impurities*) and specific tests that can vary depending on the intended use of the ingredient and the type of dosage form. ICH Q3A(R), Q3B(R), and Q3C provide a comprehensive set of approaches for controlling impurities in ingredients and dosage forms [16–18].

Although impurities may be controlled by good system suitability testing and by response factors, as often is the case in the USP *Chromatographic Purity* test, general agreement exists that optimal control for specified impurities is an *Impurity* test procedure that adequately identifies impurities. This is best accomplished when the widest range of impurity Reference Standards is available [19]. USP, as part of its mission to advance public health, is exploring options to obtain candidate impurity Reference Standards materials and to advance them to official status, recognizing that the increasing globalization of the pharmaceutical industry offers both opportunities for the rewards of free markets and the dangers of counterfeiting.

5. Assignment of content

A key conclusion of the RSC, based on collaborative testing, is the value assignment for content of an official USP Reference Standard. In the past, USP usually applied a value of 100% using either two or three significant figures.[7] Over time, this practice came to be challenged for several scientific and pragmatic reasons. To address the issue, USP formed Project Team 4 (Reference Standards), which worked in association with the Prescription/Nonprescription Stakeholder Forum.[8] In addition, USP conducted some statistical

[7] The number of significant figures in the labeled calculation value is a function of the use of the standard and the number of significant figures in the acceptance range or limit. Generally, Reference Standards used in assays are labeled with three significant figures, and standards used in limit tests have two significant figures. Reference Standards that have multiple applications in different methodologies may require separate assay-specific assignments. For calibration standards, the labeled value is determined by a statistical analysis.

[8] For members of the Project Team, see author list.

studies on this topic; the more detailed work is presented in the Appendix A.

5.1. Content assignment: ingredients

The Project Team met on four occasions (October 2002, January 2003, May 2003, and October 2003), at times concurrently with the Reference Standards Committee. Their deliberations and recommendations provide further insights into the technical details of characterizing, maintaining, and validating Reference Standards.

One of the Project Team's earliest objectives was to develop a characterization protocol to outline the type of collaborative testing required for each category of Reference Standard and to define how Reference Standard assignments should be made. Characterization protocols are based on the three ways that Reference Standards are used: (1) quantitatively; (2) nonquantitatively; and (3) special applications. The protocol tests include: identification, impurities, ROI, water, testing against Reference Standards, testing against previous lot(s) of USP Reference Standards, and monograph tests. The Project Team recommended that the assignment strategy for USP Reference Standards should be based on mass balance when possible. Comparisons to previous lot(s) of USP Reference Standards generally should not be used for assignment purposes but should be completed to confirm an assignment and to determine if Reference Standards users observe a shift in results. When mass balance is not practical and an international Reference Standard is not available, direct methods can be used for assignment as well as for assay-specific assignments. Finally, the Project Team agreed that there should be no fewer than two collaborators and that replicates should be based on Reference Standard type; further, if the variability of the two collaborators' work was greater than the relative standard deviation (R.S.D.) of the method as submitted by the monograph sponsor, further studies should be conducted.

For a quantitative Reference Standard value, the Project Team recommended using a determined (actual) value for the labeled calculation value, not a threshold above which 100% is assigned. This approach has several positive aspects, including: (1) it needs no justification (unlike the case when one is assigning thresholds) because one is assigning determined (i.e., actual) values; (2) there are no rounding issues with the use of a determined (actual) value (e.g., 99.45 → 100%); (3) for assays of active pharmaceutical ingredients (APIs), potential problems arise when one assigns threshold values in the context of tight API limits; (4) materials with obvious impurities are assigned a value <100%; (5) a slight purity change does not result in a larger assignment change (e.g., if 99.5 → 100% and the USP Reference Standard changes to 99.4%, then the labeled value changes 100 → 99.4%).

5.2. Content assignment: impurities

The Project Team evaluated impurity Reference Standard values and recommended that USP should use a threshold above which 100% is assigned (e.g., if ≥98% → 100%). Positive implications of this approach include: (1) error caused by rounding is insignificant compared to error associated with the method for impurities at low levels; (2) error in rounding an impurity Reference Standard results in a more conservative estimate of impurities in drug substances and drug products; and (3) accurate quantification of an impurity is not as critical as it is for the ingredient. The recommendation is qualified in that: (1) using thresholds is not a true scientific approach; (2) if rounding occurs near a threshold, the difference could be detectable with the method; (3) if an error in rounding an impurity Reference Standard results in a more conservative estimate of impurities in drug substances and products, one faces the potential of rejecting material that actually meets specifications.

5.3. Value reassignment

Project Team 4 considered the reassignment of labeled Reference Standard values; i.e., what is an appropriate threshold difference beyond which relabeling is required? The team concluded that no fixed threshold exists for all products—decisions must be made on a case-by-case basis. However, for a typical well-characterized small molecule, reassignment might be necessary if the change is >0.5% for a material of purity >98% or if the change is >1% for a material of purity <98%. If assignment values are three significant figures and retesting (at the USP laboratory) differs by more than 0.5%, then further data should be collected, possibly by a collaborating laboratory. How these data will be collected, assessed, and compared raises significant questions involving statistical analysis and collaborative testing. Re-evaluation typically takes place at a single USP laboratory.

6. Discussion

In offering a *USP–NF* monograph and official USP Reference Standard, USP wishes to base its activities with the latest principles of sound metrologic science. These principles speak to "fitness for purpose," as follows:

The fitness for purpose of chemical measurements is formally defined as the "degree to which data produced by a measurement process enable a user to make technically and administratively correct decisions for a stated purpose [20]." A key element in the concept is for the 'interested parties to define in advance the acceptable degree of measurement uncertainty and desired degree of identification confidence' [21]. In addition to being the criterion for assessing when any aspect of the measurement effort is adequately complete, fit-for-purpose considerations are central to the prospective design of a measurement study [22,23]. The better defined the purpose, the more realistic the forecast of analytical effort required to achieve fitness. An unrealistic, unclear, or overly broad purpose may result in unnecessary costs, delay, or failure of measurement study.

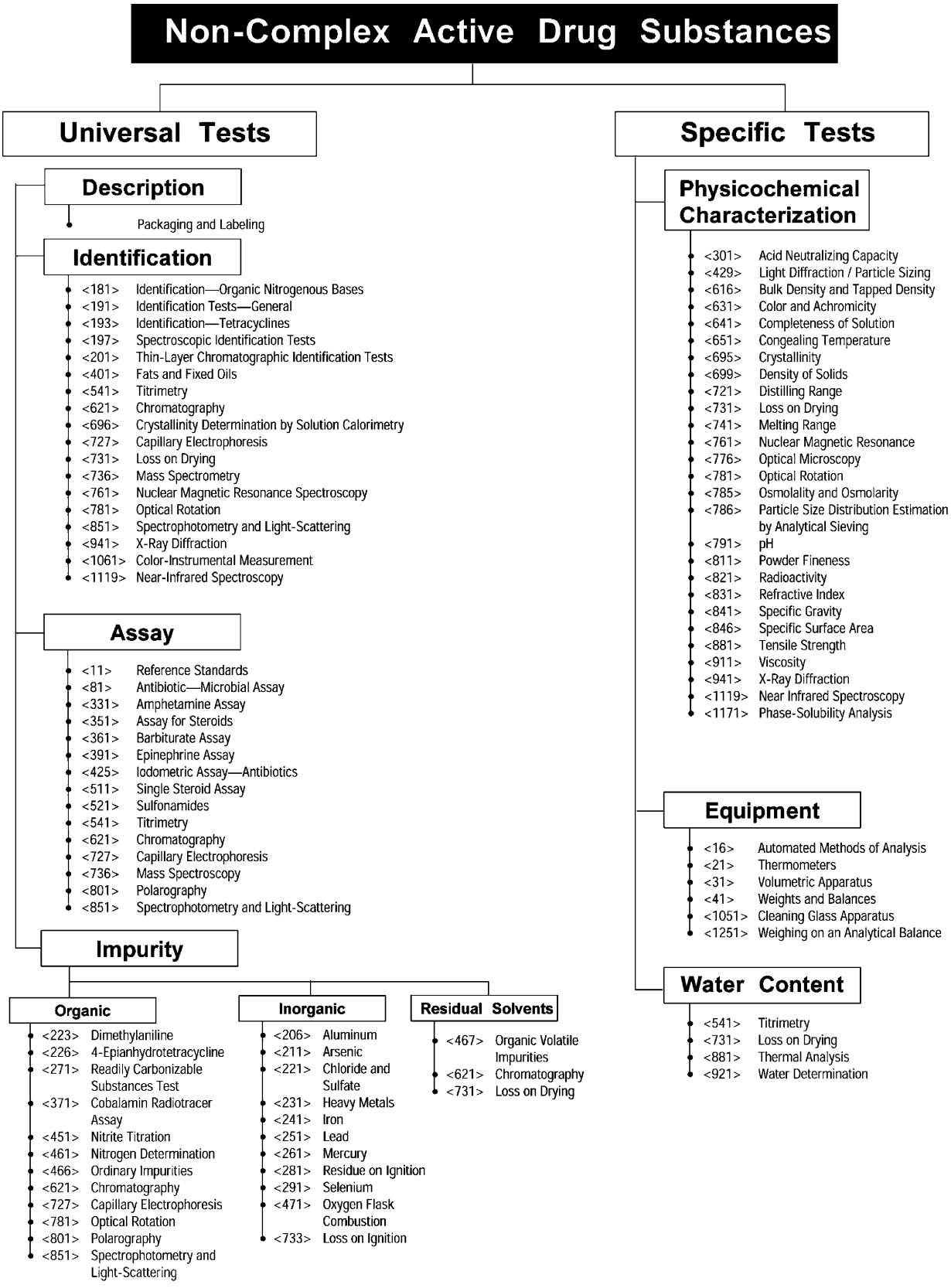

Fig. 4. Chart demonstrating some linkages between USP monographs for non-complex actives with ICH Quality approaches.

Compendial approaches have a strong role in assuring practitioners and patients – and the public at large – that a therapeutic article is "fit for purpose," i.e., is safe and/or effective in the maintenance of health and treatment of disease [13]. In this context, modern metrologic science supports USP's long-standing objective of ensuring the identity of an article via the test procedures and other standards of the monograph, regardless of who is manufacturing the article, who is testing it, and when or where it is tested.

As USP works with first-, second-, and third-party quality control laboratories, it is important to understand roles and responsibilities. USP's monographs and official USP Reference Standards are most commonly used in hypothesis testing studies by quality control laboratories. USP does not engage in hypothesis testing itself but rather provides the "measurement study" (monograph) and official USP Reference Standard to facilitate testing. The hypothesis of a quality control laboratory is that the article, when tested, yields a result that either does or does not fall within the acceptance criteria. If results fall within the acceptance criteria, the article is deemed acceptable. If not, the result may be deemed "out of specification" [24]. The article tested is an ingredient or a dosage form. The measurand is the active pharmaceutical ingredient and/or its impurity(ies) in the drug substance or dosage form. As a rule, the SI units for small molecules are almost invariably expressed in terms of mass (kilogram units). The general issue of either accepting or rejecting the testing hypothesis paradigm leads to issues of consumer and producer risk. In the past USP has not always carefully considered these two aspects of hypothesis testing. Recent work [25] has suggested this is a fruitful area of study for both manufacturers and compounding professionals, as well as practitioners and patients.

In the coming cycle, USP wishes to explore the value of estimating the uncertainty in the estimation of content in selected official USP Reference Standards. This uncertainly could be added to the overall uncertainty in hypothesis testing of, for example, measurand mass. USP and other pharmacopeias have not done this in the past on the assumption that uncertainly for a pharmacopeial reference material "is negligible in relation to the defined limits of the method-specific assays . . ." [26]. In fact, based on preliminary studies, USP believes that this may not always be the case and, further, that the definition of negligible might at times require careful consideration. USP believes that a better understanding of an uncertainty statement for an official USP Reference Standard, coupled with an uncertainty estimate of tests applied to the measurand, may help quality control laboratories reduce the likelihood of failing to meet acceptance criteria. Uncertainty in measurement should also be considered in setting acceptance criteria [3].

During the 2000–2005 cycle, USP created a *Guideline for Submitting Requests for Revision to* USP–NF [10]. The *Guideline* provides instructions to Sponsors intending to submit Requests for Revision and harmonizes many elements of the USP monograph with the ICH Quality approaches (Fig. 4). USP expects to revise this document continuously. For selected candidate reference materials, USP will add a section that provides a protocol with study design and analysis approaches. This protocol will focus initially on small-molecule ingredient and impurity candidate materials and can be expanded subsequently for other candidate materials as needed.

The integrity of the USP monograph and allied official USP Reference Standards is of paramount importance to USP. The authority to create and maintain both the documentary standards in *USP–NF* and the physical reference materials in the USP collection is a privilege in the United States. In other countries, the authority is reserved to the government. USP takes this privilege seriously and continuously works to merit it through close attention both to its processes, e.g., those of the RSC, and to the laboratory activities that support the availability of an official USP Reference Standard. In this regard, USP is especially proud of its attainment of ISO 9001 and 17025 certification during the 2000–2005 cycle. As this article demonstrates, USP also wishes to work closely with first, second, and third parties who rely on USP's official Reference Standards as a means of assuring the public that the articles they place in commerce meet optimal quality standards.

Appendix A. Statistical topics

A.1. Introduction

During the 2000–2005 cycle, USP engaged in a statistical analysis of various issues associated with collaborative testing of candidate reference material for value assignment of content for a small molecule ingredient. This effort involved theoretical considerations, a pilot study, and evaluation of data from current collaborative testing.

A.2. Theoretical considerations

The current USP collaborative study design for small-molecule Reference Standards typically uses three laboratories (L), with one experiment (E) per laboratory and 1–6 determinations (D) per laboratory. The precision of the estimated content depends on three sources of variability: 1) inter-laboratory variability; 2) inter-experiment variability within laboratory; and 3) intra-experiment variability (these can be based, e.g., on determinations, including inter-injection and inter-preparation for HPLC). If L, E, and D are all >1, all three variances can be estimated. Otherwise, only combinations can be estimated, which may be acceptable depending on the purpose of the study. For example, in the USP design for non complex ingredient Reference Standards, only one experiment per laboratory is usually performed, which precludes determination of a separate variance for experiment; any inter-experiment variability cannot be separated from inter-laboratory variability. As a means of simplifying presentation of formulas, a design is called *balanced* if there

are equal numbers of determinations per experiment and an equal number of experiments per laboratory.

For each component to be measured in a mass-balance determination, the precision of that measurement is needed. Let S^2 denote estimated variances, with subscripts L, E, and D for laboratory, experiment, and determination, respectively. The standard error (SE) of the estimated measurement for a single component of a mass-balance determination from a balanced design is then:

$$\mathrm{SE} = \sqrt{\frac{S_L^2}{L} + \frac{S_E^2}{L \times E} + \frac{S_D^2}{L \times E \times D}} \quad (1)$$

For confidence intervals, the degrees of freedom of the t-statistic is L-1, regardless of the magnitude of E and D. If the design is not balanced, then there is no simple formula for SE, and statistical software must be used.

To determine the precision of estimated content determined by mass balance, the standard errors from the various components need to be combined. Suppose there are C components to be assayed. For each component, this results in a standard error. The standard error for the content is then found by adding on the variance scale:

$$\mathrm{SE}_P = \sqrt{\mathrm{SE}_1^2 + \cdots + \mathrm{SE}_C^2}. \quad (2)$$

Here the subscript P is for the final content (purity) estimate and the other C SE's are for the C components. SE_p is the combined standard uncertainty of the estimated content.[9]

Determination of a confidence interval is more difficult. A conservative approach is to proceed as above using the t distribution with L-1 degrees of freedom. An alternative is to use Satterthwaite's approximation, as suggested by the National Institute of Standards and Technology (NIST) [3]. The confidence interval is of the form $C \pm k\mathrm{SE_p}$, where C is the estimated content and k is a coverage factor that depends on the degrees of freedom. (USP does not use the conventional choice of $k = 2$, due to the small number of degrees of freedom). The term $k\mathrm{SE_p}$ is the expanded uncertainty of the estimated content.

One implication of these considerations is that we can identify situations in which USP is unable to determine the precision of an estimated content. For example, if any non negligible component of a mass balance determination is measured by only one laboratory, it is not possible to estimate inter-laboratory variability, and hence the precision cannot be determined. Also, because in most studies the design is not balanced due to unequal numbers of determinations at the laboratories, USP needs information about the individual replicate values from all laboratories.

A second implication is an understanding of how to consider the design of these studies. Variability is addressed by increasing sample size; i.e., to increase precision, the sample size is increased with attention to each source of variability. Increasing the number of determinations, for example, does nothing to reduce the contribution from inter-laboratory or inter-experiment variability (see formula (1)). The only way to reduce the contribution from inter-laboratory variability is to increase the number of laboratories collaborating. If the goal of the study is to have a precise estimate of the content, then part of the design of the collaborative study should be specification of the desired level of precision. This could be stated in terms of a desired standard error or of a desired confidence interval width. The confidence interval is preferable because it incorporates information about how well the variances are estimated. There usually is no unique approach for choosing L, E, and D; choices typically are determined by cost, logistics, and other factors.

The following two sections summarize work conducted by USP to assess whether the current design was adequate for the purposes of assigning a content to USP Reference Standards.

A.3. Pilot study

A small pilot study of three candidate Reference Standard materials was conducted with the objective of determining the effect of increasing the number of collaborating laboratories on a) the final assignment by mass balance calculation of the content of candidate Reference Standard materials and b) the precision of the assignment.

The selection criteria for three candidate Reference Standard materials were: 1) the standards had to have an official quantitative application in an assay (thus requiring an assignment with three significant figures; 2) the standards must not yet have been evaluated in any laboratory (to avoid any potential bias); 3) the standards had to be on the USP list of priority candidates (so that the data could be submitted in a timely manner to RSC for approval); 4) the projected content was to be assigned by mass balance using results from at least two tests (practically, that meant that the standards had to be used as-is with water content being one of the elements used in the mass balance calculation; and 5) the candidates must not have been first-time standards. Study protocols included number of preparations, number of replicates, and sample size. Participating laboratories were USP's Reference Standards Laboratory and FDA laboratories; RDL and accredited contract laboratories also participated.

Candidates selected were bendroflumethiazide, sotalol hydrochloride, and valproic acid. All three materials presented special analytical challenges: (1) bendroflumethiazide testing requires two separate HPLC chromatographic purity tests and decomposes so rapidly in one of them that only singlet injections can be made from each preparation; (2) sotalol has three identified related compounds that are determined by external standard testing using the respective USP Reference Standards; and (3) valproic acid is hygroscopic, and the water content determination requires special precautions.

[9] In ISO and related documents, the combined standard uncertainty is described as a standard deviation. The standard deviation of the estimated content is more commonly called its standard error, the term the authors keep to.

Table A1
Results for bendroflumethiazide

Number of laboratories	Average impurities *D* (%)	Average other impurities (%)	Average water (%)	Estimated		Content 95% Conf. Int'l.
				Standard error	Average (%)	
5	0.05	0.26	0.21	99.47	0.04	99.37, 99.58
4	0.06	0.26	0.21	99.47	0.05	99.32, 99.62
3	0.05	0.26	0.22	99.47	0.06	99.20, 99.73
2	0.04	0.26	0.28	99.43	0.03	99.06, 99.80

Table A2
Results for sotalol hydrochloride

Number of laboratories	Average impurities (%)	Average water (%)	Estimated		Content 95% Conf. Int'l.
			Average (%)	Standard error	
5	0.17	0.08	99.76	0.02	99.70, 99.81
4	0.17	0.09	99.74	0.02	99.69, 99.80
3	0.17	0.08	99.74	0.02	99.64, 99.84
2	0.16	0.10	99.74	0.01	99.62, 99.85

Table A3
Results for valproic acid

Number of laboratories	Average impurities (%)	Average water (%)	Estimated		Content 95% Conf. Conf. Int'l.
			Average (%)	Standard error	
5	0.03	0.20	99.77	0.11	99.47, 100.06
4	0.04	0.25	99.71	0.12	99.32, 100.10
3	0.04	0.14	99.82	0.08	99.46, 100.17
2	0.06	0.06	99.87	0.03	99.46, 100.28

For each compound, the results were determined with five sets of laboratories. Results shown for two laboratories used laboratories labeled #1 and #2. Those for three laboratories used results for laboratories labeled #1–3, and similarly for four and five laboratories. Confidence intervals were determined using a *t*-distribution with degrees of freedom equal to the number of laboratories minus one.

Tables 1–3 show the results for the three compounds. Standard errors for the three compounds range from 0.02 to 0.12. All are precisely estimated.

The confidence intervals' widths reflect the magnitude of the standard error and the number of laboratories (an indicator of how well the standard error is determined). Normal expectation would be that the confidence intervals become narrower as the number of laboratories increases, and the *t*-distribution multiplier drops from 12.71 to 2.78 as the number of degrees of freedom increases from 1 (two laboratories) to 4 (five laboratories). That does not always happen here to the extent expected – Laboratories 1 and 2 are more consistent with each other than are the other laboratories; notice that the smallest standard error in each table is with two laboratories.

With three important caveats, the results of this pilot study do not support routinely increasing the number of laboratories from the current three for the type of candidate materials selected for study. The caveats are:

- the pilot study results apply only to compounds whose content is determined by mass balance,
- the pilot results refer only to three compounds that may not be representative of all compounds that may be measured by mass balance, and
- the results apply only to data collected as described here—namely with a standardized protocol and data collection form.

One can speculate that the standardized protocol and data collection may be contributing to the good precision seen here. Results from other compounds following current standard practice may be informative about this point. In any event, this pilot study brings us to a consideration of important ways in which modern metrology informs and improves

Table A4
Reviewed evaluation packages

RSCEP	Compound
719	Cefpodoxome Proxetil
808v01	Naratriptan HCl
850	Clonidine Related Compound B
856	Metoprolol Succinate
862	2E,4E-Hexadienoic Acid Isobutylamide
878	Leuprolide Aetate
1112	Amiodarone Hydrochloride
1127	Fexofenadine Hydrochloride
1130	Mangafodipir Trisodium
1143	Losartan Potassium
1144	Tolcapone
1164	Morantel Tartrate

Table A5
Confidence interval results for seven RSCEPs by mass balance determination

RSC	Water (%)		Impurities (%)		Negligibles (%)	Estimated (%)		Approx. 95% Conf. Int'l.	Monograph Accept. Range
	Average	SE	Average	SE	Average	Average	SE		
878 (Values in water columns are for acetic acid)	7.690	0.053	1.492	0.151	0.1	90.7	0.16	(90.0, 91.4)	±3%
1112	0.094	0.027	0.253	0.003	0.1	99.6	0.03	(99.4, 99.7)	
1127	0.220	0.044	0.078	0.014	0.11	99.6	0.05	(99.4, 99.8)	
1130			0.402	0.080		99.6	0.08	(99.3, 99.9)	±3%
1143	0.193	0.011	0.046	0.033		99.8	0.03	(99.6, 99.9)	−1.5%, +1.0%
1144	0.045	0.006	0.006			99.9	0.01	(99.9, 100.0)	±1.5%
1164	0.205	0.038	0.098	0.023	0.025	99.7	0.04	(99.5, 99.9)	

Negligibles are residue on ignition (878, 1112, and 1164), and sulfated ash and solvents (1127).

the processes involved in identifying, preparing, and working with Reference Materials.

A.4. Analysis of 12 representative data sets

To assess current practice, USP conducted a statistical analysis to obtain the confidence intervals for the estimated contents for 22 first-time Reference Standard candidate evaluation packages (RSCEPs) from two periods, December 2003 to January 2004 and August to October 2004. Here we consider only 12 assay standards with content determined by mass balance (Table 4). Other Reference Standards were for impurities or did not use mass balance.

Uncertainty (standard error and confidence interval) could be determined for 7 of the 12 candidates. Common problems are the lack of detail from Collaborator C, absence of replicate information about water determination from Laboratory A, and determination of non-negligible components by only a single laboratory—ether in one case and water in another. For the August–October set, USP staff obtained replicate information from Collaborators A and C. For the first set, no contact was made.

Table 5 shows confidence-interval results for the seven candidates for which confidence intervals could be determined. For three others from the first set of candidates, it seems likely that replicate information could have been obtained if sought, so there are only 2 of the 12 candidates for which uncertainty could not be determined from the studies as currently conducted.

References

[1] L. Anderson, G.J. Higby, The Spirit of Voluntarism—A Legacy of Commitment and Contribution: The United States Pharmacopeia 1820–1995, The United States Pharmacopeial Convention, Inc., Rockville, MD, 1995.
[2] H.G. Wolfe, *Charles Rice* (1841–1901), *an immigrant in pharmacy*, Paper presented at Joint meeting of American Pharmaceutical Association and American Institute of the History of Pharmacy, April 1949. Jacksonville, FL.
[3] B.N. Taylor, C.E. Kuyatt, Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results, National Institute of Standards and Technology, Gaithersburg, MD, 1994.
[4] General Chapter ⟨11⟩ USP Reference Standards. USP 28–NF 23. The United States Pharmacopeial Convention, Inc., Rockville, MD, 2005. p. 2204.
[5] W. May, R. Parris, C. Beck, J. Fassett, R. Greenberg, F. Guenther, G. Kramer, S. Wise, T. Gills, J. Colbert, R. Gettings, B. MacDonald, Definitions of Terms and Modes Used at NIST for Value-Assignment of Reference Materials for Chemical Measurement, National Institute of Standards and Technology Special Publication 260-136, Gaithersburg, MD, 2000, p. 10.
[6] S.L.R. Ellison, M. Rosslein, A. Williams (Eds.), EURACHEM/CITAC Guide: Quantifying Uncertainty in Analytical Measurement, second ed., National Institute of Standards and Technology, Gaithersburg, MD, 2000, p. 9.
[7] D.L. Deuwer, R.M. Parris, E.W. White, V.W.W. May, H. Elbaum, An Approach to the Metrologically Sound Traceable Assessment of the Chemical Purity of Organic Referenxce Materials, NIST Special Publication 1012, National Institute of Standards and Technology, Gaithersburg, MD, 2004, pp. 3.
[8] Federal Food, Drug, and Cosmetic Act (Public Law No. 75-717 52 Stat. 1040), 1938.
[9] 21 U.S.C. §321(g).
[10] *General Notices*, *USP* 28–*NF* 23, United States Pharmacopeia, Rockville, MD, 2005. p. 5.
[11] Valentino J.G., Pharm. Technol. 3 (9) (1979) 61–64.
[12] Department of Standards Development. Guideline for Submitting Requests for Revision to USP–NF. United States Pharmacopeia, Rockville, MD, 2004. Also available at http://www.usp.org/standards/revisionguideline/index.html (accessed 05/09/05).
[13] T. Layloff, M. Nasr, R. Baldwin, M. Caphart, H. Drew, J. Hanig, C. Hoiberg, S. Koepke, G. Lunn, J.T. MacGregor, Y. Mille, E. Murphy, L. Ng, R. Rajagopalan, E. Sheinin, M. Smela, M. Welschenbach, H. Winkle, R. Williams, Pharm. Technol. 24 (1) (2000) 30–42.
[14] L.T. Grady, W.M. Heller, Headquarters Column. Pharm. Forum. 11 (5) (1985) 620–621.
[15] International Conference on Harmonization. Q6A: Specifications: Test Procedures and Acceptance Criteria for New Drug Substance and New Drug Products: Chemical Substances, Geneva, Switzerland, 2000.
[16] International Conference on Harmonization. Q3A (R): Impurities in New Drug Substances, Geneva, Switzerland, 2000.
[17] International Conference on Harmonization. Q3B (R): Impurities in New Drug Products, Geneva, Switzerland, 2000.
[18] International Conference on Harmonization. Q3C: Impurities: Residual Solvents, Geneva, Switzerland, 2000.
[19] L. Bhattacharyya, H. Pappa, K.A. Russo, E. Sheinin, R.L. Williams, Pharm. Forum. 31 (3) (2005) 960–966.
[20] IUPAC. Compendium of Analytical Nomenclature, Definitive Rules 1997. ("Orange Book"), third ed. Blackwell, Oxford, 1998.

[21] R. Bethem, J. Boison, J. Gale, D. Heller, S. Lehotay, J. Loo, S. Musser, P. Price, J. Stein. J. Am. Soc. Mass Spectrometry 14 (2003) 528–541.
[22] T. Fearn, S.A. Fisher, M. Thompson, S.L.R. Ellison, Analyst 127 (2002) 818–824.
[23] L. Bhattacharyya, T. Cecil, R. Dabbah, D. Roll, S. Schuber, E.B. Sheinin, R.L. Williams, USP Council of Experts Expert Committee, Pharm. Res. 21 (10) (2004) 1725–1731.
[24] Food and Drug Administration, Center for Drug Evaluation and Research. Guidance for Industry (Draft Guidance), ANDAs: Impurities in Drug Substances, Rockville, MD, 2004.
[25] W.W. Hauck, T. Foster, E. Sheinin, T. Cecil, W. Brown, M. Marques, R.L. Williams, Pharm. Res. 22 (2) (2005) 182–187.
[26] ISO. Guide 34 General Requirements for the Competence of Reference Material Producers. second ed. Geneva, Switzerland: ISO Press; 2000:v.