F-D-C REPORTS —

FOUNDED 1939
$625 A YEAR

www.TheGoldSheet.com

"The Gold Sheet"®

PHARMACEUTICAL & BIOTECHNOLOGY QUALITY CONTROL

THE NEWS THIS ISSUE

Vol. 40, No. 6 June 2006

BIOTECH PRODUCT REFERENCE STANDARDS are drawing attention early in product development as innovator firms realize their importance to later problem solving. The benefits of well-characterized reference materials in biopharmaceutical development and manufacturing control was a central theme at a recent "CMC strategy forum." FDA and industry experts discussed the evolving strategies for selecting and qualifying reference standards for new biotech products and the regulatory implications involved. Reference standard acceptance criteria, formulation and change management are among the challenging issues that drew debate. ***[A presentation by a Biogen Idec official on the various stages of reference standard development and the biotech firm's program and procedures for handling the standards is included on pp. 11-14.]***

Early Investment in Characterization Pays Dividends, CMC Forum Participants Agree

FDA is seeing increasingly comprehensive characterization of reference standards early in the development of new biotechnology products as analytical techniques improve and companies recognize the potential advantages in later stage problem solving.

The trend was highlighted by Center for Drug Evaluation and Research (CDER) Division of Monoclonal Antibodies biologist Elena Gubina at a well-characterized biotechnology pharmaceutical (WCBP) "CMC strategy forum" on reference standards in San Francisco in late January.

While FDA understands that different stages of biopharmaceutical development can have different degrees of reference standard characterization, the agency is finding that "when companies submitted ideas for Phase I to us maybe five, seven years ago, they characterized a reference standard much less than they are doing now," Gubina reported.

It appears, she said, that "from their own experience, companies started to realize that more comprehensive characterization of the reference standard at early stages allows the problems of product development to be solved faster."

The monoclonals division, Gubina noted, has seen "quite a few cases when during clinical development the product lost its efficacy and for some reason the toxicity of the product went up, and the company needed to retrace its steps and discover what happened with the product." The problem can develop after the introduction of manufacturing changes, and "much more often it happens if the cell line is unstable" and the company "needs to go back and pick up another cell line from the freezer, and start it almost from the beginning." She emphasized that "a well-characterized reference standard may even prevent such [problems] from happening."

The CDER official cautioned that reference standard lots must be stored appropriately to maintain the highest level of purity and activity. She also stressed that they should be produced "in quite a large quantity because you will need to use this reference standard quite extensively during your product development or product life after licensure."

➤ During the discussion period at the forum, CDER Office of Biotechnology Products Acting Director Steven Kozlowski and Center for Biologics Evaluation and Research (CBER) Division of Hematology Acting Deputy Director Andrew Chang supported the emphasis on early reference standard characterization.

Chang acknowledged that during the development stage, companies have to make budgeting decisions. However, his bias as a regulator is that "the standard is such important material for your product development, it is worth it to spend money on this material." The improved knowledge of the material will result in better judgments later on, he stressed.

Kozlowski commented that "clearly there is guidance on what you need for Phase I study, and that guidance doesn't say you need to fully characterize your molecule for all of its physiochemical attributes."

However, he summarized, "from what we have been hearing today, it is a good idea. Depending on the resources of a company and what they can do, the more they can characterize and the more development lot [knowledge] the better. That doesn't mean that you can't get into an IND phase if you haven't fully characterized every disulfide bond and every oxidation. But again, it is better to know more early if you can afford those upfront costs."

The in-depth characterization of the initial primary reference standard and producing a sufficient quantity are key to avoiding the problem of having to change standards later, Kozlowski stressed.

He echoed his FDA colleague Chang in advocating the perspective that the reference material "is something you are going to use, it is something you are very dependent on, it is critical to your product development – characterize it to the best extent you can."

Kozlowski recognized that "if you did some very specific studies with complicated orthogonal steps to work out one particular variant and it turned out not to be important at all, maybe that doesn't need to be repeated at the characterization. But it would seem to me that since you [establish reference standards] infrequently and it is critical, this is something where the cost of doing the characterization well is not wasted."

CDER Office of Pharmaceutical Science Deputy Director Keith Webber also concurred, adding that "early on, the more information you have, the better off you are going to be in terms of product development later."

CMC Forum Focuses On Innovator Standards

The importance of qualified reference materials in biopharmaceutical development and manufacturing control was a central theme at the CMC strategy forum.

- Unlike the discussions at the December FDA/ NYAS/NIST meeting in Brooklyn on reference standard issues for follow-on biologics ("The Gold Sheet" May 2006), the focus of the forum was on the in-house development, refinement and use of reference materials by the innovator through the life-cycle of a biotech product.

The reference standard forum was held in conjunction with the annual WCBP conference sponsored by the California Separation Science Society (CaSSS) and FDA. Organized semiannually, the CMC strategy forums are designed to bring industry and regulators together to discuss ideas and share experiences on key issues of concern, with the goal of developing better standards and guidance in the evolving biotech arena.

The previous workshop, held last July at the National Institutes of Health (NIH) campus in Bethesda, Maryland, focused on stability programs ("The Gold Sheet" October 2005). The next forum will be held at NIH on July 20-21 and will address "changing paradigms in process validation."

The organizing committee explained the rationale and intent for the January forum on reference materials as follows:

"The implementation of robust strategies for establishing proprietary reference materials during

"The Gold Sheet"

FOUNDED 1967
F-D-C REPORTS, INC.
5635 FISHERS LANE
ROCKVILLE, MD 20852
PHONE 240-221-4500
FAX 240-221-4400

CUSTOMER SERVICE 1-800-332-2181
FDC.Customer.Service@Elsevier.com
TECHNICAL SUPPORT 1-866-353-5108
AN ELSEVIER COMPANY

Executive Editor: **Bill Paulson**
Senior Editor: **Bowman Cox**
Composition: **Ting Yang**
Production: **Phil Simms**
Advertising: **Nicole Muehr**
Customer Service: **Emily Brainard**
Marketing: **Liz Perman**
Senior Marketing Director: **Sean Smith**
Senior VP – Sales, Marketing & Business Development: **Michael Magoulias**
Director of Web Publishing: **Patrick Bobst**
Editorial Director: **Melissa Carlson**
Executive Director of Operations: **Jim Chicca**
Business Controller: **Kathy Schaffhauser**
President: **Mike Squires**

"THE GOLD SHEET" (ISSN 1530-6194) is published monthly by F-D-C Reports, Inc., 5635 Fishers Lane, Suite 6000, Rockville, MD 20852. The annual subscription rate is (1) $625 or (2) $355 for additional copies mailed in the same envelope with $625 subscription. Periodicals Postage paid at Rockville, MD and at additional mailing offices. POSTMASTER: Send address changes to "The Gold Sheet", 5635 Fishers Lane, Suite 6000, Rockville, MD 20852. **In Japan, Elsevier Science KK** is the exclusive subscription representative for "The Gold Sheet." E-mail: jp.bkinfo@elsevier.com. URL: www.elsevierjapan.com. Telephone: 03-3589-6370. Fax: 03-3589-6371.

product development serves as an important cornerstone for setting meaningful product specifications. While many regulatory documents (e.g., ICH Q6B) offer general guidance regarding the establishment of reference materials, little specific direction is provided. The symposium will provide an open forum to discuss pertinent issues in the design and implementation of reference material strategies during product development."

➤ The morning part of the forum addressed reference selection and regulatory considerations.

The session began with Gubina providing an "FDA perspective on reference standard characterization and maintenance." Presentations followed by Elena Vasilyeva, a senior scientist in Biogen Idec's QC chemistry lab, and Luke Somerlot, a director at GlaxoSmithKline, on their respective firms' reference standard programs.

In the afternoon, the participants focused on qualification and stability strategies.

CBER official Chang provided an FDA viewpoint, focusing on potency standards for therapeutic proteins. Merck Research Labs Bioprocess and Bioanalytical Development Associate Director Katherine Owen addressed the assigning of potency to reference standards for live viral vaccines, and Anne Jespersen, a scientist at Novo Nordisk, discussed her firm's strategies for qualifying reference materials for well-specified recombinant proteins.

The presentations at the two sessions were followed by extensive discussion periods in which the relevant issues were debated (*see box on p. 3*).

Standards Are Important Pre- And Post-Approval

In the opening presentation, Gubina commented further on why it is important to use the reference material both during clinical development and after licensure.

"During clinical development you usually have quite substantial lot-to-lot variation, and you make quite a few manufacturing changes. You are trying to develop your manufacturing process to the level of a commercial process. The same goes on with your analytical test methods, the same problems – quite large variations. Your analytical test methods are not validated yet. You may choose to change some analytical methods during this stage of the product development."

The reference standard, in turn: "helps to assure consistency in the level of dosing during clinical trials; helps in assessment of lot-to-lot variations and the

Key Questions For Reference Materials

The discussion at the WCBP CMC strategy forum on reference standards was focused on the following questions related to reference material selection, regulatory considerations, qualification and stability evaluation:

Reference Material Selection and Evolution

- What criteria are used to select the initial and subsequent reference material lots?
  - Impurity levels (least impurities, highest level of acceptable impurities, or something in between?)
  - Is the candidate material representative of the manufacturing process?
- What risks and/or benefits are associated with different formulations of reference material compared to drug product?
- What types of changes prompt the need to establish a new reference material?
- Is it appropriate to use pre-process validation reference materials to support commercialization?

Regulatory Considerations

- When do you establish prospective acceptance criteria for reference material?
- Should the acceptance criteria for reference materials be tighter than the acceptance criteria for lot release specifications?
- What happens if your reference material does not meet the acceptance criteria?

Qualification of Reference Materials

- How is a new reference material qualified? What additional testing should be done to ensure accuracy and consistency? How much qualification testing is appropriate?
- How do you assign reference material potency?
- How do you minimize variability in reference material potency assignment (potency drift)?

Stability Considerations

- What elements of stability programs are appropriate for reference materials?
  - Formal stability?
  - Forced degradation studies?
  - Shipping studies?
- How are re-qualification dates established?

biologic drug's safety during manufacturing development; and helps in assessment of data during analytical method development."

The agency's guidance on monoclonal antibodies, Gubina noted, states that a properly qualified in-house reference standard should be used for lot-to-lot comparisons. The standard should have known characteristics, specificity and potency, be stored under appropriate conditions, and periodically tested to ensure its integrity. It should be updated as a product evolves, but finalized by the start of Phase III trials.

When qualifying a reference standard, firms may be tempted to use the lot with highest purity. However, Gubina cautioned that doing so may lead to the problem of future clinical lots being out of spec because of the extra purity in the standard.

➤ The importance of the reference standard after licensure rests in providing consistency, the CDER regulator emphasized.

"It provides consistency of your product during day-to-day manufacturing. It provides consistency of your product and helps in assessing the impact of manufacturing changes. And sometimes, especially when you have quite a long life of your product after licensure, you may need to change some analytical equipment. You may want to change some analytical test methods, because something new was developed, for example, in the field, or you develop some new method. The reference standard will help with all these hurdles."

Reference standards should be used during analytical testing, Gubina continued. The material "should be used for potency, purity, safety and all other product characterization when possible." For example, side-by-side testing with the product sample should be done with SDS-PAGE and IEF, while HPLC is "a good example" of where sequential use is appropriate. In other cases, the analysis may involve analyzing a mixture of the reference and sample product.

Gubina pointed out that ICH's analytical methods guidance Q2A calls for well-characterized reference materials with documented purity to be used throughout the validation study, with the degree of purity depending on the intended use.

➤ The stability of the reference standard is also important and should be monitored over its life-cycle.

"When you compare your product lot to a reference standard, you need to know what is going on with your reference standard, what condition it is in," and stability programs should include tests that are able to detect reference standard degradation, the CDER official said. In general, she advised, the standard should be stored at a different temperature than product lots to prevent the same pattern of degradation.

Reference Standard Changes Present Challenges

Gubina also commented on the implementation of a new reference standard, which, she noted, may happen when manufacturing changes are made and there are biochemical differences that show up.

FDA reviewers take a case-by-case perspective, taking into consideration the stage of product development, the depth of the biochemical comparability study performed, the sensitivity of the test methods used, and the additional data available to support the safety of the reference standard change.

➤ One difficult juncture for the reference change is after clinical trials and before full-scale manufacturing.

The situation can develop, Gubina noted, when a company acquires a product from another company. "To manufacture this product, they want to make some changes or use their own facilities," and there may be some lot differences involved, with potential implications for the quality, safety and efficacy of the product going from clinical trials into a "very large, very diverse patient population.'

Changing standards between Phase I and Phase II studies may also present difficulties, she said – for example, if a company buys a product from an academic institution which manufactured some clinical material with the help of a contractor.

"Use your common sense," Gubina advised. "Remember that when you make some biochemical comparison" of pre- and post-change lots, "and you can see that there is some minor peak which appears, for example on a chromatogram, that it is difficult, at least for us, to assess the impact of these small peaks on the safety or efficacy of the product." The additional data available "may help to make this decision."

Additional Characterization Urged Also For Early Lots

Following her presentation, Gubina was asked by industry consultant Nancy Sajjadi for an interpretation

of her statement that the characterization of the reference standard should go beyond lot-release testing.

Sajjadi suggested that early lots should also include the additional characterization in order to assure the representativeness of the standard.

The consultant sees a "reluctance many times" to do the same degree of additional testing on the lots used for pre-clinical or early-phase clinical studies as is done for the reference standard "for fear that" these will then become normal lot release tests. However, she sees the additional testing as critical "because we always assume that these reference materials are representative for what we need them to do, and the only way to really know that and for them to be truly useful is to do a lot of characterization on all of the early lots, not just the reference standard."

➤ Gubina responded that her division sees a lot of IND submissions and understands that maybe only one in ten of the products will make it to market.

FDA wants the product to be characterized well even for a Phase I study, and "it would be nice" if the material used in Phase I were acceptable for licensure, she said. However, "we understand that you cannot make such a product when you are just starting development. That is why sometimes maybe our requirements are not as tough as you yourself make them….We give you some leeway just to start your clinical trials and then to catch up with all the characterization, even when we understand that characterization" of the reference material and the product "is very important and the more you characterize your product itself from the beginning of clinical trials the better."

Chiron Analytical Development Director Chulani Karunatilake stressed the importance of "trying to identify as early as possible what are product-related substances and what are product-related impurities." Although resources may be limited early on, Karunatilake advocated "investing enough effort early on to try and define some of these variant forms – even starting from the preclinical reference material."

From Interim To Primary Reference Standard

In her presentation on Biogen Idec's reference standard program, QC official Vasilyeva explained how the key milestones for the reference material "are related very tightly" to the product's development stage (*see box on pp. 11-14*).

➤ GlaxoSmithKline's Somerlot similarly addressed the "evolutionary pressures that come to play on the reference standard" as products move along the pathway from research to final commercial product.

Somerlot joined in pointing to "the many important roles" that reference standards play over the course of the biopharmaceutical life-cycle, including supporting "the generation of tox materials, test methods, release of clinical trial supplies, method and process validation, product approval and licensing, and finally commercial manufacturing." As the product development life-cycle progresses, he stressed, "the more critical it becomes to have a well-analyzed and highly characterized reference standard."

The evolutionary factors are different at each phase of product development, Somerlot pointed out. These include "different uses, different expected life spans, process [evolution], cost impacts, quantity and scale, quality compliance pressure and potential for project success."

GSK introduces an "interim reference standard" during the early stages of development, he explained, which "may be produced in a research or development lab, and not manufactured to GMP at this point" nor have all specs defined.

A primary reference standard is then prepared when: specifications are in place for test methods, release, formulation, container size and units per container; the stability program can support at least one to two years expiration; and the storage requirements are defined. These primary reference standards are used in R&D assay development and transfer to QC, fermentation and cell culture process monitoring, recovery process development, and for regulatory filings.

➤ Somerlot described GSK's process for selecting the primary reference standard:

"Essentially a candidate is brought forward by discovery research and brought into the R&D organization. If that candidate is selected, it goes through a project team. From that project team, a production campaign will be scheduled in the pilot plant. There is also an analytical subteam that is formed."

In the pilot plant, "a bulk drug substance lot is selected for the reference standard out of that schedule, and the quality control and characterization team will

thoroughly test the material against specifications and do additional characterization. That becomes the reference standard after QA has reviewed the manufacturing records and the record of analysis." The reference standard is approved and implemented, "and then finally the reference standard inventory is controlled by materials management."

Somerlot explained that while the specifications are still in the process of being refined, the Phase I reference material is "highly characterized" and compared against the interim standard. "So we do have a highly characterized analytical reference standard supporting the IND."

➤ The bioanalytical comparability of the commercial reference standard becomes, in turn, "the important step that will link your development to your commercial process."

The test methods GSK uses for this comparability study encompass a broad range of physiochemical, biochemical and biophysical techniques.

The bioanalytical comparability testing, Somerlot explained, "will be two-fold: Firstly, the Phase III reference standard, which is prepared from material manufactured at pilot scale, will be fully compared with the commercial reference standard prepared from commercial scale, typically from a PQ lot. Secondly, a further two PQ lots will be compared with the pilot scale reference standard using a subset of these test methods....And the results should be a primary reference standard for marketed products."

Typically, he said, several normal production lots have been evaluated at this point, and the stability profile is well defined. He added that multiple bulks could be pooled to make this reference standard, and that process changes or new indications may require new reference standards.

What Is Representative?

During the discussion that followed the morning presentations at the forum, the panelists were asked to expand on the criteria used to select the initial and subsequent reference material lots.

The criteria are "sometimes very loose" for the initial reference standard, reflecting the early specifications, Somerlot commented. With a limited amount of time and production schedule and the process evolving, "you have to use what you have," while making sure that "you have as much characterization data as you can possibly get." The selection criteria are "going to be mostly based on what is coming out of development."

Asked specifically about how GSK picks the criteria for a typical monoclonal for which the company has not yet started Phase I clinical trials, Somerlot explained that "the organization has a lot of history with monoclonals and it tends to build on each other. We are lucky that we have somewhat of a generic process, and so from that we apply, as much as we can, generic specifications for these reference standards. Again, those criteria are generated in the development organization and are brought forward as specifications to the analytical subteam of the project team, and then those are evaluated and decided upon there."

➤ The initial reference standard often is the same material used in the tox study, Somerlot said. "You are going to generate your safety data off that so that will continue into the evolutionary process of generating the next specification."

Continuing the conversation, industry consultant Sajjadi suggested that a distinction should be made between product types that have a lot of characterization already and "those that are new, different, where there is no history, so you have to start with more of a philosophical approach rather than data and information." She cited autologous products, which are different for each patient and yet "still need some form of standards and reference," as on the extreme of the latter side of the spectrum.

Sajjadi again raised the issue of what is "representative," pointing out that if "you have one or two lots, [that] is representative. So in the beginning, it is production driven. As you go along, representitiveness is going to be driven by what you know and how similar things are."

➤ Addressing the characterization of early lots beyond the lot release criteria, Sajjadi urged the need to "understand from the beginning what your product is" by taking "a more holistic approach, and to the best of your ability characterize the product extensively so that the information can be linked together with the safety study."

She noted that there may be some reluctance to do that in small companies due to evolutionary and financial pressures. However, in other cases, "it is more philosophical – people are afraid that if they have more data than the regulators expect then they might get in

trouble, because they might actually find out something about their product in the process."

Responding to an informal survey by the session moderator of whether the participants were doing detailed characterization of preclinical reference material, Merck Research Labs' John Hennessey suggested that the terms are relative. "You probably have limited methods," he noted, and if there are enough resources, "you are likely to throw every one of them at those first materials that are coming out. So what is detailed? If you have 100 people, it is different than if you have five."

Hennessey added that "if you want to put in something into your preclinical safety testing that you feel is representative of what you want to make, you are going to pick something that is not an outlier. It has got a spectrum of characteristics that you would like to make in future batches whether you modify the process or not. At least it gives you a target."

Tox/Reference Material Overlap Tightens Linkages

Participants concurred that, where possible, using the same lot for the tox study and reference material is beneficial.

➤ Having this overlap is "a great idea," Sajjadi maintained, "because you can link the analytical data to the biological data and you can build on that."

GlaxoSmithKline analytical scientist Linden Gledhill commented that "the ideal situation is obviously to make the reference standard on the tox material and your first clinical batches from the same campaign. That is a seamless approach where you are applying extensive characterization to material that is from the same campaign."

However, Gledhill stressed, the pressures of a particular project may not accommodate that approach. The "fall back position" at GlaxoSmithKline if there is a campaign of process demonstrations lots conducted ahead of the clinical batches is to "then place those into the tox studies and along side that we will generate the first reference standard." That reference standard can then be used to compare clinical batches against.

When time pressures are even greater, Gledhill continued, "you may only be able to get your tox material and your reference standard from a smaller scale process demonstration batch. And of course you are moving further away from your clinical batches, so there is more risk. But I think coming back to what was previously said, it is process driven – you only have a process that you can use in the time available….We all evolve our own processes, so you have to lay that foundation down with the original reference standard."

➤ Biogen Idec Analytical Development Director Rohin Mhatre concurred that "like it or not, your process is dictating your product" during the early stages. Firms have to rely "at some point" on historical knowledge and not "get so hung up on saying exactly what we would like."

GSK's Somerlot suggested keeping in mind that in the case of monoclonals there is now an EP monograph available, "so you do have a template that you can start building your criteria up around for your reference standard."

CDER biotech product official Kozlowski acknowledged the material limitations that affect early reference standard development, adding: "clearly, the more you can learn early the better." Similarly to linking the initial material to the tox study, it is preferable if later materials are linked to clinical lots, he advised – "preferably pivotal clinical lots, because those are ultimately the closest link to safety and efficacy."

The OBP official commented further that there is often a number of "comparability exercises" for the biotech products, not all involving reference material.

Noting the debates that surface over the number of lots that should be involved in these comparisons, he advised that "certainly the agency likes multiple lots to multiple lots, not just a reference standard, so you are accumulating characterization information on a variety of lots and processes over time. And that aggregate information, when combined with using clinical lot data, would seem to me to be not a bad starting point for what criteria ultimately you would use for picking the primary reference standard."

Amgen Product Quality and External Affairs Head Anthony Mire-Sluis commented that Amgen is finding that a "big advantage" in using the same lots for reference standards as were used in tox and clinical studies lies in helping establish the "product development history file" and the new common technical document (CTD). Before joining Amgen in the spring of 2005, Mire-Sluis was a senior policy advisor on biotechnology products at FDA.

Mire-Sluis noted that Europe, in particular, wants firms to be able to "show what your product looked like over

time during development – even at the earliest stages." The extensive characterization done for reference standards, together with the certificates of analysis of the clinical lots, efficiently builds the requisite development history file, he said.

Perfect Product Continuity Not The Norm

Merck Research Labs Bioprocess & Bioanalytical Research Executive Director Robert Sitrin commented that perfect continuity of the product as you move through tox and clinical development is not the norm.

"What usually happens is the later product lots look different," he said. When questioned by clinical and process people if the material is acceptable, analysts can only point to the difference. "Then comes the value judgment as to whether the difference means something."

➤ In making these judgments, knowledge of the toxicology material and retaining of samples becomes pivotal, Sitrin affirmed.

"The more you know about your toxicology lot" and the fewer surprises, "the better off you are. So that becomes a very pivotal material. Whether it is ultimately the reference standard, for much of your development time it is something you will go back to compare with so that you can answer that question if you can. And if you can't, then you may have to repeat certain studies....Obviously the safety profile in the clinic builds on that."

Echoing a prominent theme at the previous CMC strategy forum on biotech stability testing ("The Gold Sheet" October 2005), Sitrin stressed the value of having early sample material available through the product's life cycle.

"There are products that may be 10 or 15 years old where all of a sudden manufacturing questions come up" and firms would "die for a sample" of the material from the original clinical studies to compare with what they are making now, he said, quipping that "anyone who throws out a stability sample should be disposed of in the same manner."

CDER official Gubina voiced "complete agreement," noting that she has to "remind some companies to have enough retention samples to be able to compare their new lots to the old ones, and so on. And some companies, for some reason, forget to do it."

➤ Participants at the forum also focused on the definition of a reference standard, and whether the term was applicable to a standard that applied only to a specific assay.

"We might think about things slightly differently if it is a standard that is meant to act as the benchmark for all of the characterization or quality attribute assays for the product versus let's say" exclusively for potency, Merck Research Labs Vaccine Biometrics Research Director Tim Schofield commented. He generally thinks of an assay standard, for example, as different from a "reference standard" in that the latter would not reflect "minimum potency or maximum toxicity, and so forth."

Citing pH testing as an example, Gail Burnett, a senior director at Genentech, suggested that "the more generic the method…the more likely you are to use a reference material or an assay standard that differs from your primary material."

CBER official Chang commented that the importance of preserving primary reference materials and not exhausting them in early product development leads firms to develop secondary or "working standards" calibrated against the primary standard. The working standard is then used in the ongoing quality control and lot release assays.

Merck's Sitrin proposed that a semantic distinction could be drawn between an "archive" sample and a reference sample used for testing purposes. Material from the key developmental stages needs to be securely archived and not depleted as a working standard, he said.

Minimize Reference Standard Changes

The issue of what prompts a change in the reference standard was raised.

CDER pharmaceutical science official Webber stressed that a change has "to be data driven, because if you change your reference material to meet your process... there is no bridge between the previous process material and the new process. It really has to be determined on the suitability for its purpose....You don't want your process necessarily to drive your standards. You want your standards to drive the quality of your product and the process therein."

Genentech QC Clinical Development Director Wassim Nashabeh echoed the concern that "changing the reference too often when the process is changed" undermines the clinical bridging objective.

➤ Genentech, he noted, generally develops an initial reference standard that represents a pooling of a couple of lots produced for the IND tox studies, which "ideally, will take us through the Phase III program."

"All throughout we are doing comparability exercises to judge that any process changes we made throughout are still giving us acceptable material from a product quality perspective. As such there is no need to remake a new reference unless the profile is different enough from the analysts' perspective that they cannot use it as a comparator – where it gets quite difficult for them."

Beyond that initial IND tox material, Genentech develops its next reference standard from the Phase III clinical trial batches.

"We will then make a pool from typically the large-scale Phase III…BLA enabling material, and that becomes our primary reference which will then take it forward into commercialization." This becomes the primary reference against which all subsequent working reference standards are released, "so we maintain enough of the primary basically for the life of the product," Nashabeh explained.

He added a cautionary note against viewing a process change as a driver for a reference standard change, "because actually you are losing continuity throughout the process this way."

CDER's Kozlowski suggested that the broader question than what drives a change in reference material is "what is enough of a change to drive a comparability exercise of a certain level." The results of that, in turn, "would drive whether or not you need a new reference standard," he stressed.

Reformulation May Help Stability

The ensuing issue on which forum participants focused was the risk and/or benefits associated with different formulations of reference material compared to the drug product.

"Do the formulations really need to be the same?" moderator Schenerman queried, noting that there may be a better formulation for a reference standard stored at -70 degrees.

➤ Novo Nordisk scientist Jespersen maintained that at least for proteins, different formulations are appropriate where a very long stability period is sought for the reference material that the marketed product formulation will not accommodate.

Novo Nordisk has demonstrated that HPLC does not present problems when comparing the two, she reported.

The key here, Chiron's Karunatilake commented, is to make sure that no new degradation mechanisms are introduced in storing the reformulated reference material – a situation which he has seen occur.

Schenerman noted that MedImmune's standard practice is to freeze reference materials in liquid nitrogen as quickly as possible to prevent aggregation during freezing – a problem that was found to occur in a "particularly sensitive" recombinant protein. However, he noted, the firm does not typically adopt "the strategy of developing a totally separate formulation just for the purpose of freezing a reference standard."

Schenerman did agree with CDER's Gubina that a phosphored buffer may encourage aggregation formation during freezing, adding that aggregation also can be concentration dependent. "So it might be a perfectly good approach to just dilute your product to a different concentration and then freeze it. You might find it is a lot more stable under those circumstances."

Industry consultant Sally Seaver pointed out that proteins formulated either as vaccines or a combination product are often less stable in that final formulation than when they are frozen as drug substance in an appropriate buffer. When using the later as a primary reference in the potency assay, the standard has to be formulated to mimic the drug product. With the batch variations involved in formulating at a small scale, running the assay can be challenging, she pointed out.

Merck's Owen, explaining that she works on live virus vaccines rather than proteins, commented that the "broader point is that you really need to make sure that the assays are appropriately qualified for whichever matrix you are actually in."

She noted that Merck had a case where the drug product behaved differently in the assay when it was formulated versus when it was not formulated, showing a boost in potency. "So clearly the reference standard would have to reflect that sort of a characteristic. But the point is you have got to characterize the assay."

➤ Owen's Merck colleague Timothy Schofield asserted that the two basic premises involved are that the reference standard must be stable and must be representative of the analyte that you are testing.

A material should be a viable reference standard, whether it is lyophilized or frozen to keep it stable, as long as "you prove that whatever you have mixed with the material does not interfere in the assay – does not generate a different signal than you would of the

analyte you are trying to test," Schofield said. "But you have to do the studies."

Consultant Sajjadi added that the same principle applies to the container as the formulation. When using a container for freezing and doing aliquots in different sizes and head space for different assays, "you really have to make sure...that you are maintaining the comparability," she stressed.

The Process Needs To Be Representative

The forum participants also focused on the concerns surrounding the use of reference materials made before process validation to support commercialization.

Sajjadi pointed out that the concern about whether earlier stage processes are representative is heightened for more cutting edge technologies as opposed to products like monoclonal antibodies where there is significant licensing/manufacturing experience.

"If you don't have that history," she cautioned, "make sure that you are careful…pre-process validation to know really know which process you are talking about. Because you usually have a limited number of runs, without that history it becomes really important."

➤ Industry consultant Nadine Ritter echoed the point made earlier by CBER official Chang on the importance of factoring in the intended use of the reference standard.

For example, she said, "if the intended use of the material is to look at potency, where you are not typically measuring or monitoring the degradants that are increasing, –you are just looking at the decrease in the activity of the parent molecule – then...the batch could be from earlier development, because it is the activity of the parent that counts."

On the other hand, "if the intended use of that reference material is for monitoring fingerprint profiles of impurities or monitoring the production of degradants, in those cases it may be something which has got to be locked down on the last, most representative process that you had because it changes over time – preferably less impurities and less degradation. But I think you really have to ask yourself, what are you doing with the stuff? If it is activity, it may be one thing. If it is impurities or degradants it may be another thing."

Acknowledging Ritter's point on the importance of the intended use, FDAer Kozlowski cautioned that degradants "may do things you don't expect" and "aggregates can sometimes increase potency, so I think you need to be careful which ones you are excluding."

Kozlowski was seconded by fellow regulator Chang, who added that "if there is a degradation then you have to look at whether or not the degradation product will affect your potency assay. There has to be some qualification assays to justify that."

Reference Standard vs. Lot Release Specs

Moderator Schenerman shifted the discussion onto establishing the acceptance criteria for reference material vs. the lot release specs for the product.

Noting that MedImmune does use tighter criteria on the reference material, Schenerman explained the rationale: "You want to avoid the situation where, when you are qualifying from one reference standard to the next, you could be losing something – for example, potency. If you are qualifying one reference standard right at the lower limit and the next one is at the lower limit and the next one after that is at the lower limit, you could, over time, have a product with lower potency."

"You want to select a reference material which is the best reflection of your general experience with the process," Schenerman said. "So you really want to be centered in the area for the material as opposed to at the extremes of the range."

Other participants concurred. "Why create a headache? Why not pick the most normal looking lot that you have?" consultant Seaver suggested. Kozlowski pointed out that, given the variability in the two comparators, "if you have a reference standard that is on the edge, you are creating a more difficult statistical dilemma. So why not pick something as centered as possible so you have the best chance of actually passing lots that should be passed and failing lots that should be failed."

➤ CDER Office of New Drug Quality Assessment Pharmaceutical Assessment Lead Stephen Moore cited insulin as a precedent for tighter reference specifications in the case of potency.

For market release, a bioidentity test is done which is only four groups of two rabbits compared to a full insulin potency assay which uses more rabbits and gives a lower variance. "In the latter case you can actually have a range," Moore said, "whereas in a bioidentity test you only have a limit to meet."

Novo Nordisk's Jespersen noted that Novo's reference standard criteria are not tighter for the impurity profile, while "of course, there are much tighter acceptance limits on the confidence interval when we certify the content of our reference material."

FDAer Gubina cautioned that the agency often sees a drift in a product's reference standards over time toward one end of the specification, so that "most likely you need to have tighter specs to prevent" a serious problem from occurring.

Bristol-Myers Squibb Analytical Biochemistry Director Kirk Leister pointed out that the selection criteria for the reference candidate depends on "where you are in the manufacturing process history."

As firm's move into Phase III and commercial production, they gain "a lot more understanding of whether or not" the reference material established earlier is representative of the manufacturing process. "It may be representative within a three sigma of the process. However, it may be on the fringe, and many of your tests may be comparing to something that is not" really representative, Leister stressed. The question, he said, is "when you reselect a new reference standard, should you use the manufacturing process history to try to guide you in selection of something that is more centered to prevent future drift away from your manufacturing experience."

➤ Moderator Schenerman asked what participants would do if a reference standard is found during use not to meet the acceptance criteria.

Steven Wan, bioassay associate director at Centocor, cited two implications: "One is we have to investigate to see why it failed" – whether the assay is drifting giving an artificially failing result or the reference standard is "really starting to fail." A second action needed, particularly in the case of a potency assay where the result of the release data is based on the reference standard, is "to evaluate the training data to make sure," for example, that aggregation "does not cause the active potency value to be over or under estimated."

➤ Genentech's Burnett framed the discussion in terms of differentiating the role of a "reference spectra" from the reference material.

Having a reference spectra, she noted, allows the firm to go back and determine "what is really going on. Is it something that is drifting in my test methods or is it…truly my reference lot that something might have happened to?"

Burnett stressed that "there are a lot of occasions… where people use reference spectra potentially as the primary criteria to determine suitability on a day-to-day basis as opposed to using the reference material every time out." In troubleshooting, a place to start would be "to look at your reference spectra and see if in fact what you are getting today on your reference lot is still looking the same."

In the transition of the product from research, "only a limited amount of the purified drug substance is available as an initial research reference," she said. At this stage, the development of the primary structure of the research reference must be known based on peptide mapping and carbohydrate analysis.

Biogen Idec's Program for Developing Reference Standards

At the WCBP strategy forum in January, Elena Vasilyeva from Biogen Idec's QC chemistry department discussed the various stages of reference standard development and her firm's program and procedures for handling them. She noted that the talk was originally prepared by her colleague Mathias Kretschmer.

My talk will consist of two parts: The first one is about key milestones for the reference standard. Actually it is related very tightly to the stage where the product is in terms of development. In the second part I am going to talk about the system we have in place, and actually this is a reference standard SOP.

Interim Standard

The first stage is the transition of the product from research. At this point only a limited amount of purified drug substance is available as an initial research reference. At this stage of development, the…primary structure of the research reference must be known. This is based on a peptide map and carbohydrate analysis. Key assays must be defined and results documented in research reports. This includes electrophoresis, size exclusion chromatography [SEC], charge isoforms and potency assays -- in most cases it is a binding assay. The research reference must be also active in an animal model, for example, backing up the designated API, so we know that its purified molecules and not, for example, aggregates are active.

During early development, material from cell line selection, process, and purification development is compared against the research reference. Screening of formulations for drug substance and drug product is performed with results also compared against the research reference. The research reference is stored in most cases at minus 70 degrees C. Early process prototypes are compared against the research reference and may be used as intermediate standards for assay development, since again the amount of research reference is very limited.

The next stage is IND enabling tox studies. In this stage usually the cell line is selected, initial fermentation process is in place, unit operations for drug substance purification are defined, and formulation components for the drug substance and drug product have been set, even though, for example, concentration of the formulation buffer can be changed in the future. The prototype from this process is used in the first IND enabling toxicology study. A portion of drug substance from this prototype is aliquoted and characterized as the first interim reference standard. Characterization of this interim reference standard follows a pre-approved protocol per the reference standard SOP [which] I will talk about later.

The protocol for the characterization studies includes tests for identity, purity, potency and safety. All these assays will be future release assays and characterization assays also that establish primary structure and carbohydrate structure. At this point the research reference is used as comparator. For quantity, it is usually A280 assay, purity (usually SEC and electrophoresis) and safety (bioburden endotoxin). Quantitative specs are defined at this stage. Specifications are based on research data and early prototypes.

The specifications are based on tolerance limits that are calculated using a K factor. This is the formula to calculate this: Here you can see K plotted for 95% confidence and 99% acceptance rate. What is very important to keep in mind is that when you have a lot of data points, the key factor will be close to 3.0. So it is actually three standard deviations that we usually use. But when you have a limited number of data points, the key factor can be really high. This is very important to remember about this. So that is why we usually use at least six data points to establish this.

At this stage, many assays such as carbohydrate profile, charge isoform profile, oxidation and potency may have only 'report result' criteria. The sequence is verified by a mass spec peptide map and mass of intact, reduced, and reduced and deglycosylated molecules -- whatever is appropriate. Also the extinction coefficient is determined. It is good if this extinction coefficient is used throughout development unless a structural change is incurred.

Aliquoting and storage conditions -- usually, as I already mentioned, it is minus 70-degrees Celsius -- are also defined for the first time. The first interim reference standard will be used to test Phase I clinical material. The interim reference standard is placed on stability from zero to 36 months and is assigned a first re-test period of one year. After one year, the data are analyzed in a report and re-test dates extended appropriately with a maximum one-year extension.

[It is important] to use statistical analysis when a decision has to be made about the one year extension. To check the appropriateness of an extension for one reference standard for one year, we use a 95% confidence of prediction, not confidence of mean. So it will allow us to know where the next data point will be after one year for this one reference standard. I have to mention that if the decision is made about extension for one year, we continue to do real-time stability testing each three months.

GMP Standard

The next stage is the Phase I clinical study where the first GMP drug substance and drug product are produced. The first GMP reference standard is prepared and qualified per a pre-approved protocol, usually from the second manufactured drug substance batch just to give us more experience. The protocol builds on that for the interim reference standard. At this point, assays for characterization are typically the same as used before but more tests are added. For example, now we use a peptide map for disulfide assignment and an alternative peptide map for complete coverage of sequence.

If not done before, potency units should be defined for the first GMP reference standard to allow seamless scaling of subsequent batches to the first clinical material. All subsequent drug substance and drug product batches will be tested against the first GMP reference standard until either the reference standard expires or reference standard aliquots are used up, or significant process changes require a new reference standard.

The first GMP reference standard is placed on stability, also from zero to 36 months, and given an initial re-test date of one year. At this point, at least three months of real-time stability data are available from the pre-clinical interim reference standard on stability. As with the pre-clinical interim reference standard, the first GMP reference standard is followed on stability in real time. At the re-test period, a report is issued summarizing the data and assigning a new re-test date not greater than one year from the last one.

Phase II clinical studies often involve process changes. The default approach is to qualify a new reference standard for a changed process. Since the GMP reference standard from the new process will not be available until the new process is run at scale, new process material will be initially compared against the previous GMP reference standard or against a newly qualified interim reference standard.

The judgment about the transition has to be made in advance of GMP production, based on development data that we have from the new process. Significant changes that call for a new reference standard includes a new cell line, a new unit operation or any other changes with an analytically verified impact on the product. And one example is shown here: This is [a charge profile by cation exchange] test. You can see here one peak versus two peaks….And this actually might be the result of some changes in the process of purification of this material.

If material from the new process is exactly the same as the previous, and this is confirmed by comparison of the new GMP material against the previous process -- the GMP reference standard made with the old process -- in this case no new reference standard has to be qualified. If a new reference standard is qualified, the same rules for the GMP stability program and re-test dating apply as before, as I talked previously about for the interim reference standard.

The next stage is Phase III BLA enabling. In this phase, the formulation presentation and expiry of the reference standard have to be finalized and backed up with the stability data gathered during all previous phases of development. In this stage, more characterization assays have to be defined -- for example, secondary/tertiary structure and process impurities have to be well characterized so we know what we are dealing with. The reference standard stability program and re-test dating continue in real-time with a maximum one year extension.

Commercial Standard

Then it is the commercial final stage and launch with the reference standard from the commercial process using the reference standard qualifications protocol, which is submitted in the BLA. Aliquots sufficient for a three-year supply are prepared. At this point the reference standard is placed on stability and monitored in real-time with three month intervals, and the initial re-test may extend beyond 12 months based on prior GMP reference standard stability data.

The reference standard program at Biogen Idec is managed by Site Quality Operations. They provide annual stability reports to support the re-test extension. They manage the reference standard inventory and supply. They choose the replacement reference standard once the supply is exhausted. And the qualification of this reference standard if replacement has to be done follows the BLA-approved protocol. So it is all the same protocol.

Reference Standard SOP

Now I am going to talk about the system we have in place, which is the reference standard SOP. The SOP includes scope, qualification of a new reference Standard [retest dating, re-qualification], replacement of the reference standard, control and storage, and documentation.

Scope: This reference standard SOP applies to all reference standard material used by the Quality Department for release and stability testing of clinical and/or commercial material. It also applies to any proprietary reference standard materials employed at third party labs for contracted testing of clinical and/or commercial materials.

Qualification of a new reference standard: A Reference Standard qualification protocol is prepared and submitted to Quality Control for review. Successful execution of the protocol serves to qualify the propriety substance as a reference standard. A Certificate of Analysis for the newly qualified reference standard is generated. The CoA is filed by Quality Assurance.

About re-test dating of the reference standard: Newly qualified reference standard material is assigned a re-testing period of 12 months and placed on long-term stability. Re-test dates may be extended based on annual review. If at any time during the re-test period the stability data indicate the reference standard is not stable, the shelf life will be set to reflect the expiry. The established shelf life may be adopted for future replacement reference standard batches.

Part of the SOP talks about re-qualification of the reference standard. Material that has not been evaluated at the re-test period may be re-qualified. Such material must be placed under quarantine and cannot be used as reference standard until re-qualification is completed. This can happen, for example, when the project is on hold and then comes back. The batch of material to be re-qualified is tested according to the appropriate reference standard qualification protocol.

Replacement: Reference standard material…that reaches the end of its shelf life can be replaced with a new batch. The procedure is identical to new reference standard qualification and the existing protocol may be used.

For control and storage of reference standard: The reference standard is controlled at the site's Sample Control Area. The reference standard is stored under the appropriate storage conditions as directed by the qualification protocol. Sample Control shall notify local labs if the reference standard is placed under quarantine.

There is a list of information that usually is kept with reference standard: [receipt date; reference standard number; chemical/material name; batch number; manufacture date; manufacturing site; quantity; CoA; storage conditions; re-test or expiration date; and any special directions or comments].

Handling and storage of reference standard in the labs: Receipt and quantities of local supplies of reference standard are recorded in a logging system. When the local reference standard supply is near depletion, re-supply is requested from the appropriate Sample Control Area. When the shelf life is reached and with approval by Quality Control, all remaining supply has to be disposed, which is recorded in the reference standard logging system.

Documentation: QA Documentation maintains all qualification protocols, reports and CoAs, and all stability protocols. Sample Control Area maintains copies of CoAs and shipping documents. And Labs maintain copies of CoAs and shipping documents.

Qualifying Of Protein Potency Standards Addressed

To lead off the forum's afternoon session on reference material qualification and stability evaluation, CBER official Chang gave a presentation that focused on the issues as they relate to potency standards for therapeutic proteins.

In qualifying a potency standard, Chang stressed that it is "very important" to have linearity of dose response over a wide range of concentrations. For example, concentration ranges that are not linear and not parallel between the firm's testing results and lot release by CBER "will create a very bad situation," Chang cautioned.

Inter-assay and inter-laboratory variability in test results for the bulk drug substance will create problems as well; for example, when a firm opens a new manufacturing facility and discrepancies appear.

Chang noted that international or national potency standards are normally selected from two or more candidates prepared by different methods, with the intention that they will be used by different laboratories for different purposes including quality control and clinical monitoring.

➤ FDA similarly likes to see institutional or global collaborative studies for in-house standards involve at least two candidate lots and more, if possible.

It is "very important" that the calibration be done against both the international/national standard, where available, and the current in-house potency standard, Chang stressed. "You do not want to deviate not only from your in-house unitage that was established before, but also from the unitage that is carried by the international standard, because this unitage is defined and used by a variety of different laboratories."

A comparison should be done with the SOPs and methods used in the clinical labs, with different methodologies "hopefully" giving the same result.

Chang recommends that firms consider using global collaborative studies to prevent an institutional bias in the in-house standard.

A central lab should then collect and analyze the data. "You should have a mean potency estimated for each lab and also have an overall weighted mean for all the data that you collected," he advised. Results should be included from "all assays with only a few exceptions that can be justified, such as those which are statistically invalid."

Citing Biogen's program as explained by Vasilyeva as an example, Chang recommends that firms have "an institutional quality control unit on standards," preferably separate from the unit developing the standard "to reduce the bias and double check."

In minimizing the frequency of preparing new standards and the related potential for potency drift, it is important to prepare a "very stable primary potency standard," Chang stated. Most of the international standards for protein products are stored freeze-dried in ampoules, to reduce moisture and eliminate oxygen, with storage at -20 C in the dark.

The FDAer advises firms to know the development process for, and participate where possible in, the

international calibration studies to help in minimizing the drift.

FDA Spearheads Standard For Factor VIII

In the later part of his presentation, Chang reviewed the study process used to establish the international standard for von Willebrand Factor (Factor VIII).

Before FDA licensed the first product for the disease indication, the agency determined that an issue slowing product development was assay variability and the need for an international standard.

- ➤ Two phases of the study were conducted – the first involving qualification, followed by production and calibration of the standard.

For qualification, five products were selected from five different manufacturers. FDA, the United Kingdom's National Institute for Biological Standards and Control (NIBSC) and the Science and Standardization Committee under the International Society of Thrombosis and Homeostasis were selected to carry out the assessment of potency, purity, integrity, stability and assay performance of the various VWF concentrates. Two candidates from the five were chosen for further development based on the qualification work.

Stability was "the number one criteria" used in the selection process, Chang said, with parallelism of dose response curves also a primary criteria. Secondary criteria included similar results with different potency assay methodologies, a ratio between active and antigen of close to one – mirroring the native protein in the plasma – and the integrity of von Willibrand Factor multimer forms.

The slopes of the dose-response curves were not statistically different among four of the five candidates, and all five were found suitably stable for use as international standards. With the primary criteria thus not determinative, the secondary criteria were used to select the two candidates.

The two candidates and the study information were presented to the WHO Expert Committee on Biological Standardization which selected one as the first international standard for VWF concentrates.

Merck's RotaTeq Strategy Includes "Gold Standard"

Chang was followed at the podium by Merck Research Labs' Katey Owen, who addressed assigning potency to reference standards for live viral vaccines and Merck's strategy in particular with ***RotaTeq***. The vaccine was recently approved for the prevention of rotavirus gastroenteritis, which entails significant morbidity for children in the U.S. and mortality worldwide.

RotaTeq is a combination of five different viruses. It is refrigerator stable, with a proposed shelf life of two years at 2-8 degrees C and no potency loss detected to date when stored at -70.

Potency is assigned in many live virus vaccines by a plaque assay. However, Owen explained, the assay can't distinguish between serotypes of different viruses and is variable, operator dependent, time consuming and very labor intensive. The one advantage is that it is not dependant on a relative potency assignment, providing an endpoint dilution.

For RotaTeq, Merck developed "what we like to consider our next generation infectivity assay or potency assay" using PCR technology to quantitate the amount of replicated DNA/RNA and then report the potency based on relative potency to a standard.

The quantitative PCR-based assay provides the desired multivalent specificity, is easier and faster to operate, with "very high throughput," resulting in reduced standard deviation.

- ➤ Merck's strategy was to link the primary reference standard for the potency assay to the clinical material, with the working standards calibrated back to the primary standard.

In developing the standard, Owen reported, Merck started with a "pivotal reference standard lot" manufactured with GMP supplies, which had "sister lots" that went into the pivotal clinical safety studies. Potency was assigned to this pivotal lot based on monovalent standards with specified potency using the plaque assay to provide an endpoint dilution. This involved 100 assays over a six-month period.

The pivotal reference standard lot was then used to assign the potency to the pivotal dose confirmation clinical lot, which was also used for the pivotal expiry study for the product. The reference lot was also used for process validation studies.

Because Merck began to run out of this material, potency needed to be assigned to a new reference

standard lot, for which 60 assays were conducted. The new reference standard was then used in assessing the product launch lots.

The "important point," Owen stressed, is that "the pivotal reference standard lot had a very close tie to the clinical studies" as well as the manufacturing process.

➤ Another part of Merck's strategy involved creating a "gold standard" from a manufactured lot to which potency was assigned based on the pivotal reference standard lot.

The gold standard is stored in liquid nitrogen using a different container from that used in the final product. "We have no expiration date assigned, but clearly we are testing it across time," Owen explained, with the expectation that "this particular standard will last us for something like 100 years of manufacture."

This gold standard is used to calibrate the potency of the working reference standards. The latter are stored at -70 with a five-year targeted use period, after which another working reference lot will be manufactured.

Summarizing the RotaTeq reference standard strategy, Owen stressed that the "tie to the clinic is really of primary importance" and that "the actual numbers are not nearly as important as the relative potency as characterized compared to clinical material."

She also emphasized the importance of trending the standard data over time. With the cell-based assays for live virus vaccines, "we tend to see cycling of the assay itself. You can see something start to head down and you are not really sure, do we have a stability issue or do we have an assay issue? So trending across time helps us get to the bottom of those questions."

➤ Owen also addressed the objectives of Merck's reference standard development program for the HIV adenovius vaccine on which the firm is working.

"The goals for reference standards in early development are slightly different from the goals in later development and launch," the Merck official pointed out. Echoing CBER's Chang, she explained that in ensuring product consistency, "sometimes you actually need to compare across companies or across laboratories, so reference standards can help with that." They also assist in "clearly being able to compare across phases of clinical studies, making sure the Phase II material is the same as the Phase I material, or at least behaves the same."

Another role is ensuring product consistency when multiple lots are needed to supply a particular clinical study. "You start a clinical study. You think you have enough material. Turns out the clinical study goes a lot longer than you expect, so your supplies either expire or you run out of your supplies and you have to make a new lot." Owen emphasized that making sure that the second lot is identical to the first "is especially important" for the final expiry study in defining the vaccine end-potency.

Further, she added, reference standards in development "clearly need to ultimately be a bridge between your development activities and manufacture."

Novo Nordisk Emphasizes Stability, Homogeneity

Anne Jespersen followed Owen with a presentation on Novo Nordisk's strategies for qualifying reference material candidates for well-specified recombinant proteins.

Critical issues in this regard, she stressed, include:
● the amount of characterization necessary for primary vs. secondary/working reference materials
● homogeneity, particularly of the working standard
● how the content and potency of the primary reference is determined ● the mass balance of the primary reference material ● how changes from one reference material batch to another are handled with respect to calibration, and ● how stability is followed on the reference material.

➤ Jespersen explained that Novo Nordisk generally uses freeze-dried protein containing no excipients for its primary reference material in order to allow for the determination of mass balance of the total protein. If possible, a released batch of drug substance is selected. "It must be as stable and homogeneous as possible," she stressed.

The secondary reference material or "working standards" are in the ready-to-use formulation from a drug product batch, if possible – again, as stable and homogenous as possible.

Identification for the primary reference material, which is "at the top of the hierarchy," is "very important, so

we do an extra n-terminal amino acid sequence and mass spectroscopy," Jespersen explained. The identification of the working standard is less extensive, "but we do compare it to the primary reference material by a peptide map or HPLC," she said.

Novo Nordisk does not generally do additional types of purity analysis on the reference material beyond what is deemed sufficient for products, and the impurity specification limits are the same.

Jespersen explained that the degree of homogeneity or vial-to-vial variation can be taken into account when calibrating the secondary reference material against the primary standard. "That means that if you have a bit higher variation between vials on your primary, you can introduce more vials to the calibration study, and this way compensate for the variation."

However, she stressed, for the secondary material in daily use, the homogeneity of the batch is "very important." Otherwise, "if you use maybe one vial in each analytical setup, then your analytical results will vary from day to day corresponding to this variation."

There are two ways to determine this homogeneity, Jespersen said: by weight variation during filling as outlined in the pharmaceopieas; and by performing content uniformity on the filled containers to determine the content of the vials. Novo Nordisk uses the latter as the better approach, she said.

➤ Jespersen stressed that when setting product specifications, firms should be aware that they need "to correlate with the homogeneity of your reference material."

In ensuring precision and accuracy, she highlighted the importance of using qualified laboratories that "really know the assay and are trained in it." If Novo Nordisk has only one lab, "then we vary the days, the equipment, the personnel and the columns," she said. Six or more independent setups are often used, with the amount of samples dependent on the homogeneity and analytical uncertainty.

The shelf life specifications on the reference material should include the analytical variation of the method and the acceptable decrease or increase, Jespersen stated. Novo Nordisk evaluates any significant trend "even if all the stability study results are within the shelf-life specification." She added that "of course, the shelf-life specifications of the drug substance and the drug product ought to include the allowed change in content when changing the reference material batch."

Another important component, Jespersen stressed, is having "a clear strategy" on how a change in the primary reference material batch will be handled.

Potency Definition Impacts Reference Changes

In the discussion period following Jespersen's presentation, the issue was raised of how potency is defined and how that impacts on changing reference standards.

LGC official Helen Parkes commented that the WHO is working with the International Bureau of Weights and Measures "to try and get an alignment of SI units with international units with the recognition that companies do actually have problems with the international units and batch variability and not being able to trace back to an original standard." She noted that IBWM "obviously is interested in traceability to the SI. So the organizations are actually working together at the moment to try and come to some level of resolution in this particular area."

Before using mass to label the product dose, FDAer Chang asserted that firms should know "exactly what specific activity you have." He underscored a point made by Jespersen that "if the specific activity varies significantly, to label the product with a mass often is not proper." From a regulatory point of view, the CBER official cautioned, "it could be considered misbranded."

Jespersen noted that for Novo Nordisk's product Factor VII "there is a WHO standard which is defined in units, which we use as our primary reference material for the unitage. But it is not defined in milligrams. So we have to have our own company reference material in milligrams." The result is that Novo Nordisk in effect has "two parallel running primary reference materials."

Amgen scientist Venkat Mukku asked Chang about using the international standard for calibration when the candidate primary standard is being produced from a different process. "What if the dose response curves are not parallel?" Mukku queried, noting that "most of the time, the assignment of potency is based on the premise what you are comparing [has] a similar dose response curve. How do you deal with that?"

Chang responded that the question is "a very good" one. The standards setting groups were sensitive to the issue in developing the international von Willibrand Factor standard and included parallelism of the marketed products in making its selection.

- Recognizing that the lack of parallelism presents a tough situation, Chang advised that firms should "give a lot of thought to the best way not to deviate your unitage" from the international standard. "Reduce the assay variation" and try to "establish one that can give you a parallel relationship."

LCG's Parkes, who has been actively involved in the development of international standards, expressed concern "about the way that the international unit is derived" and "the inherent bias which I believe can be introduced into the assignment of a value by the WHO's use of many methods." By contrast, "progressive biopharmaceutical companies," she stressed, "would not think… of not using the leading edge analytical method."

Parkes suggested that "taking the best performing method for doing method assignment" would be a step towards "giving an assignable value which is perhaps more in the units that most of us would use."

Chang commented that he had prepared a concept paper for the VWF international standard that called for not only generating material, but a standardized method as well. The reason the latter was rejected, he explained, was concern that the standard method might systematically measure the wrong thing – that "before you know what should be measured in a consistent way that reflects a clinical outcome, you should not develop a method that is standardized because you do not know whether or not you standardized a wrong thing."

Parkes responded by acknowledging that the "recognition that the clinical application of the method to what you are actually trying to find…is a positive part of the WHO standards." However, she maintained that it is possible to move beyond a case-by-case approach, as NIBSC is trying to do, "in areas such as protein quantification, nucleic acid testing, where perhaps there is a more clear cut root to coming up with some consensus assignment values that might give us traceability longer term."

Document Qualification Rationale

In addressing the issue of how a reference standard should be qualified, CBER's Chang suggested that "for your identity test or potency test, you should have some predefined criteria that suit your intended purpose….Then when you start to qualify your candidate, you probably will learn additional things." Firms should try to stick to the criteria that has been defined prospectively based on the intended use, and any modification would need to be justified, he said.

Lilly Principle Regulatory Scientist John Dougherty, who was moderating the afternoon discussion, pointed to general agreement that "as we go to qualify these things, we want to make sure that we are very clear on, and very succinctly state the purpose and the use of the reference standard, and then drive that qualification protocol off that stated purpose."

Morning discussion moderator Schenerman added that documentation was referenced as important in this context. "As you document your qualification process for your reference material, it would probably be worthwhile to also document your rationale – what was your thought process for selecting each of these tests and whatever ranges you are establishing, and how did you come to that decision as well."

- Industry consultant Sally Seaver stressed the inherent problem in trying to meet the objective of relating the assay assignment to how the product is working *in vivo*.

"Very often a good therapeutic has multiple modes of action" that are not always well understood, she noted. "A lot of times," the assay that is chosen is the one "that works most reliably in our QC lab."

Chang responded that not precisely knowing the mechanism is "why in our regulation, clinical evaluation can be considered as an *in vivo* assessment for the potency" and the product licensed accordingly. However, he pointed out, "without knowing the mechanism of action, it becomes very difficult to handle manufacturing changes after licensure." The emphasis is placed then on "sticking with the process" and the related standard.

CDER's Webber concurred that for "a lot of products, we don't know what the method of action is, and in those cases certainly you do the best you can. And you do want to try to find an assay which is precise and accurate." But where possible, "try to correlate it as best you can to your expected method of action, and [don't] choose an assay...just because it is simple and easy to run."

Ad page.

➤ USP official Larry Callahan asked the participants to comment on "what the best strategy is" for assigning the potency value on a second-generation reference standard.

"Are you better off doing it in your best lab with your best technician and coming up with a relative value for the reference standard, or are you better off sending it…to 10 labs all around the world and trying to average" and weight the results "and see which ones you should throw out or not throw out?" Callahan queried.

Genentech's Nashabeh said that his firm's practice: "If we have multiple labs, multiple countries or multiple sites that are utilizing or testing a product, then the qualification of a new reference is done across sites." A statistical analysis is done "as far as the number of replicates." These are "usually equally weighted across the multiple sites."

Genentech then analyzes the data and assigns the actual value. "We do not round it down to a 100 or one. The value that we actually measure is what goes on our CFR which we will then use for all future reference," Nashabeh explained. "Typically, we will have a minimum of 20 runs performed across multiple sites."

Moderator Dougherty commented that the number of labs involved depends on the stage of development. "Earlier on you are going to have one lab running" the method. On the other hand "at Lilly, by the time we are hitting the commercialization phase, we have got this running in multiple laboratories….We try to build as much variability into that equation as we can to make sure we are capturing all the elements of the variability of the method."

Callahan pointed out that qualifying the reference standard is different from qualifying the method. For the former, he said, "you only want to assign one number to it and have it as close up to reality as you can. You want to do it under your best circumstances. You are not looking at a ruggedness or a robustness test. You are really trying to get your best number."

In summarizing the form discussions around qualifying and assigning potency to the reference standard, Dougherty pointed to agreement on the importance of what is on the label. Whether "you are filling your product based on protein content or based upon biological activity...will certainly at the end of the day drive your strategy around this," he said.

➤ The forum stability discussions, Dougherty noted, focused on the importance of appropriately storing the reference standards to maintain their quality and monitoring them over time.

"The material provides a consistent point of reference for variability and process and analytical methods, so you need to be able to distinguish these changes from assay changes over time." he said. "So the stability data is in fact critical."

A show of hands among participants indicated that most firms are executing formal stability studies as part of their qualification protocols for the reference materials, with some placing the materials in structured, forced degradation and accelerated stability conditions. Some participants indicated that their firms are also focusing on shipping and handling concerns, either prospectively or when problems occur.

For protein standards, there was agreement that the main focus during stability is on factors other than potency, which is not generally a leading indicator of problems.

Novo Nordisk's Jespersen commented that the focus for proteins is on degradation and impurities which will allow much earlier detection of problems rather than potency testing, with its inherent variability.

Genentech's Nashabeh concurred that for well-characterized proteins "you have on your stability program a multitude of very sensitive analytical assays that are quantitative in nature and" indicate a change in the charge or aggregation profile or fragmentation. "So these are what is really giving you the stability of your reference." While potency is part of the regimen, he asserted that potency is never "the first one going down."

Merck's Owen pointed out that the situation is different for viral vaccines. "Generally potency is the first thing to go," she said. ♦♦