

Research Article

On the Shelf Life of Pharmaceutical Products

Robert Capen,^{1,13,14} David Christopher,¹ Patrick Forenzo,² Charles Ireland,³ Oscar Liu,⁴ Svetlana Lyapustina,⁵ John O'Neill,⁶ Nate Patterson,⁷ Michelle Quinlan,⁸ Dennis Sandell,⁹ James Schwenke,¹⁰ Walter Stroup,¹¹ and Terrence Tougas¹²

Received 15 June 2011; accepted 8 June 2012; published online 23 June 2012

Abstract. This article proposes new terminology that distinguishes between different concepts involved in the discussion of the shelf life of pharmaceutical products. Such comprehensive and common language is currently lacking from various guidelines, which confuses implementation and impedes comparisons of different methodologies. The five new terms that are necessary for a coherent discussion of shelf life are: true shelf life, estimated shelf life, supported shelf life, maximum shelf life, and labeled shelf life. These concepts are already in use, but not named as such. The article discusses various levels of “product” on which different stakeholders tend to focus (e.g., a single-dosage unit, a batch, a production process, etc.). The article also highlights a key missing element in the discussion of shelf life—a Quality Statement, which defines the quality standard for all key stakeholders. Arguments are presented that for regulatory and statistical reasons the true product shelf life should be defined in terms of a suitably small quantile (e.g., fifth) of the distribution of batch shelf lives. The choice of quantile translates to an upper bound on the probability that a randomly selected batch will be nonconforming when tested at the storage time defined by the labeled shelf life. For this strategy, a random-batch model is required. This approach, unlike a fixed-batch model, allows estimation of both within- and between-batch variability, and allows inferences to be made about the entire production process. This work was conducted by the Stability Shelf Life Working Group of the Product Quality Research Institute.

KEY WORDS: ICH method; quantile for distribution of batch shelf lives; random-batch model; shelf life terminology; stability.

INTRODUCTION

Since 1979, the Food and Drug Administration (FDA) has required that all prescription drugs have a shelf life (or expiration date) indicated directly on the container label.

Similar requirements are in place in the European Union and around the world. The International Conference on Harmonisation (ICH) of Technical Requirements for the Registration of Pharmaceuticals for Human Use guidance document Q1A(R2) (1) (ICH Q1A) defines shelf life as, “The time period during which a drug product is expected to remain within the approved shelf life specification, provided that it is stored under the conditions defined on the container label.” Although this is an accepted definition, a crucial-for-implementation first question that arises is: what is meant by “drug product”? A manufacturer may think it is the entire collection of individual units (e.g., tablets) released as one batch. An inspector may think it is the particular sample of units taken from the batch and placed on stability. A patient may think it is an individual dosage unit. This is an important question since it relates to how shelf life should be defined, which in turn guides how the data should be analyzed and how the results should be interpreted.

Somewhat surprisingly, an implementable definition of the term “product” is lacking from all standard-setting documents and even legal statutes, to the best of the authors' knowledge. For example, the US Food, Drug and Cosmetics Act (FDCA) defines the term “Drug” as “articles recognized in the United States Pharmacopeia or National Formulary” [FDCA 201(g)(1)(A)], without any clarification—either in the FDCA or USP/NF as to the amount of a particular article that

¹ Nonclinical and Pharmaceutical Sciences Statistics, Merck, West Point, Pennsylvania, USA.

² PDU4 Team, Novartis, East Hanover, New Jersey, USA.

³ Regulatory Affairs, Colgate Palmolive, Piscataway, New Jersey, USA.

⁴ Pharmaceutical Sciences and Clinical Supplies, Merck, Summit, New Jersey, USA.

⁵ Pharmaceutical Practice Group, Drinker Biddle & Reath, Washington, District of Columbia, USA.

⁶ Corporate Quality, Boston Scientific, Natick, Massachusetts, USA.

⁷ Inspiration Biopharmaceuticals, Laguna Niguel, California, USA.

⁸ Clinical Pharmacology Biostatistics, Novartis Oncology, Florham Park, New Jersey, USA.

⁹ S5 Consulting, Lund, Sweden.

¹⁰ Applied Research Consultants, New Milford, Connecticut, USA.

¹¹ Statistics, University of Nebraska-Lincoln, Lincoln, Nebraska, USA.

¹² Analytical Development, Boehringer Ingelheim, Ridgefield, Connecticut, USA.

¹³ Merck & Co. Inc., 770 Sumneytown Pike, WP37C-305, West Point, Pennsylvania 19486, USA.

¹⁴ To whom correspondence should be addressed. (e-mail: robert.capen@merck.com)

would constitute a “product”. The lack of a quantifiable definition could mean, in one extreme interpretation, that all instances of a drug article (including all of those from different manufacturers) collectively are considered “product”, which creates obvious practical and regulatory difficulties. In another extreme interpretation, it could mean that each instance of an article (e.g., each individual tablet) is considered a “product”, which also leads to logical absurdities (e.g., one tablet = one product; two tablets from the same bottle = two products?). A common, clear, explicit, and implementable definition of “shelf life” and “product” should be promoted among all stakeholders to avoid miscommunication of essential information.

In 2006, the Product Quality Research Institute (PQRI) established a Stability Shelf Life Working Group (referred to as the “Working Group” in this article) with the mandate to investigate current statistical methods for estimating shelf life based on stability data, and if possible to investigate and develop an improved method (2). The Working Group comprises pharmaceutical, regulatory, and statistical scientists from industry, government, and academia. As one of its first actions, the Working Group reviewed available literature and applicable guidelines, and discussed current industry and regulatory practices related to determining the shelf life for pharmaceutical products. Different issues with current practices were discussed along with possible statistical approaches to resolve them. It soon became apparent that the term “shelf life” is used to describe different concepts in the scientific literature, and that there is no formal agreement on how shelf life should be mathematically defined.

A formal and mathematically strict definition of shelf life is required as the basis for the development of statistical techniques for shelf life estimation. Without such a definition, it is difficult to compare different estimation approaches because it is not clear what is to be estimated. The Working Group engaged in discussions to review and summarize available descriptions of shelf life, evaluating their benefits, drawbacks, and consequences in order to better target the appropriate research question for statistical discussions. Key results from these discussions are presented here to raise public awareness of the existing different interpretations of shelf life and to stimulate a broader public discussion on this topic, which is relevant for drug products, drug substance, clinical supplies, etc. In this process, the Working Group has considered existing guidelines but sometimes taken the liberty to question elements of these for the purpose of potentially developing an improvement.

Customer Expectations

The customer has a reasonable expectation that a prescribed drug is labeled clearly, performs as expected throughout its labeled shelf life, is safe and effective, and is available when needed. The quality of a commercial pharmaceutical product is a direct result of using quality raw materials in a well-designed, understood and executed manufacturing process. Prior to regulatory approval of a product, an agency expects the manufacturer to propose and justify the specific quality attributes to evaluate, how each attribute will be tested and what the acceptance criteria will be that each attribute has to meet. The manufacturer must provide convincing evidence

to the agency that upon release of a drug product batch, the selected attributes and corresponding acceptance criteria are sufficient to ensure that customers' expectations are satisfied with high confidence. Yet because of ambiguity in how acceptance criteria are sometimes defined, especially in regards to establishing shelf life, what constitutes “convincing evidence” is not well-defined.

Specification, Acceptance Criteria, and Test Plan

A key difference between the pharmaceutical industry and most other industries is that while a pharmaceutical specification (3) includes acceptance criteria and perhaps the corresponding test plans (also called test protocols or sampling plans), it generally does not include the underlying Quality Statement that describes the manufacturer's commitment to the customer. A test plan stipulates how much data should be collected (i.e., sample size), how they should be obtained and analyzed, and the level of statistical risk (or confidence) considered acceptable. An acceptance criterion listed in a specification can be constructed in many different ways, but it is inextricably linked to a particular test plan. The same numerical limits/acceptance criteria applied to different test plans may imply drastically different quality requirements. The test plan and acceptance criteria should ideally be designed based on statistical concepts that relate these requirements in a known way to the Quality Statement for the product.

Quality Statement

All commercial products, including pharmaceuticals, should have a specific requirement for each controlled quality attribute—a clear, transparent statement, independent of test plan, which defines the quality standard for that product for all pertinent stakeholders. A Quality Statement must be both achievable and testable, providing maximum and practical assurance of the acceptability of the quality attribute. The Quality Statement should form the fundamental basis for developing release and stability acceptance criteria for the quality attribute. There is much that the pharmaceutical industry could gain by considering and adapting quality systems and approaches employed in other industries for making quality decisions that have critical consequences. Excellent examples of established industry standards that provide scientifically valid acceptance criteria designed for a range of typical Quality Statements have been issued by ISO, ANSI and other organizations (4–7) and in the literature (8). For shelf life determination, the content of a Quality Statement is important in that the claims of the Quality Statement are part of defining the shelf life. This, along with an understanding of how the stability-limiting attribute (e.g., level of impurities or degradants) changes over time, drives the statistical analysis of the data.

Conformance to the Quality Statement

For every quality attribute, a manufacturer should develop a test plan to check conformance with the corresponding Quality Statement, as part of an overall quality strategy. For example, the Quality Statement “true batch mean within 90–110% of label claim and true within-batch standard deviation not more

than 5%” could be verified by the test plan requirement “average of 10 test results within 92.3–107.7% of label claim and standard deviation of 10 test results not more than 2.7%.” Further, the sampling plan and testing requirements should be designed based on statistical concepts that relate the requirements in a known way to a quality standard. For example, “if the test plan requirements are fulfilled, the confidence is 95% that the Quality Statement holds.” The sampling plan and testing requirements then inform and enable good decision making for the disposition of a particular batch based on its likelihood of conforming or not conforming to the Quality Statement.

ICH guidance documents Q1A and Q1E (9) provide recommendations for establishing testing requirements to evaluate stability data and estimate shelf life. A succinct description of these requirements is:

Sample a minimum of three batches, measure the critical attribute(s) over the storage time periods recommended in ICH Q1A, perform a statistical analysis of the stability data as described in ICH Q1E, estimate the shelf life as the storage time when a 95% confidence limit crosses the acceptance boundary.

Shelf life is an inherent property of the pharmaceutical production process and is therefore defined independently of the sample size used to estimate it. The estimate of shelf life, using ICH Q1E methodology, does depend on sample size (in particular, the number of batches) as well as the level of confidence. While no explicit quality statement is provided, the intent of the ICH Q1E strategy is to establish the storage time during which the critical attribute(s) will be considered acceptable for all “future batches manufactured, packaged, and stored under similar circumstances.” Unfortunately, as will be discussed in detail later, the statistical methodology recommended in this guidance document is incompatible with this intent.

A NEW LANGUAGE FOR SHELF LIFE

The ICH guidance documents provide a narrow framework for considering the shelf life of a pharmaceutical (or drug) product. Indeed, the term “shelf life” itself is not well-defined unless placed in the proper context. A broader understanding requires a clear terminology that distinguishes between different concepts often involved in the discussion of shelf life. Five terms are presented here to enable a coherent discussion about shelf life in its various contexts. Some of these terms already exist in scientific discourse but are rarely, if ever, recognized as distinct and different entities, leading to misuse by industry, regulatory agencies and academia. In casual conversations, when little care is given to precise terminology and nomenclature, the same vague term “shelf life” is applied loosely to all these different concepts, creating confusion and preventing progress.

The proposed terms, which are further explained below, are the following:

- true shelf life
- estimated shelf life
- supported shelf life
- maximum shelf life
- labeled shelf life

The *true shelf life* is the true but unknown limit on the period of storage time during which the pharmaceutical or drug product is considered fit for use and effective. In this context, the true shelf life can also be referred to as the *true product shelf life*, to be most specific. It is this unknown storage time, the true product shelf life, which is to be estimated through a stability study. Further discussion of what the Working Group means by “drug product” is provided below. Note that because the true product shelf life applies to current and future batches, it only has meaning when the manufacturing process is under a state of statistical control. Otherwise, batches manufactured today may not be representative of batches manufactured in the future.

A stability study is a designed experiment where the pharmaceutical product is stored in environmental chambers and followed for a prescribed amount of storage time. Periodically, the product is sampled to measure a series of stability limiting properties. From these data, an estimate of the true product shelf life is obtained. In general, this estimate of the true product shelf life is called the *estimated shelf life* or the *estimated product shelf life*.

Any suitably conservative estimate of the true product shelf life, as supported by statistical calculations, is called the *supported shelf life*. It is intended to be a conservative estimate of product shelf life to help assure that a high proportion of product remains fit for use up to that estimated storage time. For example, in the ICH Q1E guidance, the supported shelf life is the time point where the 95% confidence limits (one-sided or two-sided depending on the properties of the stability limiting characteristic being measured) intersect the acceptance limit(s), as illustrated in Fig. 1 for the case of a response that increases over time.

The *maximum shelf life* is the maximum allowed extrapolated product shelf life estimate based on the decision tree provided in ICH Q1E. The decision tree provided in ICH Q1E is a series of questions resulting in a limit to how far an estimated product shelf life can be extrapolated beyond the maximum storage time measured in a stability study. For example, if a 12-month stability study was being considered, any shelf life estimate may be limited to a maximum of 18 months of storage time (1.5 times the length of storage time considered in the stability study) by following the ICH decision tree. Note that the maximum shelf life is not dependent on data, but rather only on the length of storage time considered in the stability study.

Because the intention of the shelf life claim made by a manufacturer is that the true product shelf life is equal to or longer than the *labeled shelf life* (with high confidence), the labeled shelf life must be defined as the shorter of the supported shelf life and maximum shelf life. The labeled shelf life is what is printed on the drug product's label and is used to calculate the expiry date.

Batch and Product Shelf Life

A pharmaceutical product is typically manufactured in batches. A batch is a fixed quantity of product, for example, 100,000 tablets. Each batch is distinguished by its own shelf life which can be called the *true batch shelf life*. The batch is a single sample of the pharmaceutical product's manufacturing process at a specific point in time. More precisely, it is a

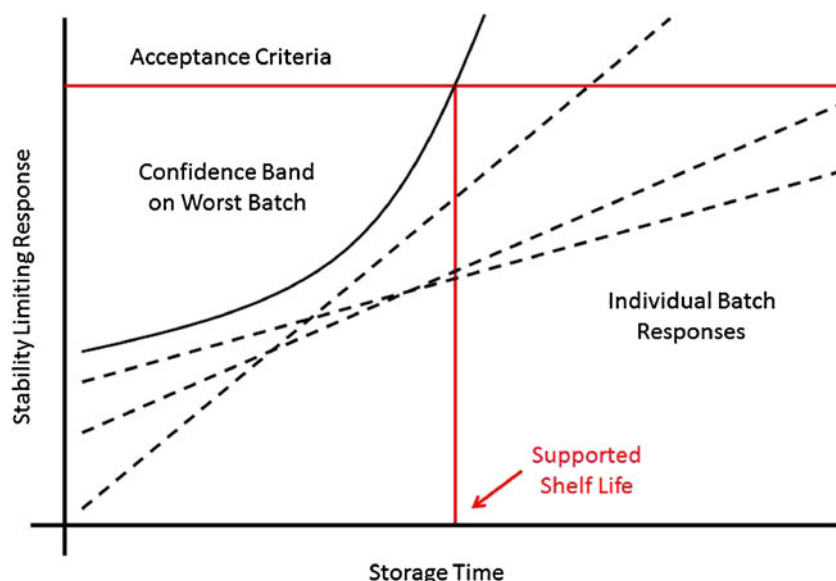


Fig. 1. ICH recommended method for estimating the supported shelf Life without pooling batches. The *Y-axis* represents the value of the mean response of the stability limiting characteristic and the *X-axis* represents storage time in months. The *dashed lines* represent the estimated regression of the change over time for three hypothetical batches. An upper acceptance criterion is needed since the batch trends depict a stability limiting characteristic that is increasing over time. This example shows the consequences of a case where, following the ICH suggested analysis, batch data cannot be pooled. This leads to using the “worst-case” batch, the batch intersecting the acceptance criteria first, to estimate the supported shelf life

realization of the manufacturing process subject to random variation. A specific batch may have stability characteristics that are slightly better or slightly worse than those of other batches from the production process. As a result, the true batch shelf life is a random quantity that varies from batch to batch. The variation among the true batch shelf lives defines a distribution as illustrated in Fig. 2. In general, this distribution will be right-skewed (10).

Similarly, each individual single-dosage unit, such as a tablet, has its own true tablet shelf life. A bottle containing 100 tablets has a true bottle shelf life. As each tablet (or bottle) is a unique realization of the pharmaceutical product's manufacturing process, the true tablet shelf life will vary from tablet to tablet and the true bottle shelf life will vary from bottle to bottle. This discussion can go on by naming various other packaging types, each having its own unique true shelf life that will vary from one package to another.

Conceptually, the entire production of pharmaceutical product consisting of a number of already manufactured batches, as well as an unknown number of future batches, is characterized by the true product shelf life. It is this entire production that the Working Group believes best defines what is meant by “drug product” in the ICH Q1A definition of shelf life. The true product shelf life is never known but can be estimated. If the estimation method is unbiased and precise, the estimate should be close to the true value and collecting more data would further improve that estimate. While the definition of drug product should explicitly acknowledge the immediate container closure system, this is not critical for the purpose of this paper since the concepts presented here do not depend on what container closure system is used.

As there is a hierarchy of units (e.g., batches, bottles and tablets) related to the true shelf life, the level of focus should

be defined. ICH Q1A refers to the shelf life of the drug product (without explicitly defining “product”). In ICH Q1E, the analysis is focused on individual stability batches and is based on regression analysis (estimating the true batch intercept and slope over time). Other bases for determining the appropriate unit to consider include the actual manufacturing process, release testing and stability assessment. As the unit for release is a batch, and it is individual batches (not individual tablets or bottles) that are studied over time, batch shelf life is the lowest level of hierarchy that should be evaluated.

STATISTICAL METHODOLOGY

ICH Guidelines

For quantitative stability limiting attributes, ICH Q1E suggests using linear or nonlinear regression and statistical modeling through “poolability” tests for determining the estimated shelf life of a drug product. To do this, test results from at least three stability registration batches are obtained at predetermined storage times. For a simple linear regression model, the analysis follows in a stepwise fashion to determine which of the following alternative regression models is most appropriate for characterizing the response of the batches over storage time and estimating the shelf life: (a) common intercept and common slope, (b) separate intercepts and common slope, or (c) separate intercepts and separate slopes. In practice, a simple linear regression model with common intercept and differing slopes among batches is also considered.

The estimation methodology suggested in the ICH guidelines follows a fixed-batch approach to estimate shelf life. A fixed-batch estimation philosophy assumes that the batches

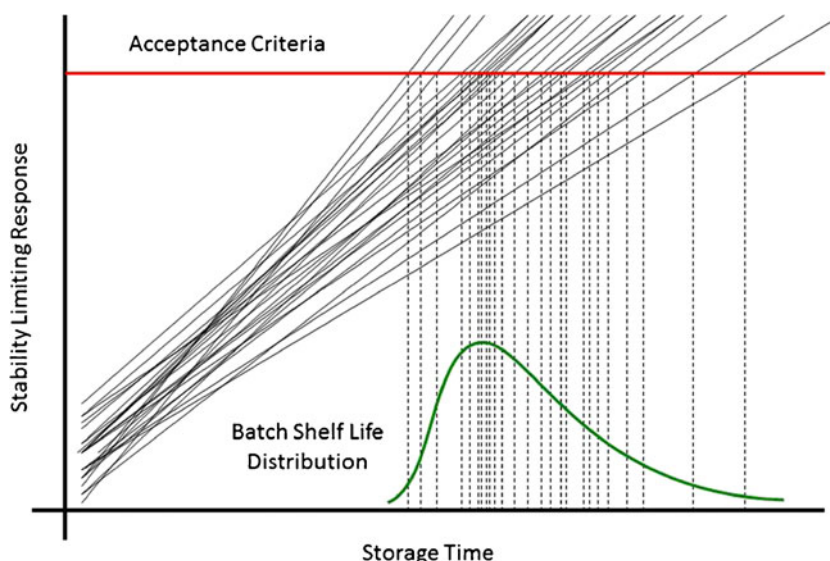


Fig. 2. Distribution of true batch shelf lives. The *Y-axis* represents the value of the mean response of the stability limiting characteristic and the *X-axis* represents storage time in months. Each *solid line* represents the true regression of mean batch response over storage time for a particular stability limiting response variable. True batch shelf life is defined as the calibrated (*dashed line*) storage time associated with the point where the true regression line for any batch intersects the acceptance criterion. Each individual batch is associated with a unique true batch shelf life, subject to the random variation among batches. Taken together, the set of individual true batch shelf lives define a distribution of batch shelf lives in terms of storage months

used in the stability study are entirely representative of the pharmaceutical product's distribution in terms of the product's manufacturing process. Following a fixed-batch estimation philosophy permits only the estimation of within-batch variation in response; among-batch variation cannot be estimated. The among-batch variation measures the amount of variation observed as batch-to-batch differences. The within-batch variation measures the difference in the response data within each batch. In a fixed-batch analysis, these two sources of variation are combined together and used as a single measure of overall variation.

The basic assumption of a stability study is that the registration batches included in the stability study arise from a manufacturing process that is in a state of control. A well-controlled process is defined by an overall mean response where individual batch responses represent a realization of that process mean subject to random variation. Arguably, then, a regression model utilizing a common intercept and common slope is the most meaningful model to describe the mean response of a process in control. Estimating the common intercept and slope should then utilize all available data from the stability study to provide the most precise estimate of the overall mean response over time. But increasing the precision of this estimate leads to the consequence that even if the regression model with a common intercept and common slope among batches is the most appropriate model, the statistical calculations will tend to select it least frequently if the ICH strategy is followed because increasing the precision of the estimates of intercept and slope increases the power of the ICH tests to detect small, even inconsequential, differences among them. This is problematic since a justification to pool the batch stability data is a result of not rejecting the null hypothesis that a common intercept and slope model is

appropriate. Increasing the number of batches to include in a stability trial only exacerbates the problem. The ramification of this is that by following the ICH guidelines, there is no incentive for the manufacturer to include additional measurements or batches in the stability study because, most often, one of the regression models allowing for differing intercepts and/or slopes among batches is selected. For these regression models, the estimated product shelf life is dictated by the ICH guidelines as the shortest of the estimates of shelf life from the individual batches. These methods result in an estimate of product shelf life that is based primarily on data from the "worst-case" batch. This is counterintuitive to fundamental statistical philosophy and principles, where an increase in sample size should ensure a more precise estimate of the true product shelf life, not essentially guarantee an estimate that is biased towards a shorter storage time.

The intent of the ICH Q1E strategy is to establish the storage time during which the critical attribute(s) will be considered acceptable for all "future batches manufactured, packaged, and stored under similar circumstances." Unfortunately, the statistical methodology recommended in this guidance document is incompatible with this intent. Regardless of what regression model is used, the shelf life being estimated by the ICH methods only applies to those batches used in the stability study because the variation among batches cannot be estimated. In other words, no legitimate inference can be made by statistical methods to future batches of the pharmaceutical product because the among-batch variance component is not estimated. To infer to the entire pharmaceutical product requires information about batch-to-batch performance which requires a measure of the variation among batches. If the ICH Q1E methodology is extended so that the batches used in stability studies are treated as a random

sample of batches taken from the entire production of the pharmaceutical product, there is still a problem with estimating the true product shelf life as the minimum of the estimated batch shelf lives. Since the distribution of true batch shelf lives is considered to be continuous and nonnegative, the minimum of the estimated batch shelf lives will tend to reflect an ever shorter estimate of the true product shelf life.

Another possibility is to define the true product shelf life as the (central) quantile of the distribution of true batch shelf lives corresponding to the mean of the product distribution. However, as depicted in Fig. 3, this causes a serious limitation in the interpretation of true product shelf life—by the time the true product shelf life is reached, half of the product will by definition have already exceeded the acceptance criteria. This is contradictory to the intent of the definition of shelf life.

An Alternative Strategy to ICH Guidance

The Working Group proposes an alternative estimation strategy of the true product shelf life that is more consistent with the philosophy and intent of the ICH definition. This alternative proposal uses the same definition of true batch shelf life as used in the ICH approach for determining the labeled shelf life but assumes that the batches included in the stability study are a random sample of batches from the production process. An analysis strategy assuming random batch effects has a number of important benefits because it:

- allows estimation of among-batch variation separately from the within-batch variation;
- provides the information needed for making inferences to future batches of the pharmaceutical product;
- avoids the “poolability” testing issue faced by a fixed-batch analysis;

- avoids the problem of estimating the true product shelf life based on data from the “worse-case” batch;
- eliminates the counterintuitive notion that including additional batches in the stability study increases the likelihood of obtaining a shorter estimate of true product shelf life; and
- provides the manufacturer with the incentive to include additional batches in the stability trial to obtain a better estimate of the true product shelf life.

The random-batch model uses the same samples and measurements as the ICH approach, and the true batch shelf life is estimated using estimates of the intercept and slope of each batch (and the uncertainty of those estimates) (11).

Understanding the random variation in the true batch shelf lives is crucial to understanding true product shelf life. With ICH Q1A in mind, the critical question is: what is the implementable interpretation of “drug product batches are expected to remain within specification”? It is clear that this question refers to some characteristic of the product distribution and how that relates to a corresponding characteristic of the distribution of true batch shelf lives. In other words, the true product shelf life is defined to be the characteristic of the distribution of true batch shelf lives giving operational meaning to Q1A. Because of the concerns associated with using either the minimum or a central quantile of the distribution of batch shelf lives to define the true product shelf life, the Working Group proposes to utilize the lower p th quantile of this distribution. The choice of quantile has to be carefully considered because it translates to an upper bound on the probability that a randomly selected batch will be nonconforming when tested at expiry. For example, if we choose $p=5$, then there is a 5% chance that a given batch's shelf life is less than the true product shelf life and the probability will be less than 5% that a randomly selected batch will be

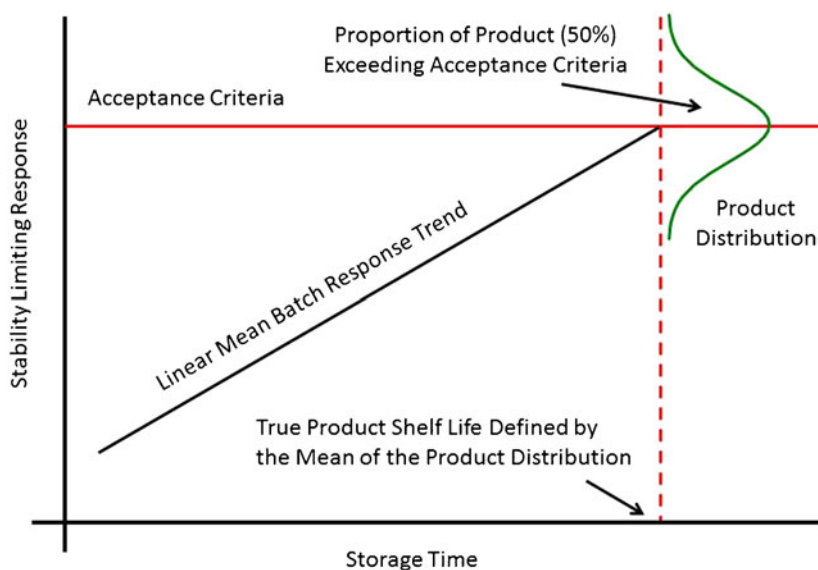


Fig. 3. Consequences of relating the true product shelf life to the mean of the product distribution. The *Y-axis* represents the value of the mean response of the stability limiting characteristic and the *X-axis* represents storage time in months. The product distribution is the theoretical distribution of mean responses for all batches in the production process obtained at a particular storage time. The shelf life corresponding to the mean of the product distribution leads to a quality statement of the form “no more than 50% of all batches are nonconforming at expiry”

nonconforming when tested at its labeled shelf life. A Quality Statement relating to the proportion of nonconforming batches deemed acceptable at expiry is both mathematically tractable and captures the intent behind the recommendations provided in the ICH guidance documents. The Working Group's proposed definition of shelf life is illustrated in Fig. 4.

Estimating Batch and Product Shelf Life

If stability studies consist of batches understood to be a random sample from the population of batches, as described above, then the resulting data can be used to estimate the true shelf lives of the batches in the studies. However, the ultimate objective of a stability trial is to assemble this information into an estimate of the true product shelf life and based on these data to select and justify a suitably labeled shelf life. A production batch is assumed to arise from a manufacturing process that is in a state of statistical control at the time the batch is made and is expected to stay in such a state. Thus, the population of batches has two characteristics relevant to this discussion (see Fig. 2):

- There is random variation among the batches in terms of initial levels and trends over time (intercepts and slopes for a linear model).
- As a consequence of this random variation, true batch shelf life varies among batches

The true product shelf life and the distribution of true batch shelf lives are both unknown. Their characteristics must be estimated from statistical analysis of stability trials. ICH guidance recommends linear regression analysis, often as part of an overall analysis of covariance, but allows for polynomial

or nonlinear regression models to describe the temporal nature of the stability limiting characteristic. While the recommended analyses assume a fixed-batch model, they can be readily amended to handle the random-batch model, which is necessary if inference is to be made to the entire production process. The focus of these analyses is on the mean response. Alternatively, quantile regression or mixed model tolerance interval methods are particularly attractive due to their versatility (12). These approaches analyze the observed data from the sampled batches and model a quantile of the distribution of true batch shelf lives [Quinlan, M.; Stroup, W.; Christopher, D.; Schwenke, J. On Estimating Shelf Life Using Mixed Model Quantile Regression (in review). 2011]. For example, instead of focusing on the minimum of the distribution, which is mathematically intractable, the focus could instead be on some lower quantile of the distribution (e.g., 1st or 5th quantile). Quantile regression or tolerance intervals are natural approaches to address the intent of the ICH guidelines. One inherent consequence of these alternative approaches is they may require that more than the customary three batches be placed on stability.

CONCLUSIONS

To enable research and discussion of methods for establishing shelf life of pharmaceutical products, a consensus agreement is needed on the formal definition of true product shelf life, the relationship between overall product and individual batch response and the construction of a Quality Statement to reflect the desired properties of a shelf life estimate for a pharmaceutical product. On this basis, an appropriate set of statistical tools that verify conformance to a Quality

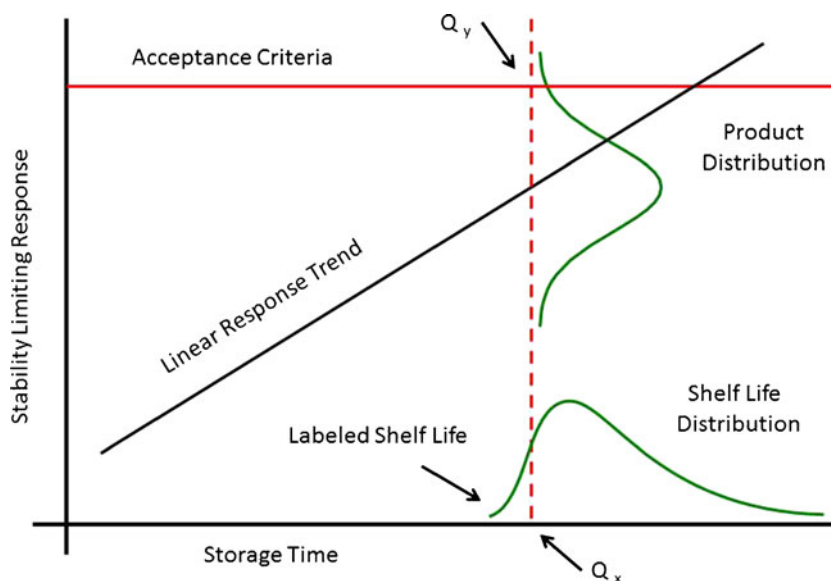


Fig. 4. The Working Group's proposed definition of shelf life based on an acceptably small proportion of product exceeding an acceptance criterion. The linear response trend displayed is the overall mean response for the stability limiting characteristic for all batches in the production process. Q_x is a quantile of the distribution of batch shelf lives. The labeled shelf life can be defined as a lower bound corresponding to Q_x , which could be estimated through quantile regression or as a lower limit of a tolerance interval. The product distribution is defined by both the between-batch and within-batch variation. Q_y is the quantile of the product distribution associated with an acceptable proportion of product exceeding the acceptance criteria at the labeled shelf life

Statement can be designed and compared to select the most suitable estimation methods. Because of the limitations associated with using the minimum or central quantile of the distribution of true batch shelf lives to define the true product shelf life, the Working Group proposes that the true product shelf life should be defined in terms of a suitably small quantile of this distribution. Gaining information about both the product and shelf life distributions and the relationship between the two distributions, through replicate or historical batch response, then allows determining the proportion of the shelf life distribution to be considered for defining the estimated product shelf life.

ACKNOWLEDGMENTS

The Working Group is grateful to PQRI for supporting this project, to Suntara Cahya and Paula Hudson (Eli Lilly), and David Thomas (Johnson & Johnson) who participated in the early stages of this work and to Abhay Gupta (FDA) for his contributions to the manuscript development and discussions as an active PQRI Stability Shelf Life Working Group member.

REFERENCES

1. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Q1A(R2): Stability testing of new drug substances and products; 2003.
2. PQRI Stability Shelf Life Working Group. http://www.pqri.org/commworking/minutes/pdfs/dptc/sslwg/Addl/2007_MBSW.pdf. Additional presentations from SSL WG are available at <http://www.pqri.org/structure/wg.asp#sslwg>, 2007; Accessed 27 March 2012.
3. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Q6A: specifications: test procedures and acceptance criteria for new drug substances and new drug products: Chemical Substances; 1999.
4. International Organization for Standardization. ISO 2859. Sampling procedures for inspection by attributes, parts 0–4. ISO 2859-0:1995; ISO 2859-1:1999; ISO 2859-2:1985; ISO 2859-3:1991; ISO 2859-4:2002. http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=7865. Accessed 18 June 2012.
5. International Organization for Standardization. ISO 3951:1989. Sampling procedures and charts for inspection by variables for percent nonconforming. http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=9602. Accessed 18 June 2012.
6. US Department of Defense. MIL-STD-690D Failure rate sampling plans and procedures. [http://www.variation.com/anonftp/pub/MIL-STD-690D%20\(10%20June%202005\).pdf](http://www.variation.com/anonftp/pub/MIL-STD-690D%20(10%20June%202005).pdf). Accessed 15 June 2011; 2005.
7. American National Standards Institute. ANSI/ASQC Z1.4-2008 Sampling procedures and tables for inspection by attributes; 2008. <http://webstore.ansi.org/RecordDetail.aspx?sku=ANSI%2FASQ+Z1.4-2008>. Accessed 18 June 2012.
8. Larner G, Cooper A, Lyapustina S, Leiner S, Christopher D, Strickland H, *et al.* Challenges and opportunities in implementing the FDA default parametric tolerance interval two one-sided test for delivered dose uniformity of orally inhaled products. *AAPS PharmSciTech*. 2011;12(4):1144–56.
9. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Q1E evaluation of stability data; 2004.
10. Quinlan M, Stroup W, Christopher D, Schwenke J. On the distribution of batch shelf lives. (Accepted for publication in *J Biopharm Stat*. 2011).
11. Quinlan M, Stroup W, Schwenke J, Christopher D. Evaluating the performance of the ICH guidelines for shelf life estimation. (Accepted for publication in *J Biopharm Stat*. 2011).
12. Stroup W, Quinlan, M. Alternative shelf life estimation methodologies. In *JSM Proceedings*, Biopharmaceutical Section. Alexandria, VA: American Statistical Association. 2010;2056–x2066.