EXHIBIT D

IN THE UNITED STATES DISTRICT COURT
FOR THE NORTHERN DISTRICT OF ILLINOIS
EASTERN DIVISION

HOSPIRA, INC., Plaintiff, v. FRESENIUS KABI USA, LLC, Defendant.	Civil Action Nos. 1:16-cv-00651 1:17-cv-07903 Hon. Judge Rebecca R. Pallmeyer

DECLARATION OF DR. STEPHAN OGENSTAD IN OPPOSITION TO DEFENDANT'S MOTION *IN LIMINE*

I, Stephan Ogenstad, Ph.D., do hereby declare as follows:

1. I have a Ph.D. in statistics and have been retained by Plaintiff Hospira, Inc. ("Hospira") in the above actions. I submit this Declaration in support of Hospira's Opposition to Defendant's Motion *In Limine*.

2. I received my B.Sc. in Mathematics, Statistics, and Computing from the University of Stockholm in 1974. I received my Ph.D. in Statistics from the University of Stockholm in 1982. From 1975-1982, I was Chief Statistical Advisor and reviewer to the Nobel Prize Committee for Medicine and Physiology.

3. I have more than 40 years' experience working as a biostatistician in the pharmaceutical industry and academia. For a decade, I was a statistician at Vertex Pharmaceuticals where I was responsible for management of clinical data managers, biostatisticians, and statistical programmers in the pre-clinical, non-clinical, and clinical areas. Before Vertex, I held management positions in biostatistical divisions of several

biopharmaceutical companies, including Amgen and AstraZeneca. In my roles at these companies, I frequently worked with the FDA in connection with regulatory IND applications and submissions (mainly NDAs), including INDs and NDAs for local anesthetics and analgesics, diabetes, central nervous system, hiv/aids, oncology, infectious diseases, cardiology, urology, rheumatoid arthritis, neurology, dentistry, and psoriasis. I also developed study protocols and statistical analysis plans, and conducted statistical analysis for NDA submissions.

4. Since 2006, I have been President of Statogen Consulting LLC, where I consult with companies in the biopharmaceutical and medical device industries that are seeking to bring new products to market. I provide statistical analysis throughout the entire product development process, including with respect to clinical studies and the development of dossiers requesting regulatory approval to commercially market drug products.

5. For the last ten years, I have also been an Adjunct Faculty Member and Professor of Biostatistics at Georgia Southern University. Earlier in my career, I was Lecturer and Professor of Statistics at the University of Stockholm, lecturer at the Swedish Academy of Pharmaceutical Sciences, and the Karolinska Institute. I have taught courses on, among other things, probability theory, Analysis of Variance, and regression analysis.

6. I have decades of experience with the evaluation of stability data.

7. The sampling methodology I used in my report, relying on a random number generator to simulate data, recommended in FDA's guideline (Q1E), is very valuable to demonstrate whether or not the statistical properties of the analysis methods are appropriate, is widely-accepted and reliable. (Exh. A.) The statistical methods used are classical statistical

methods developed over more than 100 years by the founders of statistics such as Karl Pearson, Ronald Fisher, Jerzy Neyman, Harald Cramer, and David Cox, just to mention a few.

8. For example, the following articles and book chapters, all published in refereed journals, use this very technique.

a. Bakshi, M. and Singh, S. Development of validated stability-indicating assay methods—critical review. Volume 28, Issue 6, 2002, pp. 1011-1040. (Exh. B.) At page 1035, recommends using simulation to generate data to evaluate wehther stability assays for pharmaceutical products will meet regulatory requirements.

b. Magari, R.T. Assessing Shelf Life Using Real-Time and Accelerated Stability Tests. BioPharm International, Volume 16, Issue 11, 2003. (Exh. C.) Uses simulated data on drug degradation to study whether accelerated stability testing can substitute for real-time proof.

c. Wongpoowarak, W. et al. Computer Simulation for Studying Complexation between a Model Drug and a Model Protein. (Exh. D.) Uses simulation data to illustrate problems with data sets that contain variable data.

d. Kleinman et al. In silico prediction of pharmaceutical degradation pathways: a benchmarking study. Mol Pharm. 2014 Nov 3;11(11):4179-88. (Exh. E.) An article evaluating predictions regarding drug degradation products.

e. Torres et al. The application of electrochemistry to pharmaceutical stability testing--comparison with in silico prediction and chemical forced degradation approaches. J Pharm Biomed Anal. 2015 Nov 10;115:487-501. (Exh. F.) Also uses predications to regarding degradation products to evaluate stability. '*In silico*' meaning via computer simulation.

9. I used prediction intervals in my expert report. A prediction interval is similar to a confidence interval and is the interval in which there is a 95% probability (for a 95% prediction interval) the next measurement will occur.

10. Prediction intervals play no role in the simulation methodology I used. A prediction interval is calculated from the actual data and is independent of simulated data.

11. I have published frequently using simulation data in the pharmaceutical area. For example, the following articles all use simulation data to evaluate aspects of pharmaceutical science.

a. Ette, E.I., Godfrey, C.J., Ogenstad, S., Williams, P., Analysis of Simulated Clinical Trials: In: Simulation for Designing Clinical Trials: A Pharmacokinetic-Pharmacodynamic Modeling Perspective. IBSN: 0-8247- 0862-8, Marcel-Dekker 2002. (Ex. G.)

b. Westfall P.H., Tsai K., Ogenstad S., Tomoiaga A., Moseley S., and Lu Y. Clinical Trials Simulation: A Statistical Approach. Journal of Biopharmaceutical Statistics 18, 611–630, 2008. (Ex. H.)

c. Ogenstad, S. A Statistical Approach to Clinical Trial Simulations. In: Biopharmaceutical Applied Statistics Symposium. Springer. 2018.

I declare under penalty of perjury that the foregoing is true and correct to the best of my knowledge and belief.

Executed on June 22, 2018

/s/ Stephan Ogenstad
Stephan Ogenstad

Ogenstad Declaration Exhibit A

Guidance for Industry

Q1E Evaluation of Stability Data

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

June 2004
ICH

Guidance for Industry

Q1E Evaluation of Stability Data

Additional copies are available from:

Office of Training and Communication
Division of Drug Information, HFD-240
Center for Drug Evaluation and Research
Food and Drug Administration
5600 Fishers Lane
Rockville, MD 20857
(Tel) 301-827-4573
http://www.fda.gov/cder/guidance/index.htm

Office of Communication, Training, and
Manufacturers Assistance, HFM-40
Center for Biologics Evaluation and Research
Food and Drug Administration
1401 Rockville Pike, Rockville, MD 20852-1448
http://www.fda.gov/cber/guidelines.htm.
(Tel) Voice Information System at 800-835-4709 or 301-827-1800

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

June 2004
ICH

TABLE OF CONTENTS

Contains Nonbinding Recommendations

Guidance for Industry[1]
Q1E Evaluation of Stability Data

This guidance represents the Food and Drug Administration's (FDA's) current thinking on this topic. It does not create or confer any rights for or on any person and does not operate to bind FDA or the public. An alternative approach may be used if such approach satisfies the requirements of the applicable statutes and regulations. If you want to discuss an alternative approach, contact the FDA staff responsible for implementing this guidance. If you cannot identify the appropriate FDA staff, call the appropriate number listed on the title page of this guidance.

I. INTRODUCTION (1.0)[2]

This guidance provides recommendations on how to use stability data generated in accordance with the principles detailed in the ICH guidance Q1A(R2) *Stability Testing of New Drug Substances and Products* (parent guidance) to propose a retest period or shelf life in a registration application. This guidance describes when and how extrapolation can be considered when proposing a retest period for a drug substance or a shelf life for a drug product that extends beyond the period covered by available data from the stability study under the long-term storage condition (long-term data).

FDA's guidance documents, including this guidance, do not establish legally enforceable responsibilities. Instead, guidances describe the Agency's current thinking on a topic and should be viewed only as recommendations, unless specific regulatory or statutory requirements are cited. The use of the word *should* in Agency guidances means that something is suggested or recommended, but not required.

The recommendations in the evaluation and statistical analysis of stability data provided in the parent guidance are brief in nature and limited in scope. The parent guidance states that regression analysis is an appropriate approach to analyzing quantitative stability data for retest

[1] This guidance was developed within the Expert Working Group (Quality) of the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) and has been subject to consultation by the regulatory parties, in accordance with the ICH process. This document has been endorsed by the ICH Steering Committee at *Step 4* of the ICH process, February 2003. At *Step 4* of the process, the final draft is recommended for adoption to the regulatory bodies of the European Union, Japan, and the United States.

[2] Arabic numbers reflect the organizational breakdown in the document endorsed by the ICH Steering Committee at Step 4 of the ICH process, February 2003.

period or shelf life estimation and recommends that a statistical test for batch poolability be performed using a level of significance of 0.25. However, the parent guidance includes few details and does not cover situations where multiple factors are involved in a full- or reduced-design study. This guidance expands the recommendations presented in the evaluation sections of the parent guidance.

This guidance covers:

- The evaluation of stability data that should be submitted in registration applications for new molecular entities and associated drug products.
- Recommendations on the establishment of retest periods and shelf lives for drug substances and drug products intended for storage at or below *room temperature*.*
- Stability studies using single- or multi-factor designs and full or reduced designs.

***Note:** The term *room temperature* refers to the general customary environment and should not be inferred to be the storage statement for labeling.

ICH Q6A and Q6B should be consulted for recommendations on the setting and justification of acceptance criteria, and ICH Q1D should be referenced for recommendations on the use of full- versus reduced-design studies.

II. EVALUATION OF STABILITY DATA (2.0)

A. General Principles (2.1)

The design and execution of formal stability studies should follow the principles outlined in the parent guidance. The purpose of a stability study is to establish, based on testing a minimum of three batches of the drug substance or product, a retest period or shelf life and label storage instructions applicable to all future batches manufactured and packaged under similar circumstances. The degree of variability of individual batches affects the confidence that a future production batch will remain within acceptance criteria throughout its retest period or shelf life.

Although normal manufacturing and analytical variations are to be expected, it is important that the drug product be formulated with the intent to provide 100 percent of the labeled amount of the drug substance at the time of batch release. If the assay values of the batches used to support the registration application are higher than 100 percent of label claim at the time of batch release, after taking into account manufacturing and analytical variations, the shelf life proposed in the application can be overestimated. On the other hand, if the assay value of a batch is lower than 100 percent of label claim at the time of batch release, it might fall below the lower acceptance criterion before the end of the proposed shelf life.

A systematic approach should be adopted in the presentation and evaluation of the stability information. The stability information should include, as appropriate, results from the physical, chemical, biological, and microbiological tests, including those related to particular attributes of the dosage form (for example, dissolution rate for solid oral dosage forms). The adequacy of the

mass balance should be assessed. Factors that can cause an apparent lack of mass balance should be considered, including, for example, the mechanisms of degradation and the stability-indicating capability and inherent variability of the analytical procedures.

The basic concepts of stability data evaluation are the same for single- versus multi-factor studies and for full- versus reduced-design studies. Data from formal stability studies and, as appropriate, supporting data should be evaluated to determine the critical quality attributes likely to influence the quality and performance of the drug substance or product. Each attribute should be assessed separately, and an overall assessment should be made of the findings for the purpose of proposing a retest period or shelf life. The retest period or shelf life proposed should not exceed that predicted for any single attribute.

The decision tree in Appendix A outlines a stepwise approach to stability data evaluation and when and how much extrapolation can be considered for a proposed retest period or shelf life. Appendix B provides (1) information on how to analyze long-term data for appropriate quantitative test attributes from a study with a multi-factor, full, or reduced design, (2) information on how to use regression analysis for retest period or shelf life estimation, and (3) examples of statistical procedures to determine poolability of data from different batches or other factors. Additional guidance can be found in the references listed; however, the examples and references do not cover all applicable statistical approaches.

In general, certain quantitative chemical attributes (e.g., assay, degradation products, preservative content) for a drug substance or product can be assumed to follow zero-order kinetics during long-term storage (Carstensen 1977). Data for these attributes are therefore amenable to the type of statistical analysis described in Appendix B, including linear regression and poolability testing. Although the kinetics of other quantitative attributes (e.g., pH, dissolution) is generally not known, the same statistical analysis can be applied, if appropriate. Qualitative attributes and microbiological attributes are not amenable to this kind of statistical analysis.

The recommendations on statistical approaches in this guidance are not intended to imply that use of statistical evaluation is preferred when it can be justified to be unnecessary. However, statistical analysis can be useful in supporting the extrapolation of retest periods or shelf lives in certain situations and can be critical in verifying the proposed retest periods or shelf lives in other cases.

B. Data Presentation (2.2)

Data for all attributes should be presented in an appropriate format (e.g., tabular, graphical, narrative) and an evaluation of such data should be included in the application. The values of quantitative attributes at all time points should be reported as measured (e.g., assay as percent of label claim). If a statistical analysis is performed, the procedure used and the assumptions underlying the model should be stated and justified. A tabulated summary of the outcome of statistical analysis and/or graphical presentation of the long-term data should be included.

C. Extrapolation (2.3)

Extrapolation is the practice of using a known data set to infer information about future data. Extrapolation to extend the retest period or shelf life beyond the period covered by long-term data can be proposed in the application, particularly if no significant change is observed at the accelerated condition. Whether extrapolation of stability data is appropriate depends on the extent of knowledge about the change pattern, the goodness of fit of any mathematical model, and the existence of relevant supporting data. Any extrapolation should be performed in such a way that the extended retest period or shelf life will be valid for a future batch released with test results close to the release acceptance criteria.

An extrapolation of stability data assumes that the same change pattern will continue to apply beyond the period covered by long-term data. The correctness of the assumed change pattern is critical when extrapolation is considered. When estimating a regression line or curve to fit the long-term data, the data themselves provide a check on the correctness of the assumed change pattern, and statistical methods can be applied to test the goodness of fit of the data to the assumed line or curve. No such internal check is possible beyond the period covered by long-term data. Thus, a retest period or shelf life granted on the basis of extrapolation should always be verified by additional long-term stability data as soon as these data become available. Care should be taken to include in the protocol for commitment batches a time point that corresponds to the end of the extrapolated retest period or shelf life.

D. Data Evaluation for Retest Period or Shelf Life Estimation for Drug Substances or Products Intended for Room Temperature Storage (2.4)

A systematic evaluation of the data from formal stability studies should be performed as illustrated in this section. Stability data for each attribute should be assessed sequentially. For drug substances or products intended for storage at room temperature, the assessment should begin with any significant change at the accelerated condition and, if appropriate, at the intermediate condition, and progress through the trends and variability of the long-term data. The circumstances are delineated under which extrapolation of retest period or shelf life beyond the period covered by long-term data can be appropriate. A decision tree is provided in Appendix A as an aid.

1. No significant change at accelerated condition (2.4.1)

Where no significant change occurs at the accelerated condition, the retest period or shelf life would depend on the nature of the long-term and accelerated data.

a. Long-term and accelerated data showing little or no change over time and little or no variability (2.4.1.1)

Where the long-term data and accelerated data for an attribute show little or no change over time and little or no variability, it might be apparent that the drug substance or product will remain well within the acceptance criteria for that attribute during the proposed retest period or shelf life. In these circumstances, a statistical analysis is normally considered unnecessary but

justification for the omission should be provided. Justification can include a discussion of the change pattern or lack of change, relevance of the accelerated data, mass balance, and/or other supporting data as described in the parent guidance. Extrapolation of the retest period or shelf life beyond the period covered by long-term data can be proposed. The proposed retest period or shelf life can be up to twice as long as, but should not be more than 12 months beyond, the period covered by long-term data.

b. Long-term or accelerated data showing change over time and/or variability (2.4.1.2)

If the long-term or accelerated data for an attribute show change over time and/or variability within a factor or among factors, statistical analysis of the long-term data can be useful in establishing a retest period or shelf life. Where there are differences in stability observed among batches or among other factors (e.g., strength, container size, and/or fill) or factor combinations (e.g., strength-by-container size and/or fill) that preclude the combining of data, the proposed retest period or shelf life should not exceed the shortest period supported by any batch, other factor, or factor combination. Alternatively, where the differences are readily attributed to a particular factor (e.g., strength), different shelf lives can be assigned to different levels within the factor (e.g., different strengths). A discussion should be provided to address the cause for the differences and the overall significance of such differences on the product. Extrapolation beyond the period covered by long-term data can be proposed; however, the extent of extrapolation would depend on whether long-term data for the attribute are amenable to statistical analysis.

- Data not amenable to statistical analysis

Where long-term data are not amenable to statistical analysis, but relevant supporting data are provided, the proposed retest period or shelf life can be up to one-and-a-half times as long as, but should not be more than 6 months beyond, the period covered by long-term data. Relevant supporting data include satisfactory long-term data from development batches that are (1) made with a closely related formulation to, (2) manufactured on a smaller scale than, or (3) packaged in a container closure system similar to, that of the primary stability batches.

- Data amenable to statistical analysis

If long-term data are amenable to statistical analysis but no analysis is performed, the extent of extrapolation should be the same as when data are not amenable to statistical analysis. However, if a statistical analysis is performed, it can be appropriate to propose a retest period or shelf life of up to twice as long as, but not more than 12 months beyond, the period covered by long-term data, when the proposal is backed by the result of the analysis and relevant supporting data.

2. Significant change at accelerated condition (2.4.2)

Where significant change* occurs at the accelerated condition, the retest period or shelf life will depend on the outcome of stability testing at the intermediate condition, as well as at the long-term condition.

***Note:** The following physical changes can be expected to occur at the accelerated condition and would not be considered significant change that calls for intermediate testing if there is no other significant change:

- Softening of a suppository that is designed to melt at 37ºC, if the melting point is clearly demonstrated.
- Failure to meet acceptance criteria for dissolution for 12 units of a gelatin capsule or gel-coated tablet if the failure can be unequivocally attributed to cross-linking.

However, if phase separation of a semi-solid dosage form occurs at the accelerated condition, testing at the intermediate condition should be performed. Potential interaction effects should also be considered in establishing that there is no other significant change.

a. No significant change at intermediate condition (2.4.2.1)

If there is no significant change at the intermediate condition, extrapolation beyond the period covered by long-term data can be proposed; however, the extent of extrapolation would depend on whether long-term data for the attribute are amenable to statistical analysis.

- Data not amenable to statistical analysis

When the long-term data for an attribute are not amenable to statistical analysis, the proposed retest period or shelf life can be up to 3 months beyond the period covered by long-term data, if backed by relevant supporting data.

- Data amenable to statistical analysis

When the long-term data for an attribute are amenable to statistical analysis but no analysis is performed, the extent of extrapolation should be the same as when data are not amenable to statistical analysis. However, if a statistical analysis is performed, the proposed retest period or shelf life can be up to one-and-half times as long as, but should not be more than 6 months beyond, the period covered by long-term data, when backed by statistical analysis and relevant supporting data.

b. Significant change at intermediate condition (2.4.2.2)

Where significant change occurs at the intermediate condition, the proposed retest period or shelf life should not exceed the period covered by long-term data. In addition, a retest period or shelf life shorter than the period covered by long-term data can be appropriate.

E. Data Evaluation for Retest Period or Shelf Life Estimation for Drug Substances or Products Intended for Storage Below Room Temperature (2.5)

1. Drug substances or products intended for storage in a refrigerator (2.5.1)

Data from drug substances or products intended to be stored in a refrigerator should be assessed according to the same principles as described in section II.D for drug substances or products intended for room temperature storage, except where explicitly noted in the section below. The decision tree in Appendix A can be used as an aid.

a. No significant change at accelerated condition (2.5.1.1)

Where no significant change occurs at the accelerated condition, extrapolation of retest period or shelf life beyond the period covered by long-term data can be proposed based on the principles outlined in section II.D.1, except that the extent of extrapolation should be more limited.

If the long-term and accelerated data show little change over time and little variability, the proposed retest period or shelf life can be up to one-and-a-half times as long as, but should not be more than 6 months beyond, the period covered by long-term data normally without the support of statistical analysis.

Where the long-term or accelerated data show change over time and/or variability, the proposed retest period or shelf life can be up to 3 months beyond the period covered by long-term data if (1) the long-term data are amenable to statistical analysis but a statistical analysis is not performed, or (2) the long-term data are not amenable to statistical analysis but relevant supporting data are provided.

Where the long-term or accelerated data show change over time and/or variability, the proposed retest period or shelf life can be up to one-and-a-half times as long as, but should not be more than 6 months beyond, the period covered by long-term data if (1) the long-term data are amenable to statistical analysis and a statistical analysis is performed, and (2) the proposal is backed by the result of the analysis and relevant supporting data.

b. Significant change at accelerated condition (2.5.1.2)

If significant change occurs between 3 and 6 months' testing at the accelerated storage condition, the proposed retest period or shelf life should be based on the long-term data. Extrapolation is not considered appropriate. In addition, a retest period or shelf life shorter than the period covered by long-term data could be appropriate. If the long-term data show variability, verification of the proposed retest period or shelf life by statistical analysis can be appropriate.

If significant change occurs within the first 3 months' testing at the accelerated storage condition, the proposed retest period or shelf life should be based on long-term data. Extrapolation is not considered appropriate. A retest period or shelf life shorter than the period covered by long-term data could be appropriate. If the long-term data show variability, verification of the proposed retest period or shelf life by statistical analysis can be appropriate. In addition, a discussion should be provided to address the effect of short-term excursions outside the label storage condition (e.g., during shipping or handling). This discussion can be

supported, if appropriate, by further testing on a single batch of the drug substance or product at the accelerated condition for a period shorter than 3 months.

2. Drug substances or products intended for storage in a freezer (2.5.2)

For drug substances or products intended for storage in a freezer, the retest period or shelf life should be based on long-term data. In the absence of an accelerated storage condition for drug substances or products intended to be stored in a freezer, testing on a single batch at an elevated temperature (e.g., 5°C ± 3°C or 25°C ± 2°C) for an appropriate time period should be conducted to address the effect of short-term excursions outside the proposed label storage condition (e.g., during shipping or handling).

3. Drug substances or products intended for storage below -20°C (2.5.3)

For drug substances or products intended for storage below -20°C, the retest period or shelf life should be based on long-term data and should be assessed on a case-by-case basis.

F. General Statistical Approaches (2.6)

Where applicable, an appropriate statistical method should be employed to analyze the long-term primary stability data in an original application. The purpose of this analysis is to establish, with a high degree of confidence, a retest period or shelf life during which a quantitative attribute will remain within acceptance criteria for all future batches manufactured, packaged, and stored under similar circumstances.

In cases where a statistical analysis was employed to evaluate long-term data due to a change over time and/or variability, the same statistical method should also be used to analyze data from commitment batches to verify or extend the originally approved retest period or shelf life.

Regression analysis is considered an appropriate approach to evaluating the stability data for a quantitative attribute and establishing a retest period or shelf life. The nature of the relationship between an attribute and time will determine whether data should be transformed for linear regression analysis. The relationship can be represented by a linear or nonlinear function on an arithmetic or logarithmic scale. In some cases, a nonlinear regression can better reflect the true relationship.

An appropriate approach to retest period or shelf life estimation is to analyze a quantitative attribute (e.g., assay, degradation products) by determining the earliest time at which the 95 percent confidence limit for the mean intersects the proposed acceptance criterion.

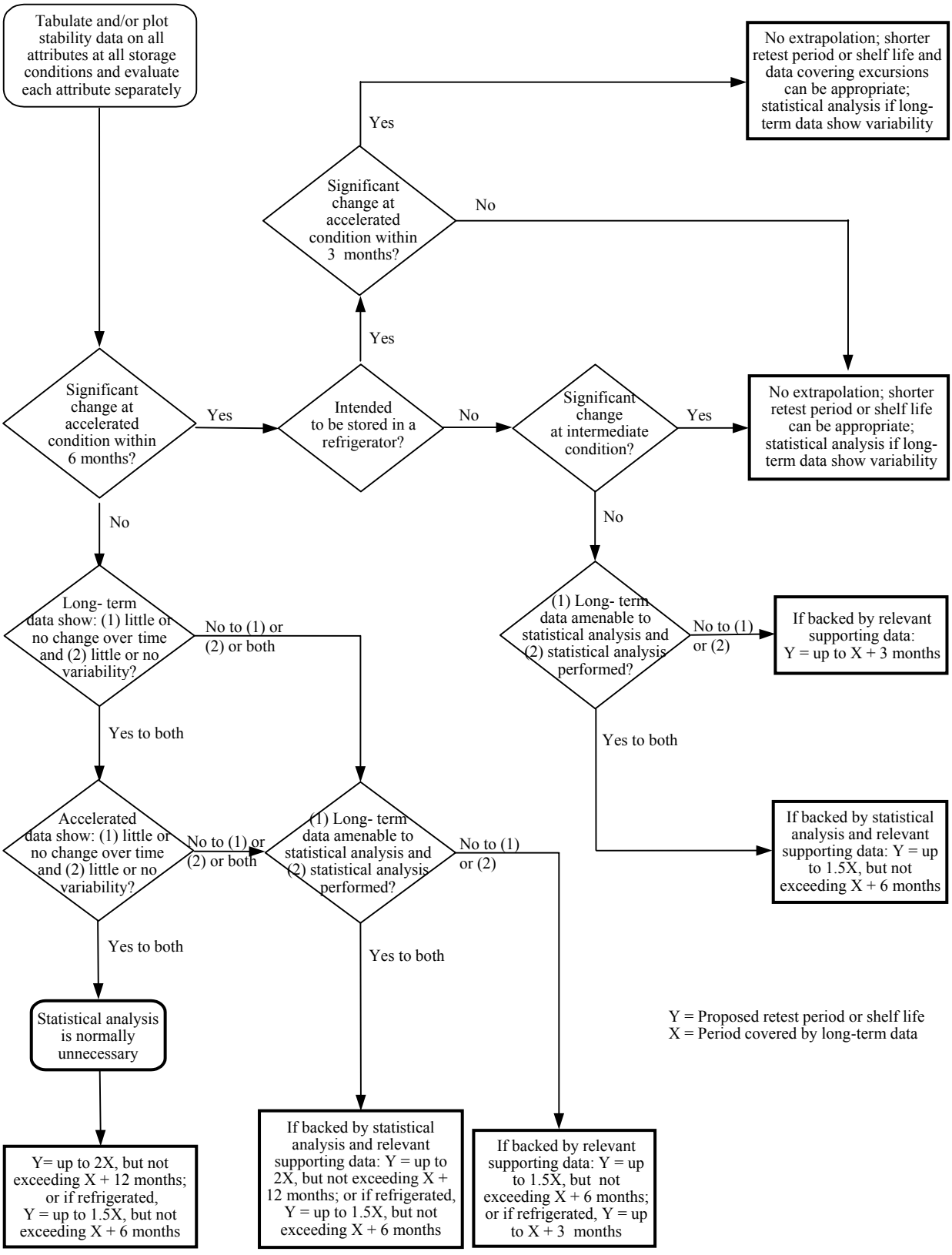
For an attribute known to decrease with time, the lower one-sided 95 percent confidence limit should be compared to the acceptance criterion. For an attribute known to increase with time, the upper one-sided 95 percent confidence limit should be compared to the acceptance criterion. For an attribute that can either increase or decrease, or whose direction of change is not known, two-sided 95 percent confidence limits should be calculated and compared to the upper and lower acceptance criteria.

The statistical method used for data analysis should take into account the stability study design to provide a valid statistical inference for the estimated retest period or shelf life. The approach described above can be used to estimate the retest period or shelf life for a single batch or for multiple batches when the data are combined after an appropriate statistical test. Examples of statistical approaches to the analysis of stability data from single or multi-factor, full- or reduced-design studies are included in Appendix B. References to current literature sources can be found in Appendix B.6.

APPENDICES (3)

Contains Nonbinding Recommendations

Appendix A: Decision Tree for Data Evaluation for Retest Period or Shelf Life Estimation for Drug Substances or Products (Excluding Frozen Products)



Appendix B: Examples of Statistical Approaches to Stability Data Analysis

Linear regression, poolability tests, and statistical modeling, described below, are examples of statistical methods and procedures that can be used in the analysis of stability data that are amenable to statistical analysis for a quantitative attribute for which there is a proposed acceptance criterion.

B.1 Data Analysis for a Single Batch

In general, the relationship between certain quantitative attributes and time is assumed to be linear (Carstensen 1977). Figure 1 (page 18) shows the regression line for assay of a drug product with upper and lower acceptance criteria of 105 percent and 95 percent of label claim, respectively, with 12 months of long-term data and a proposed shelf life of 24 months. In this example, two-sided 95 percent confidence limits for the mean are applied because it is not known ahead of time whether the assay would increase or decrease with time (e.g., in the case of an aqueous-based product packaged in a semi-permeable container). The lower confidence limit intersects the lower acceptance criterion at 30 months, while the upper confidence limit does not intersect with the upper acceptance criterion until later. Therefore, the proposed shelf life of 24 months can be supported by the statistical analysis of the assay, provided the recommendations in sections II.D and II.E are followed.

When data for an attribute with only an upper or a lower acceptance criterion are analyzed, the corresponding one-sided 95 percent confidence limit for the mean is recommended. Figure 2 (page 18) shows the regression line for a degradation product in a drug product with 12 months of long-term data and a proposed shelf life of 24 months, where the acceptance criterion is not more than 1.4 percent. The upper one-sided 95 percent confidence limit for the mean intersects the acceptance criterion at 31 months. Therefore, the proposed shelf life of 24 months can be supported by statistical analysis of the degradation product data, provided the recommendations in sections II.D and II.E are followed.

If the above approach is used, the mean value of the quantitative attribute (e.g., assay, degradation products) can be expected to remain within the acceptance criteria through the end of the retest period or shelf life at a confidence level of 95 percent.

The approach described above can be used to estimate the retest period or shelf life for a single batch, individual batches, or multiple batches when combined after appropriate statistical tests described in Appendix sections B.2 through B.5.

B.2 Data Analysis for One-Factor, Full-Design Studies

For a drug substance or for a drug product available in a single strength and a single container size and/or fill, the retest period or shelf life is generally estimated based on the stability data from a minimum of three batches. When analyzing data from such one-factor, batch-only, full-design studies, two statistical approaches can be considered.

- The objective of the first approach is to determine whether the data from all batches support the proposed retest period or shelf life.
- The objective of the second approach, testing for poolability, is to determine whether the data from different batches can be combined for an overall estimate of a single retest period or shelf life.

B.2.1 Evaluating whether all batches support the proposed retest period or shelf life

The objective of this approach is to evaluate whether the estimated retest periods or shelf lives from all batches are longer than the one proposed. Retest periods or shelf lives for individual batches should first be estimated using the procedure described in Appendix section B.1 with individual intercepts, individual slopes, and the pooled mean square error calculated from all batches. If each batch has an estimated retest period or shelf life longer than that proposed, the proposed retest period or shelf life will generally be considered appropriate, as long as the guidance for extrapolation in sections II.D and II.E is followed. There is generally no need to perform poolability tests or identify the most reduced model. If, however, one or more of the estimated retest periods or shelf lives are shorter than that proposed, poolability tests can be performed to determine whether the batches can be combined to estimate a longer retest period or shelf life.

Alternatively, the above approach can be taken during the pooling process described in Appendix section B.2.2. If the regression lines for the batches are found to have a common slope and the estimated retest periods or shelf lives based on the common slope and individual intercepts are all longer than the proposed retest period or shelf life, there is generally no need to continue to test the intercepts for poolability.

B.2.2 Testing for poolability of batches

B.2.2.1 Analysis of covariance

Before pooling the data from several batches to estimate a retest period or shelf life, a preliminary statistical test should be performed to determine whether the regression lines from different batches have a common slope and a common time-zero intercept. Analysis of covariance (ANCOVA) can be employed, where time is considered the covariate, to test the differences in slopes and intercepts of the regression lines among batches. Each of these tests should be conducted using a significance level of 0.25 to compensate for the expected low power of the design due to the relatively limited sample size in a typical formal stability study.

If the test rejects the hypothesis of equality of slopes (i.e., if there is a significant difference in slopes among batches), it is not considered appropriate to combine the data from all batches. The retest periods or shelf lives for individual batches in the stability study can be estimated by applying the approach described in Appendix section B.1 using individual intercepts and individual slopes and the pooled mean square error calculated from all batches. The shortest estimate among the batches should be chosen as the retest period or shelf life for all batches.

If the test rejects the hypothesis of equality of intercepts but fails to reject that the slopes are equal (i.e., if there is a significant difference in intercepts but no significant difference in slopes among the batches), the data can be combined for the purpose of estimating the common slope. The retest periods or shelf lives for individual batches in the stability study should be estimated by applying the approach described in Appendix section B.1, using the common slope and individual intercepts. The shortest estimate among the batches should be chosen as the retest period or shelf life for all batches.

If the tests for equality of slopes and equality of intercepts do not result in rejection at a level of significance of 0.25 (i.e., if there is no significant difference in slope and intercepts among the batches), the data from all batches can be combined. A single retest period or shelf life can be estimated from the combined data by using the approach described in Appendix section B.1 and applied to all batches. The estimated retest period or shelf life from the combined data is usually longer than that from individual batches because the width of the confidence limit(s) for the mean will become narrower as the amount of data increases when batches are combined.

The pooling tests described above should be performed in a proper order such that the slope terms are tested before the intercept terms. The most reduced model (i.e., individual slopes, common slope with individual intercepts, or common slope with common intercept, as appropriate) can be selected for retest period or shelf life estimation.

B.2.2.2 Other methods

Statistical procedures (Ruberg and Stegeman 1991; Ruberg and Hsu 1992; Shao and Chow 1994; Murphy and Weisman 1990; Yoshioka et al. 1997) other than those described above can be used in retest period or shelf life estimation. For example, if it is possible to decide in advance the acceptable difference in slope or in mean retest period or shelf life among batches, an appropriate procedure for assessing the equivalence in slope or in mean retest period or shelf life can be used to determine the data poolability. However, such a procedure should be prospectively defined, evaluated, and justified and, where appropriate, discussed with the regulatory authority. A simulation study can be useful, if applicable, to demonstrate that the statistical properties of the alternative procedure selected are appropriate (Chen et al. 1997).

B.3 Data Analysis for Multi-Factor, Full-Design Studies

The stability of the drug product could differ to a certain degree among different factor combinations in a multi-factor, full-design study. Two approaches can be considered when analyzing such data.

- The objective of the first approach is to determine whether the data from all factor combinations support the proposed shelf life.
- The objective of the second approach, testing for poolability, is to determine whether the data from different factor combinations can be combined for an overall estimate of a single shelf life.

B.3.1 Evaluating whether all factor combinations support the proposed shelf life

The objective of this approach is to evaluate whether the estimated shelf lives from all factor combinations are longer than the one proposed. A statistical model that includes all appropriate factors and factor combinations should be constructed as described in Appendix section B.3.2.2.1, and the shelf life should be estimated for each level of each factor and factor combination.

If all shelf lives estimated by the original model are longer than the proposed shelf life, further model building is considered unnecessary and the proposed shelf life will generally be appropriate as long as the guidance in sections II.D and II.E is followed. If one or more of the estimated shelf lives fall short of the proposed shelf life, model building as described in Appendix section B.3.2.2.1 can be employed. However, it is considered unnecessary to identify the final model before evaluating whether the data support the proposed shelf life. Shelf lives can be estimated at each stage of the model building process, and if all shelf lives at any stage are longer than the one proposed, further attempts to reduce the model are considered unnecessary.

This approach can simplify the data analysis of a complicated multi-factor stability study compared to the data analysis described in Appendix section B.3.2.2.1.

B.3.2 Testing for poolability

The stability data from different combinations of factors should not be combined unless supported by statistical tests for poolability.

B.3.2.1 Testing for poolability of batch factor only

If each factor combination is considered separately, the stability data can be tested for poolability of batches only, and the shelf life for each non-batch factor combination can be estimated separately by applying the procedure described in Appendix section B.2. For example, for a drug product available in two strengths and four container sizes, eight sets of data from the 2x4 strength-size combinations can be analyzed and eight separate shelf lives should be estimated accordingly. If a single shelf life is desired, the shortest estimated shelf life among all factor combinations should become the shelf life for the product. However, this approach does not take advantage of the available data from all factor combinations, thus generally resulting in shorter shelf lives than does the approach in Appendix section B.3.2.2.

B.3.2.2 Testing for poolability of all factors and factor combinations

If the stability data are tested for poolability of all factors and factor combinations and the results show that the data can be combined, a single shelf life longer than that estimated based on individual factor combinations is generally obtainable. The shelf life is longer because the width of the confidence limit(s) for the mean will become narrower as the amount of data increases when different factors, such as batches, strengths, container sizes, and/or fills, are combined.

B.3.2.2.1 Analysis of covariance

Analysis of covariance can be employed to test the difference in slopes and intercepts of the regression lines among factors and factor combinations (Chen et al. 1997; Fairweather et al. 1995). The purpose of the procedure is to determine whether data from multiple factor combinations can be combined for the estimation of a single shelf life.

The full statistical model should include the intercept and slope terms of all main effects and interaction effects and a term reflecting the random error of measurement. If it can be justified that the higher order interactions are very small, there is generally no need to include these terms in the model. In cases where the analytical results at the initial time point are obtained from the finished dosage form prior to its packaging, the container intercept term can be excluded from the full model because the results are common among the different container sizes and/or fills.

The tests for poolability should be specified to determine whether there are statistically significant differences among factors and factor combinations. Generally, the pooling tests should be performed in a proper order such that the slope terms are tested before the intercept terms and the interaction effects are tested before the main effects. For example, the tests can start with the slope and then the intercept terms of the highest order interaction, and proceed to the slope and then the intercept terms of the simple main effects. The most reduced model, obtained when all remaining terms are found to be statistically significant, can be used to estimate the shelf lives.

All tests should be conducted using appropriate levels of significance. It is recommended that a significance level of 0.25 be used for batch-related terms, and a significance level of 0.05 be used for non-batch-related terms. If the tests for poolability show that the data from different factor combinations can be combined, the shelf life can be estimated according to the procedure described in Appendix section B.1 using the combined data.

If the tests for poolability show that the data from certain factors or factor combinations should not be combined, either of two alternatives can be applied: (1) a separate shelf life can be estimated for each level of the factors and of the factor combinations remaining in the model; or (2) a single shelf life can be estimated based on the shortest estimated shelf life among all levels of factors and factor combinations remaining in the model.

B.3.2.2.2 Other methods

Alternative statistical procedures (Ruberg and Stegeman 1991; Ruberg and Hsu 1992; Shao and Chow 1994; Murphy and Weisman 1990; Yoshioka et al. 1997) to those described above can be applied. For example, an appropriate procedure for assessing the equivalence in slope or in mean shelf life can be used to determine the data poolability. However, such a procedure should be prospectively defined, evaluated, properly justified, and, where appropriate, discussed with the regulatory authority. A simulation study can be useful, if applicable, to demonstrate that the statistical properties of the alternative procedure selected are appropriate (Chen et al. 1997).

B.4 Data Analysis For Bracketing Design Studies

The statistical procedures described in Appendix section B.3 can be applied to the analysis of stability data obtained from a bracketing design study. For example, for a drug product available in three strengths (S1, S2, and S3) and three container sizes (P1, P2, and P3) and studied according to a bracketing design where only the two extremes of the container sizes (P1 and P3) are tested, six sets of data from the 3x2 strength-size combinations will be obtained. The data can be analyzed separately for each of the six combinations for shelf life estimation according to Appendix section B.3.2.1, or tested for poolability prior to shelf life estimation according to Appendix section B.3.2.2.

The bracketing design assumes that the stability of the intermediate strengths or sizes is represented by the stability at the extremes. If the statistical analysis indicates that the stability of the extreme strengths or sizes is different, the intermediate strengths or sizes should be considered no more stable than the least stable extreme. For example, if P1 from the above bracketing design is found to be less stable than P3, the shelf life for P2 should not exceed that for P1. No interpolation between P1 and P3 should be considered.

B.5 Data Analysis For Matrixing Design Studies

A matrixing design has only a fraction of the total number of samples tested at any specified time point. Therefore, it is important to ascertain that all factors and factor combinations that can have an impact on shelf life estimation have been appropriately tested. For a meaningful interpretation of the study results and shelf life estimation, certain assumptions should be made and justified. For instance, the assumption that the stability of the samples tested represents the stability of all samples should be valid. In addition, if the design is not balanced, some factors or factor interactions might not be estimable. Furthermore, for different levels of factor combinations to be poolable, it might have to be assumed that the higher order factor interactions are negligible. Because it is usually impossible to statistically test the assumption that the higher order terms are negligible, a matrixing design should be used only when it is reasonable to assume that these interactions are indeed very small, based on supporting data.

The statistical procedure described in Appendix section B.3 can be applied to the analysis of stability data obtained from a matrixing design study. The statistical analysis should clearly identify the procedure and assumptions used. For instance, the assumptions underlying the model in which interaction terms are negligible should be stated. If a preliminary test is performed for the purpose of eliminating factor interactions from the model, the procedure used should be provided and justified. The final model on which the estimation of shelf life will be based should be stated. The estimation of shelf life should be performed for each of the terms remaining in the model. The use of a matrixing design can result in an estimated shelf life shorter than that resulting from a full design.

Where bracketing and matrixing are combined in one design, the statistical procedure described in Appendix section B.3 can be applied.

B.6 References

Carstensen, J.T., "Stability and Dating of Solid Dosage Forms," *Pharmaceutics of Solids and Solid Dosage Forms*, Wiley-Interscience, 182-185, 1977.

Chen, J.J., Ahn, H., and Tsong, Y., "Shelf-life Estimation for Multifactor Stability Studies," *Drug Inf. Journal*, 31:573-587, 1997.

Fairweather, W., Lin, T.D., and Kelly, R., "Regulatory, Design, and Analysis Aspects of Complex Stability Studies," *J. Pharm. Sci*., 84:1322-1326, 1995.

Murphy, J.R. and Weisman, D., "Using Random Slopes for Estimating Shelf-life," *Proceedings of American Statistical Association of the Biopharmaceutical Section*, 196-200, 1990.

Ruberg, S.J. and Hsu, J.C., "Multiple Comparison Procedures for Pooling Batches in Stability Studies," *Technometrics*, 34:465-472, 1992.

Ruberg, S.J. and Stegeman, J.W., "Pooling Data for Stability Studies: Testing the Equality of Batch Degradation Slopes," *Biometrics*, 47:1059-1069, 1991.

Shao, J. and Chow, S.C., "Statistical Inference in Stability Analysis," *Biometrics*, 50:753-763, 1994.

Yoshioka, S., Aso, Y., and Kojima, S., "Assessment of Shelf-life Equivalence of Pharmaceutical Products," *Chem. Pharm. Bull.*, 45:1482-1484, 1997.

B.7 Figures

Figure 1

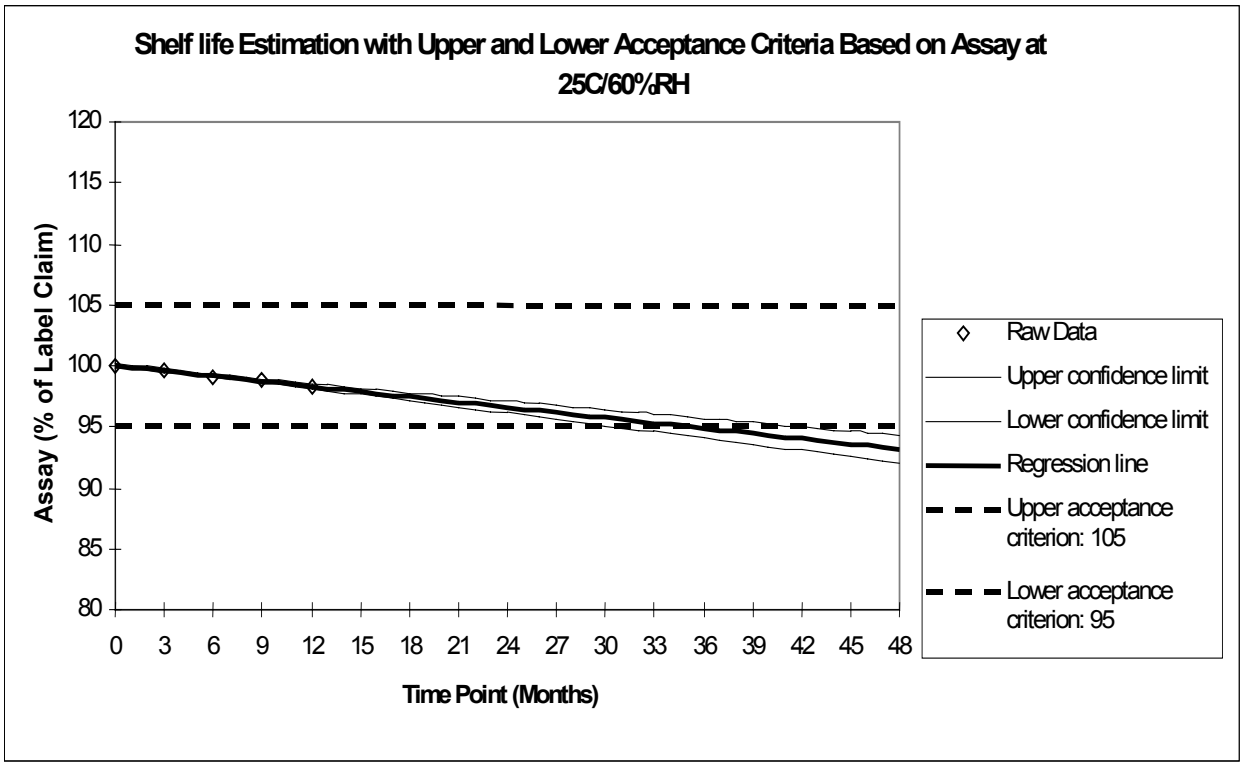
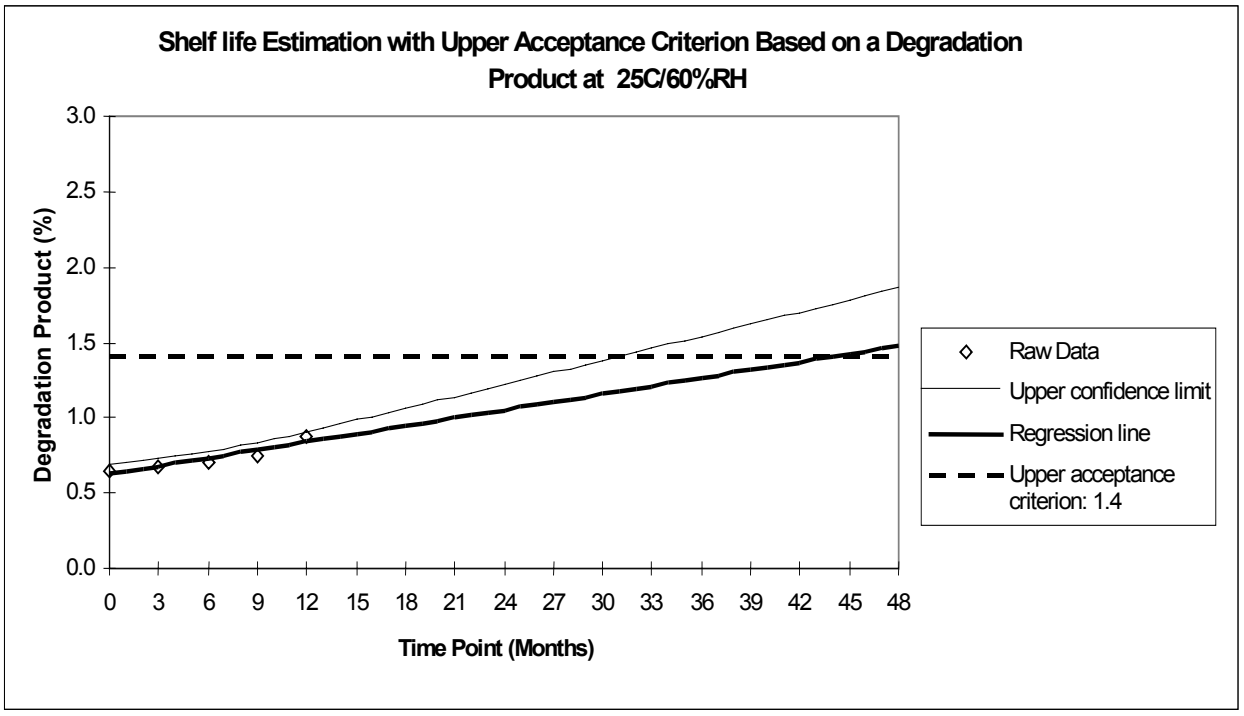


Figure 2



Ogenstad Declaration Exhibit B

ELSEVIER

Journal of Pharmaceutical and Biomedical Analysis
28 (2002) 1011–1040

JOURNAL OF
PHARMACEUTICAL
AND BIOMEDICAL
ANALYSIS

www.elsevier.com/locate/jpba

Review

Development of validated stability-indicating assay methods—critical review

Monika Bakshi, Saranjit Singh *

Department of Pharmaceutical Analysis, National Institute of Pharmaceutical Education and Research, Sector 67, S.A.S. Nagar, Mohali 160 062, Punjab, India

Received 11 October 2001; received in revised form 4 January 2002; accepted 19 January 2002

Abstract

This write-up provides a review on the development of validated stability-indicating assay methods (SIAMs) for drug substances and products. The shortcomings of reported methods with respect to regulatory requirements are highlighted. A systematic approach for the development of stability-indicating methods is discussed. Critical issues related to development of SIAMs, such as separation of all degradation products, establishment of mass balance, stress testing of formulations, development of SIAMs for combination products, etc. are also addressed. The applicability of pharmacopoeial methods for the analysis of stability samples is discussed. The requirements of SIAMs for stability study of biotechnological substances and products are also touched upon.

Keywords: Stability-indicating assay methods; Regulatory requirements; Development; Critical issues

1. Introduction

The stability-indicating assay is a method that is employed for the analysis of stability samples in pharmaceutical industry. With the advent of International Conference on Harmonisation (ICH) guidelines, the requirement of establishment of stability-indicating assay method (SIAM) has become more clearly mandated. The guidelines explicitly require conduct of forced decomposition studies under a variety of conditions, like pH, light, oxidation, dry heat, etc. and separation of drug from degradation products. The method is expected to allow analysis of individual degradation products.

A review of literature reveals a large number of methods reported over the period of last 3–4 decades under the nomenclature 'stability-indicating'. However, most of the reported methods fall short in meeting the current regulatory requirements.

Accordingly, the purpose of this write-up is to suggest a systematic approach for the development of validated SIAMs that should meet the current ICH and regulatory requirements. The discussion also touches upon various critical issues, such as the extent of separation of degrada-

* Corresponding author. Tel.: +91-172-214682; fax: +91-172-214692.
E-mail address: ssingh@niper.ac.in (S. Singh).

0731-7085/02/$ - see front matter
PII: S0731-7085(02)00047-X

tion products, establishment of mass balance, etc., which are important with respect to the development of stability-indicating assays, but are not yet fully resolved. Some other aspects like suitability of pharmacopoeial methods for the purpose and the role of SIAMs in stability evaluation of biological/biotechnological substances and products are also delved upon.

2. Regulatory status of stability-indicating assays

The ICH guidelines have been incorporated as law in the EU, Japan and in the US, but in reality, besides these other countries are also using them. As these guidelines reflect the current inspectional tendencies, they carry the de facto force of regulation. The ICH guideline Q1A on Stability Testing of New Drug Substances and Products [1] emphasizes that the testing of those features which are susceptible to change during storage and are likely to influence quality, safety and/or efficacy must be done by validated stability-indicating testing methods. It is also mentioned that forced decomposition studies (stress testing) at temperatures in 10 °C increments above the accelerated temperatures, extremes of pH and under oxidative and photolytic conditions should be carried out on the drug substance so as to establish the inherent stability characteristics and degradation pathways to support the suitability of the proposed analytical procedures. The ICH guideline Q3B entitled 'Impurities in New Drug Products' emphasizes on providing documented evidence that analytical procedures are validated and suitable for the detection and quantitation of degradation products [2]. It is also required that analytical methods should be validated to demonstrate that impurities unique to the new drug substance do not interfere with or are separated from specified and unspecified degradation products in the drug product. The ICH guideline Q6A, which provides note for guidance on specifications [3], also mentions the requirement of stability-indicating assays under Universal Tests/Criteria for both drug substances and drug products. The same is also a requirement in the guideline Q5C on Stability Testing of Biotechnological/Biological Products [4]. Since there is no single assay or parameter that profiles the stability characteristics of such products, the onus has been put on the manufacturer to propose a stability-indicating profile that provides assurance on detection of changes in identity, purity and potency of the product.

Unfortunately, none of the ICH guidelines provides an exact definition of a stability-indicating method. Elaborate definitions of stability-indicating methodology are, however, provided in the United States-Food and Drug Administration (US-FDA) stability guideline of 1987 [5] and the draft guideline of 1998 [6]. Stability-indicating methods according to 1987 guideline were defined as the '*quantitative analytical methods that are based on the characteristic structural, chemical or biological properties of each active ingredient of a drug product and that will distinguish each active ingredient from its degradation products so that the active ingredient content can be accurately measured.*' This definition in the draft guideline of 1998 reads as: '*validated quantitative analytical methods that can detect the changes with time in the chemical, physical, or microbiological properties of the drug substance and drug product, and that are specific so that the contents of active ingredient, degradation products, and other components of interest can be accurately measured without interference.*' The major changes brought in the new guideline are with respect to (i) introduction of the requirement of validation, and (ii) the requirement of analysis of degradation products and other components, apart from the active ingredients(s).

The requirement is also listed in World Health Organization (WHO), European Committee for Proprietary Medicinal Products and Canadian Therapeutic Products Directorate's guidelines on stability testing of well established or existing drug substances and products [7–9].

Even the United States Pharmacopoeia (USP) has a requirement listed under 'Stability Studies in Manufacturing', which says that samples of the products should be assayed for potency by the use of a stability-indicating assay [10]. The requirement in such explicit manner is, however, absent in other pharmacopoeias.

Current ICH guideline on Good Manufacturing Practices for Active Pharmaceutical Ingredients (Q7A), which is under adoption by WHO, also clearly mentions that the test procedures used in stability testing should be validated and be stability-indicating [11].

3. Review of the literature on stability-indicating assays

In absence of any guidance from regulatory agencies on practical steps to be followed for establishment of stability-indicating assays, a search was done on the available information in literature. The literature was found to be replete with publications on development of stability-indicating assays of specific drugs. A general review was published as early as 1971, and it gave general principles and discussed the methods developed till that period [12]. Kumar and Sunder also discussed the perspective of stability-indicating testing procedures [13]. Subsequently, Ho and Chen [14,15] reviewed stability-indicating high-performance liquid chromatography (HPLC) assay methods reported till 1996. A compilation of stability-indicating assays (>500) for various drugs was published in 1999 by Xu and Trissel [16]. A more recent publication is in the form of a chapter in the book 'Drug Stability: Principles and Practices' by Carstensen and Rhodes [17], which provides general discussion on HPLC method development and validation, with emphasis on stability-indicating assays. On whole, a critical guidance document on the topic, which encompasses current ICH requirements and discusses various critical issues, is still elusive.

4. An assessment of the extent to which the reported methods meet current regulatory requirements

A review of various literature reports shows that very few methods that are titled or claimed to be stability-indicating fit into the current definition of a stability-indicating assay in true sense. While the current requirement is of subjecting the drug substance to variety of stress conditions and then separation of drug from all degradation products, many studies have just shown the separation of drug from known synthetic impurities and/or potential degradation products without subjecting it to any type of stress (Table 1). There are also reports in which drug has been decomposed by exposing it to one (Table 2), two (Table 3), three (Table 4), four (Table 5) or more (Table 6) conditions among acidic, neutral or alkaline hydrolysis, photolysis, oxidation and thermal stress. Thus very few studies are truly stability-indicating, where drug has been exposed to all types of stress conditions and attempts have been made to separate the drug from degradation products and the latter among themselves. Different approaches have been employed in these cases, in absence of any defined requirements. There are some reports where directly the formulation, instead of the drug substance, has been subjected to stress studies for establishment of the stability-indicating behavior (Table 7). A few reports exist even on combinations of drugs (Table 8).

It may be pertinent to add here that the examples cited in Tables 1–8 are only representative and do not mean comprehensive coverage of all literature reports.

5. Techniques employed in literature reports

If one critically evaluates the literature reports, titrimetric, spectrophotometric and chromatographic techniques have been commonly employed in analysis of stability samples. There are also sporadic reports of the use of miscellaneous techniques.

5.1. Titrimetric and spectrophotometric

In these methods, usually the objective is the analysis of the drug of interest alone in the matrix of excipients, additives, degradation products, impurities, etc., and also other drugs in case of the combination products. Their advantage is the low cost and simplicity, though sometimes they are not sensitive. Due to limitation of specificity, there are hardly any reports these days on their

use for the assay of stability samples. However, a few reports involving derivative spectroscopy have been published lately [122–125].

5.2. Chromatographic

Because of the very nature of requirement of separation of multiple components during analysis of stability samples, chromatographic methods have taken precedence over the conventional methods of analysis. Other than separation of multiple components, the advantage of chromatographic methods is that these possess greater accuracy and sensitivity for even small quantities of degradation products produced. Various chromatographic methods that have been used are thin-layer chromatography (TLC), high-performance thin-layer chromatography (HPTLC), gas chromatography (GC), HPLC and newer technique like capillary electrophoresis (CE).

TLC is a simple technique that has been used in the past for developing SIAMs [126–128]. Its disadvantages, such as variability and non-quantitative nature, limit its use as a basic technique for SIAM development. However, it is very much used, especially during initial degradation [129] and stress studies to study the number of degradation products formed, to identify the products formed through matching studies using standards, and even for isolation where preparative TLC is employed.

A large number of publications have appeared in the last decade on the use of HPTLC for stability-indicating method development

Table 1
Selected reports of 'stability-indicating' methods where no stress testing has been done

	Drug	Methodology	Ref.
Separation from process impurities	Benazepril hydrochloride	HPLC	[18]
	Ribavirin	HPLC	[19]
Separation from known/potential degradation product(s)	Betamethasone 17-benzoate	HPLC	[20]
	Canrenone	HPLC	[21]
	Erythromycin estolate	HPLC	[22]
	Ethacrynic acid	HPLC	[23]
	Phenylbutazone	HPLC	[24]
	Sibutramine hydrochloride	HPLC	[25]
	Sulphacetamide	HPLC	[26]
	Homatropine methylbromide	UV spectrophotometry	[27]
	Cholesterol lowering drug	Micellar electrokinetic chromatography	[28]
	Felodipine	Supercritical-fluid chromatography	[29]
Separation from known/potential degradation products and process impurities	Benzodiazepines	HPLC	[30]
	Ranitidine	HPLC	[31]
	Temazepam	HPLC	[32]
	Melphalan	HPLC	[33]
	Piroxicam	HPTLC	[34]
	Tinidazole	HPTLC	[35]
	Fenclorac	GLC	[36]
	Azathioprine	CE	[37]

Table 2
Selected reports of 'stability-indicating' methods where only one stress condition has been employed

Stress condition	Drug	Methodology	Ref.
Acid	Dyclonine hydrochloride	HPLC	[38]
	Flunarizine dihydrochloride	HPLC	[39]
	Lisinopril	UV spectrophotometry	[40]
	Norfloxacin	UV spectrophotometry	[41]
	Lisinopril	Derivative UV spectrophotometry	[42]
Alkali	Allantoin	HPLC	[43]
	Meperidine hydrochloride	HPLC	[44]
	Metronidazole	HPLC	[45]
	Benazepril hydrochloride	UV spectrophotometry	[46]
	Carbachol	IR spectrophotometry	[47]
Neutral	Physostigmine salicylate	HPLC	[48]
Oxidation	Nortriptyline hydrochloride	UV spectrophotometry	[49]
Light	Atenolol	HPLC	[50]
	Danazol	HPLC	[51]
	Trifluoperazine hydrochloride	HPLC	[52]
	Nifedipine	HPTLC	[53]
	Ranitidine hydrochloride	Spectrodensitometric TLC	[54]
	Piroxicam	HPLC, HPTLC, CE	[55]

[34,35,53,63,69,74]. This technique overcomes the shortcomings of TLC, and is reliable, fast and accurate for quantitative drug analysis. Moreover, many samples can be run simultaneously using a small quantity of mobile phase, thus minimizing analysis time and cost per analysis. Unfortunately, its limitation is that the equipment is not routinely available in every laboratory.

GC is stability-indicating but it is not very versatile, as the drug substance may be non-volatile or thermally unstable. Further any attempt to increase the volatility of the drug and components by increasing the temperature may lead to degradation or racemization. Therefore, there are very few reports on the use of GC [130–132] for the purpose of establishment of SIAMs.

In comparison, HPLC has been very widely employed. It has gained popularity in stability studies due to its high-resolution capacity, sensitivity and specificity. Non-volatile, thermally unstable or polar/ionic compounds can also be analyzed by this technique. Therefore, most of the SIAMs have been established using HPLC, which is evident from the lists given in Tables 1–8.

5.3. Miscellaneous

A few studies have also reported the use of proton nuclear magnetic resonance (NMR) spectroscopy for the development of SIAMs [133–135]. CE is the latest entry to the techniques for the development of SIAMs [37,136–138]. It has the advantage of high sensitivity, resolution and high efficiencies with minimal peak dispersion.

Table 3
Selected reports of 'stability-indicating' methods where two stress conditions have been employed

Stress conditions	Drug	Methodology	Ref.
Acid, alkali	Betaxolol hydrochloride	HPLC	[56]
	Captopril	HPLC	[57]
	Cephalexin	HPLC	[58]
	Ciprofloxacin	HPLC	[59]
	Indapamide	HPLC	[60]
	Omeprazole	HPLC	[61]
	Yohimbine hydrochloride	HPLC	[62]
	Nimesulide	HPTLC	[63]
Acid, light	Trimethoprim	HPLC	[64]

Table 4
Selected reports of 'stability-indicating' methods where three stress conditions have been employed

Stress condition	Drug	Methodology	Ref.
Acid, neutral, alkali	Diaziridinyl benzoquinone	HPLC	[65]
	Xilobam	UV assay supported by TLC and HPLC	[66]
Acid, alkali, light	Carprofen	HPLC	[67]
	Nitrendipine	HPLC, HPTLC and UV spectrophotometry	[68]
	Estradiol	HPTLC	[69]
Acid, alkali, oxidation	Mefenamic acid	HPLC	[70]
	Morphine	HPLC	[71]
Acid, alkali, dry heat	Amphotericin B	HPLC	[72]
	Hydrochlorthiazide	HPLC	[73]
	Timolol maleate	HPTLC	[74]

There are several publications involving use of hyphenated GC-MS [139], LC-MS [140–147], LC-MS-MS [141,142,145,148], LC-NMR [141,147] and CE-MS [149,150] techniques for identity confirmation of known and unknown degradation products and their selective determination.

6. Development of validated SIAMs that are likely to meet regulatory requirements

Though the requirements with respect to SIAM have been spelt out in regulatory documents, information on the basic steps to be followed for the development and validation of stability-indicating methods is neither provided in the regulatory guidelines nor in the pharmacopoeias. Therefore, the practical steps involved in the development of SIAMs are discussed below. It is expected that by following the steps, one should be in a position to develop a SIAM that would meet the regulatory requirements. Our discussion is typically oriented towards development of SIAMs by HPLC, as it is found that 85–90% of the methods reported in literature are by this technique.

6.1. Step I: critical study of the drug structure to assess the likely decomposition route(s)

This should be the first element whenever one takes up the project on establishment of a SIAM. Much information can simply be gained from the structure, by study of the functional groups and other key components. There are definite functional group categories, like amides, esters, lactams, lactones, etc. that undergo hydrolysis [151], others like thiols, thioethers, etc. undergo oxidation [152], and compounds like olefins, aryl halo derivatives, aryl acetic acids, and those with aromatic nitro groups, N-oxides undergo photodecomposition [153].

Most of the new drugs are congeners of existing drug molecules, and there are very few new drugs, which originate from absolutely new leads. For a new congener, its degradation chemistry can be easily postulated based on the reported behavior of other drugs in the series. For example, there are more than 40 penicillins in clinical practice today and almost all of them follow the same degradation behavior at the beta-lactam moiety. Most of them also follow similar subsequent reactions (Fig. 1). Similarly, studies in our laboratory have shown that three alpha-adrenergic blockers (prazosin, doxazosin and terazosin) that have similar parent structure follow the same hydrolysis route involving breakage of the amide bond (Fig. 2).

However, a word of caution is that in some congeners there might even be existence of a totally new degradation behavior, particularly when there is overwhelming influence of the substituent. A typical example here is that of aminopenicillins, which show formation of polymers [154]. Other example is that of 2-methyl-5-ni-

troimidazole series of drugs, including metronidazole, ornidazole, tinidazole and secnidazole. Fig. 3 shows the known decomposition behavior of these drugs. Tinidazole and secnidazole are reported to decompose in alkaline conditions to the parent nucleus 2-methyl-5-nitroimidazole [155,156], while metronidazole degrades to simple compounds like ammonia and acetic acid [157]. In contrast, ornidazole neither yields the nucleus nor undergoes complete decomposition. It yields ornidazole diol via an intermediate ornidazole epoxide in alkaline medium [158].

Thus one can have a good starting point from the study of degradation behavior of congeners, but critical requirement here is the conduct of an in-depth literature survey. For information on degradation chemistry of like drugs, one can look into the treatises like *Analytical Profiles of Drug Substances* [159] and the monographs provided by Connors et al. [151]. Specific searches can even be made through the use of abstracts, the internet search engines and the Chemweb.

6.2. Step II: collection of information on physicochemical properties

Before method development is taken up, it is generally important to know various physicochemical parameters like p*K*a, log *P*, solubility, absorptivity and wavelength maximum of the drug in question. The knowledge of p*K*a is important as most of the pH-related changes in retention occur at pH values within ± 1.5 units of the p*K*a value. The ionization value also helps in selecting the pH of the buffer to be used in the mobile phase [160]. The knowledge of log *P* for the drug and the identified degradation products provides good insight into the separation behavior likely to be obtained on a particular stationary phase. p*K*a and log *P* can be practically determined or even theoretically calculated using the commercial software, such as Pallas (CompuDrug Chemistry Ltd., Budapest, Hungary), CLOGP (Pamona College, Pamona, USA), etc.

The analysis of the drug or degradation products requires that they are soluble in HPLC compatible solvents in the first place. The availability of the solubility data in aqueous, organic and commonly used HPLC solvents and their combinations can thus prove to be very useful in the selection of the sample solvent and the mobile phase.

As the HPLC analysis employing a UV detector is usually carried out at the wavelength maximum or at a wavelength where all components show good absorbance, therefore, the necessity to know the wavelength maxima and extinction of the drug and degradation products in different solvents and at different pH becomes an absolute requirement. This may be an easy exercise when the degradation products are known and available in the pure form. But when it is a new drug for which degradation pattern has not yet been estab-

Table 5
Selected reports of 'stability-indicating' methods where four stress conditions have been employed

Stress conditions	Drug	Methodology	Ref.
Acid, neutral, alkali, oxidation	Clonazepam	HPLC	[75]
	Dipyridamole	HPLC	[76]
	Esmolol hydrochloride	HPLC	[77]
Acid, alkali, oxidation, dry heat	Suprofen	HPLC	[78]
Acid, neutral, alkali, light	Guanabenz	UV spectrophotometry	[79]
	Tolmetin sodium	UV spectrophotometry	[80]
Acid, alkali, light, thermal (methanolic solution)	Retinoic acid	HPLC	[81]
Acid, alkali, oxidation, light	Trimetazidine	HPLC	[82]
	Trimetazidine	HPTLC	[83]
Acid, oxidation, light, dry heat	Fentanyl	HPLC	[84]

Table 6
Selected reports of 'stability-indicating' methods where five (and additional) stress conditions have been employed

Stress conditions	Drug	Methodology	Ref.
Acid, alkali, oxidation, dry heat, light	Sodium levothyroxine	HPLC	[85]
	Enalapril maleate	HPLC	[86]
Acid, alkali, oxidation, dry heat, light (separation from synthetic impurities also seen)	Sildenafil citrate	HPLC	[87]
Acid, neutral, alkali, oxidation, light	Nicardipine hydrochloride	HPLC	[88]
Acid, alkali, oxidation, dry heat, wet heat, light dry, light wet	Paroxetine	HPLC	[89]
Acid, alkali, oxidation, dry heat, light, reduction	Cyproterone acetate	HPLC	[90]
Acid, alkali, light, oxidation, dry heat, moisture, sonication	Buspirone hydrochloride	HPLC	[91]

lished, the same might prove to be a difficult exercise. In the latter case, the suggested way is to subject the drug to stress studies (see Section 6.3 for details) and to observe changes in the spectrum [161], first individually in each reaction solution and then in a mixture of all the solutions. This gives a fair idea (though not absolutely) on the shifts in wavelength spectra during the reaction and also guides on the best wavelength for analysis. If necessary, more than one wavelength can be selected for analysis, but taking the benefit of the same requires a dual or multi-wavelength detector. The best choice thence is using a photodiode array (PDA) detector, which allows recording of UV–visible spectrum of the components, as they get resolved on the stationary phase. Later necessary inferences can be made based on the records.

6.3. Step III: stress (forced decomposition) studies

The next step in the development of SIAM is the conduct of forced decomposition studies to generate degradation products of the drug. The ICH guideline Q1A suggests the following conditions to be employed: (i) 10 °C increments above the accelerated temperatures (e.g. 50 °C, 60 °C, etc.), (ii) humidity where appropriate (e.g. 75% or greater), (iii) hydrolysis across a wide range of pH values, (iv) oxidation and (v) photolysis. However, the guideline provides no details on how hydrolytic, photolytic and oxidative studies have to be actually performed. On the other hand, the information is available in literature but in a staggered way, with suggested approaches differing a lot from one another [162,163]. A comprehensive document providing guidance on the practical conduct and issues related to stress testing under variety of ICH prescribed conditions has been published lately [164]. This report from the authors proposes a classification scheme and offers decision trees to help in the selection of the right type of stress condition in a minimum number of attempts.

The hydrolytic degradation of a new drug in acidic and alkaline conditions can be studied by refluxing the drug in 0.1 N HCl/NaOH for 8 h. If reasonable degradation is seen, testing can be stopped at this point. However, in case no degradation is seen under these conditions, the drug should be refluxed in acid/alkali of higher strengths and for longer duration. Alternatively, if total degradation is seen after subjecting the drug to initial conditions, acid/alkali strength can be decreased along with decrease in the reaction temperature. In a similar manner, degradation under neutral conditions can be started by refluxing the drug in water for 12 h. Reflux time should be increased if no degradation is seen. If the drug is found to degrade completely, both time and temperature of study can be decreased.

To test for oxidation, it is suggested to use hydrogen peroxide in the concentration range of 3–30%. The photolytic studies should be carried out by exposure to light, using either a combination of cool white and ultraviolet fluorescent lamps, or one among the xenon and metal halide lamps. Exposure energy should be minimum of 1.2 million lux h fluorescent light and 200 W h/m^2 UV and if decomposition is not seen, the intensity should be increased by five times. In case still no decomposition takes place, the drug can be declared photostable.

A minimum of four samples should be generated for every stress condition, viz. the blank solution stored under normal conditions, the blank subjected to stress in the same manner as the drug solution, zero time sample containing the drug which is stored under normal conditions and the drug solution subjected to stress treatment. The comparison of the results of these provides real assessment of the changes. Furthermore, it is advised to withdraw samples at different time periods for each reaction condition. By doing so, one can get a clear idea on the number of products formed, their relative strengths and whether they are stable or unstable, resulting further in newer products. This information is essential in establishment of SIAMs.

The studies should be initiated at a concentration of 1 mg/ml. If solubility is a limitation, varying amounts of methanol may be used to get a clear solution or even the testing can be done on a suspension [165]. By using drug concentration of 1 mg/ml, it is usually possible to get even minor decomposition products in the range of detection. If several degradation products are formed in different conditions, the establishment of SIAM may involve a lot of development work. For this, repeat injections of reaction solutions might be required. Therefore, the volume of samples subjected to stress studies should be in sufficient quantity and also enough sample volume should be drawn at each period. The withdrawn samples can be stored in cold cabinets to stop further reaction. The aliquots might be diluted or neutralized before injecting into HPLC.

Table 7
Reports of 'stability-indicating' methods on drug formulations

Stress conditions	Drug	Dosage form	Methodology	Ref.
Acid	Fluconazole	Admixtures	GC	[92]
	Flucytosine	Extemporaneous solutions	HPLC	[93]
	Levothyroxine sodium	Tablets	HPLC	[94]
	Ipratropium bromide	Metered dose inhalers and inhalation solutions	HPLC	[95]
Acid, alkali	Ganciclovir	Capsules	HPLC	[96]
Light, thermal	Sodium levothyroxine	Tablets	HPLC	[85]
Acid, alkali, oxidation	Pentoxifylline	Suspension	HPLC	[97]
	Granisetron hydrochloride	Injection	HPLC	[98]
Acid, alkali, oxidation, thermal	Chlorobutanol	Ointment	HPLC	[99]
Acid, alkali, oxidation, light	Fotemustine	5% dextrose	HPLC	[100]
	Efavirenz	Capsules	HPLC	[101]
Acid, oxidation, light, thermal	Fentanyl	Injection	HPLC	[84]
Acid, alkali, oxidation, thermal, light	Cyclosporine	Oral solution	HPLC	[102]
Acid, alkali, thermal, light, 45 °C/75% RH for 2 weeks	Aspirin and warfarin sodium	Tablets	HPLC	[103]
Aged samples (3 years at 40 °C and 75% RH)	Losartan	Tablets	HPLC and LC-MS	[104]

Table 8
Selected reports of 'stability-indicating' methods for combination of different drugs

Stress condition(s)	Drug	Dosage form	Methodology	Type of study	Ref.
No stress	Otilonium bromide and diazepam	Finished pharmaceutical dosage forms	HPLC	Separation from related compounds and potential degradation products of both drugs shown	[105]
	Naphazoline and tetrahydrozoline	Ophthalmic preparations	HPLC	Separation seen using potential degradation products of both drugs	[106]
	Captopril and hydrochlorthiazide	Tablets	HPLC	Separation seen using potential degradation products of both drugs	[107]
	Sulfonamides and erythromycin ethyl succinate	Oral suspension	HPLC	Separation seen using potential degradation products of both drugs	[108]
	Hydroxyzine hydrochloride and benzyl alcohol	Injection	HPLC	Separation seen using potential degradation products of both drugs	[109]
Oxidation	Oxycodone and lidocaine	Rectal gel	HPLC	Stress studies done on individual drugs	[110]
Acid, alkali	Ramipril and hydrochlorthiazide	Different dosage forms	HPLC	Only ramipril exposed to stress conditions. Both drugs analyzed in presence of ramipril degradation products	[111]
Thermal	Minocycline hydrochloride and rifampicin	Intravenous solutions	HPLC	Stress studies done on individual drugs. Samples analyzed individually using different stability-indicating methods for each drug	[112]
Acid, neutral, alkali	Hydrochlorthiazide and triamterene	Capsules	HPLC	Capsules subjected to stress studies	[113]
Acid, alkali, oxidation	Pseudoephedrine and cetrizine	Pharmaceutical dosage forms	HPLC	Stress studies done on individual drugs	[114]
	Cisatracurium besylate and propofol	Mixtures	HPLC	Mixtures subjected to stress studies	[115]
Acid, alkali, thermal, light, 45 °C/75% RH for 2 weeks	Aspirin and warfarin sodium	Tablets	HPLC	Tablets subjected to stress studies and spiking with warfarin sodium related substances done	[103]

Table 8 (*Continued*)

Stress condition(s)	Drug	Dosage form	Methodology	Type of study	Ref.
Acid, alkali, thermal, light	Phenylepherine hydrochloride, phenylpropanolamine hydrochloride and guaifenesin	Capsules	HPLC	Well-mixed capsule material exposed to acidic and alkaline hydrolysis, thermal degradation and photolysis	[116]
Acid, alkali, oxidation, light	Paclitaxel with ondansetron hydrochloride or ranitidine hydrochloride	Admixtures in 5% dextrose	HPLC	Only paclitaxel subjected stress studies. Drugs analyzed by individual assays	[117]
Acid, alkali, thermal, oxidation, light	Acetazolamide, allopurinol, azathioprine, clonazepam, flucytosine	Extemporaneous oral solution	HPLC	Stress studies done on individual drugs and composite chromatogram after degradation in all conditions shown. Samples analyzed individually using different stability-indicating methods for each drug	[118]
	Baclofen, captopril, diltiazem hydrochloride, dipyridamole and flecainide acetate	Extemporaneous oral solution	HPLC	Stress studies done on individual drugs and composite chromatogram after degradation in all conditions shown. Samples analyzed individually using different stability-indicating methods for each drug	[119]
	Hydrocodone bitartarate and acetaminophen	Tablets	HPLC	Stress studies performed only on hydrocodone	[120]
Acid, neutral, alkali, thermal, light and/or high humidity	Norgestimate and ethinyl estradiol	Tablets	HPLC	Norgestimate subjected to acid, neutral, alkali, thermal and light degradation. Ethinyl estradiol subjected to thermal and high humidity stress. Separation shown for each drug in the presence of degradation products of other	[121]

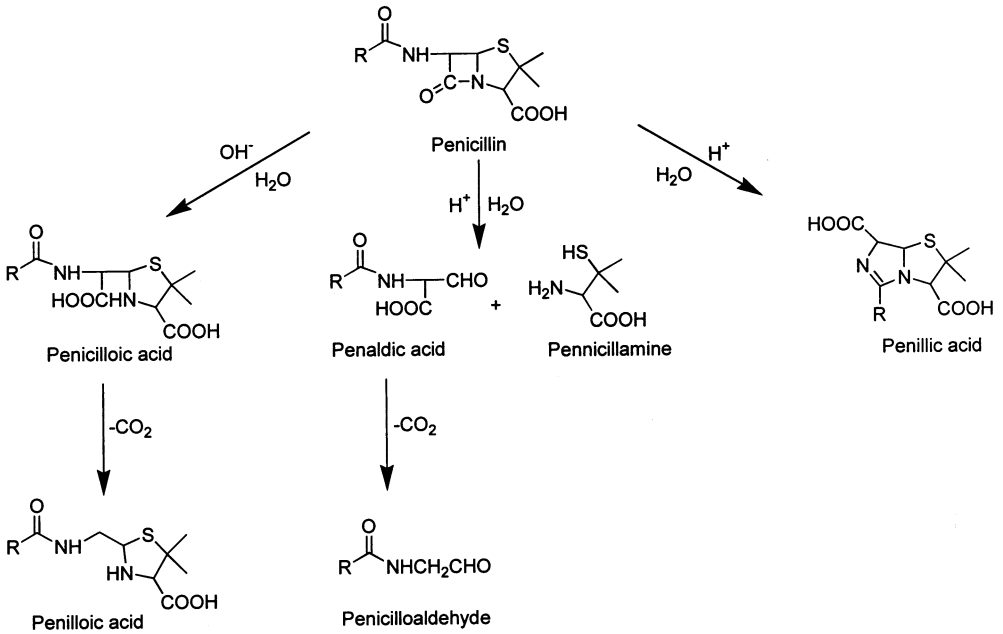


Fig. 1. Known degradation behavior of penicillins under different hydrolytic conditions.

6.4. Step IV: preliminary separation studies on stressed samples

The stress samples so obtained are subjected to preliminary analyses to study the number and types of degradation products formed under various conditions. For doing so, the simplest way is to start with a reversed-phase octadecyl column, preferably a new or the one in a healthy condition. Well-separated and good quality peaks at the outset provide better confidence because of the unknown nature of products formed during stressing. It should be preferred to use water–methanol or water–acetonitrile as the mobile phase in the initial stages. The use of buffers is not suggested at this stage because as is normally required, one can extend the buffer-free mobile phase to preparative LC or LC-MS studies. Between methanol and acetonitrile, the former should be preferred due to its low cost. The wisdom from previous studies on the development of assay method for the drug can also be applied here and the organic modifier can be chosen accordingly. The solvent can be changed, if the peak shape or separation problems are seen.

Initially, water:organic modifier ratio can be fixed at 50:50 or can be suitably modified so as to obtain the capacity factor of around 5–10 for the drug. As degradation products from drugs are generally polar in nature (of course with exceptions), pushing the drug peak to say $\sim$15 min or somewhat more in a 25-cm column, can result in separation of even several degradation products, when formed. The retention time can be brought earlier or pushed further by changing the mobile phase but it should not be pushed very far, as though it might lead to an overall increase in resolution (and ruggedness), but oppositely the peaks flatten out resulting in a decrease in sensitivity. Normally, the total run time should be 2.5 times more than the drug peak, at least in initial studies, and this long period is to show up any peak that would elute later to the drug peak.

The detection wavelength can be set, based on the study of spectral behavior of degraded samples, as discussed earlier. The injection volume and the flow rate can be suitably adjusted based on the length of the column.

Using these chromatographic conditions, one should follow the changes in all the stress sam-

ples, at various time periods. The results should be critically compared with the blank solutions injected in a similar manner. It should be observed whether the fall in drug peak is quantitatively followed by a corresponding rise in the degradation product peaks. It should not be taken as a surprise if the peak rise is not in correspondence to fall of the drug. This is because the drug and its products can have very different extinction values. Even there can be situations where no additional peak appears in the chromatogram, other than the drug. A typical example is given in Fig. 4 where the drug fall is clearly seen, but with no additional peak rise. Such a situation can either arise due to the formation of non-chromophoric products or due to decomposition of drug to low molecular weight fractions. In such situations, the detection at multiple wavelengths or the use of LC-MS becomes necessary. Sometimes the absence of simultaneous rise in degradation product peak might also be due to total insolubility of the product in the reaction solutions, which can be confirmed through physical observation of the reaction mixture. In such case the product can be separated and can be injected separately using the solvent in which it is soluble to find out its retention time (RT) in the chromatogram. Later, during the final method development changes can be made in mobile phase or the sample solvent to have the product shown up in the chromatogram. Even the absence of degradation peak can happen when the product is colored and shows no UV absorption at a particular wavelength at which the analysis has been conducted. This can be verified by simple observation whether any color has developed in the reaction solution. Here also suitable adjustment in the wavelength of analysis can be made for the product to appear in the chromatogram.

6.5. Step V: final method development and optimization

Subsequent to preliminary chromatographic studies, the RT and relative retention times (RRT) of all products formed should be tabulated for each reaction condition. Special attention is then paid to those components whose RT or RRT is very close. PDA spectra or LC-MS profile of such components are obtained and critically evaluated to ascertain whether the products are same or different. It has happened with us once that what we were considering as a drug peak, proved rather to be due to the degradation product. The drug peak appeared at a particular RT in acid conditions, but when the reaction was done in alkali, again a peak appeared at almost same RT. However, the LC-MS studies indicated it to be a different product. It was later established that the drug was almost instantly converted when brought in contact with the alkali and the product was formed quantitatively. Therefore, if PDA or

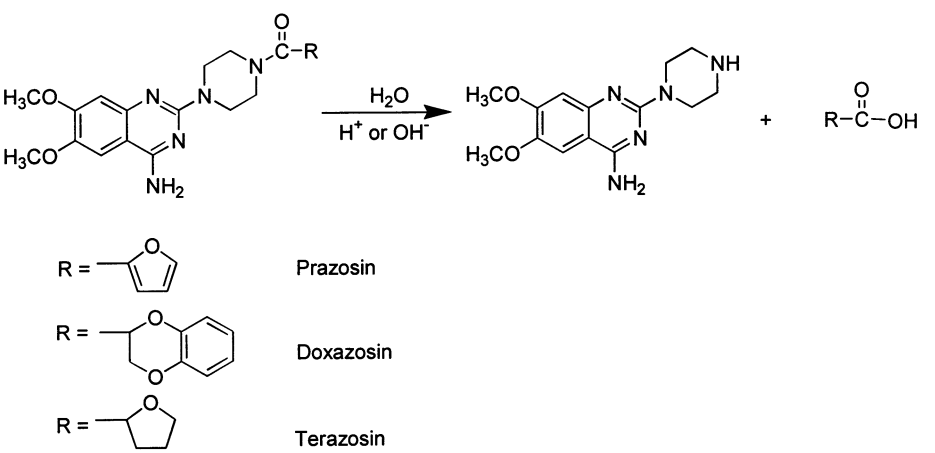


Fig. 2. Hydrolysis of different alpha-adrenergic blockers.

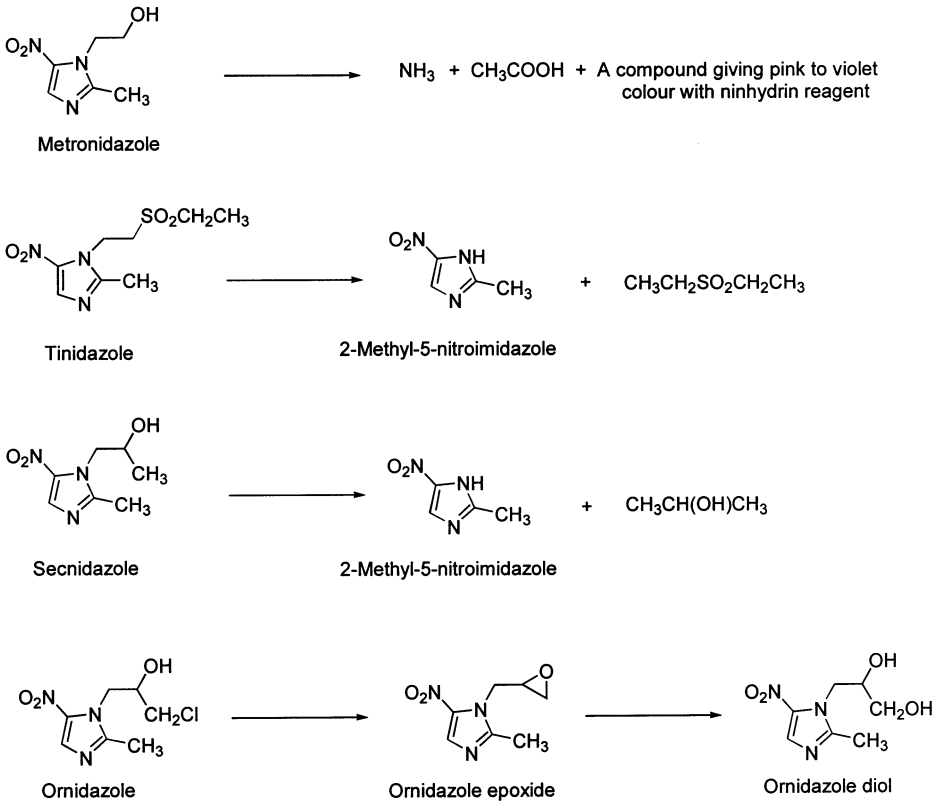


Fig. 3. Reported degradation products of different 2-methyl-5-nitroimidazoles.

LC-MS results suggest that any of the products are different but are co-eluting, then suitable modification should be done in the chromatographic method to achieve a satisfactory resolution.

In the final step, a mixture of the reaction solutions is prepared, and subjected again to resolution behavior study. While making this mixture, it is not always necessary to add all reaction solutions withdrawn at different time for all conditions. That would make the situation too complex. Rather, only those solutions are mixed where different products are formed in sufficient quantity. Resolution in the mixture is studied closely, to see whether the resolution is similar to that obtained in individual samples. This is important to rule out any changes that can happen when reaction solutions of different pH and media (3–30% hydrogen peroxide solution) are mixed. There might be a situation where products show different chromatographic behavior in a mixture.

To separate close or co-eluting peaks, the method is optimized, by changing the mobile phase ratio, pH, gradient, flow rate, temperature, solvent type, and the column and its type. Details of the basic issues in method development are not discussed here, as they are covered elsewhere [17].

A typical example of the study in author's laboratory where the desired separation was achieved in a mixture of various reaction solutions is shown in Fig. 5. The steps undertaken in optimization of the developed method can be found in a published report [166].

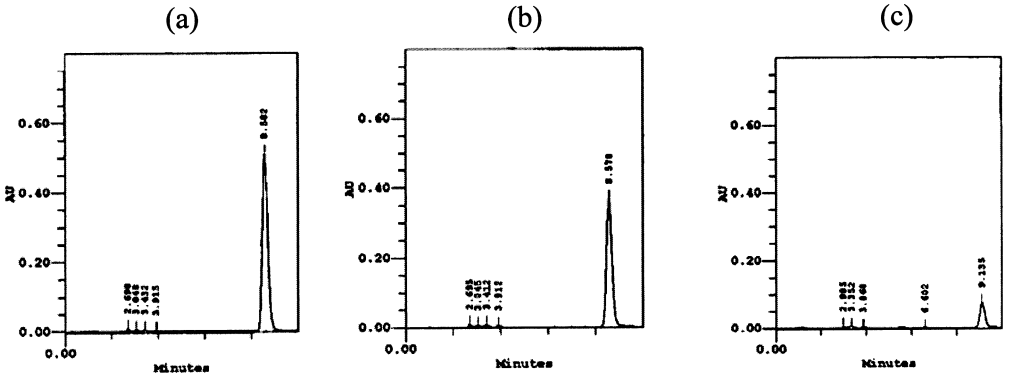


Fig. 4. Degradation of metronidazole in acidic conditions under light: (a) initial sample (b) 3 day sample (c) 12 day sample.

6.6. Step VI: identification and characterization of degradation products, and preparation of standards

Before moving to the validation of a SIAM, it is necessary to identify the drug degradation products and arrange for their standards. These are required to establish specificity/selectivity of the method. The work on this aspect can even be initiated once an idea on the nature and number of degradation products formed under different degradation conditions is obtained from preliminary separation studies.

To identify the resolved products, a conventional way is to isolate them and determine the structure through spectral (MS, NMR, IR, etc.) and elemental analysis. However, this approach is tedious and time consuming when multiple degradation products are formed. Against it, the modern approach is to use hyphenated LC techniques coupled with mass spectrometry. This strategy integrates in a single instrument approach, analytical HPLC, UV detection, full scan mass spectrometry (LC-MS) and tandem mass spectrometry (LC-MS-MS) and provides a fair idea on identity of resolving components. These days a further integrated approach is becoming popular wherein LC-MS or LC-MS-MS is employed to obtain molecular weight and fragmentation information, and further detailed structural information is obtained through LC-NMR analysis. The integrated approach provides rapid and unambiguous identification of several degradation products at one time.

Regarding the product standards, a direct way is to procure them from commercial sources (Section 8). However, in case they are not available commercially, they have to be either isolated from the degradation reaction solutions or synthesized in the laboratory. To isolate a product, the best way is to identify a reaction condition where it is formed selectively. If the product precipitates or

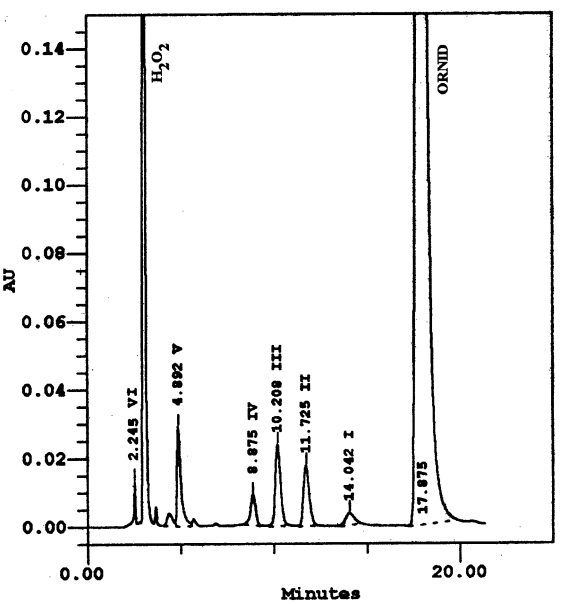


Fig. 5. Example of a 'Selective SIAM' showing separation of different degradation products of ornidazole in a mixture of reaction solutions. (Reprinted from Ref. [166] with permission from authors and Elsevier Science).

crystallizes on its own on completion of the reaction, it can be recovered simply. Otherwise, the reaction mixture can be lyophilized directly. If freeze-drying is done after neutralization of the reaction mixture, the product can be recovered by extraction with dry methanol or any other suitable dry solvent. The recovery can also be made by selective extraction with an organic solvent after acidification, neutralization or basification of the solution, depending upon the initial pH. Subsequently the extraction solvent can be evaporated to recover the product. A note of caution here is that one must check whether the product of interest decomposes further on change of pH, as it so happens frequently, as experienced by the authors.

When no condition is identified where the product is formed quantitatively into single entity, then the product can be isolated from the mixture by selective solubility based extraction, preparative TLC or preparative HPLC. Use can also be made here of normal column chromatography, medium-pressure liquid chromatography, chromatotron, flash chromatography, etc.

If the identity of the products has been previously established through sophisticated LC-MS and/or LC-NMR studies, the envisaged molecules can be synthesized, characterized and the presence confirmed through spiking in the degraded sample. The synthesis route has the advantage that it results in a much neater product than can be obtained through isolation.

6.7. Step VII: validation of SIAMs

Validation of analytical methods, in general, has been extensively covered in the ICH guidelines Q2A and Q2B [167,168], in the FDA guidance [169] and by USP [170]. There are several other reports in literature, which have reviewed the concept, either in general [171,172], or specifically the validation of spectroscopic [173], non-chromatographic [174] and chromatographic [175] methods. Numerous other investigations on development of SIAMs on different drugs also encompass validation steps, and a critical study of these reports give a fair idea on how validation can be carried out practically.

Overall, there are two stages in the validation of a SIAM. First stage is early in the development cycle when drug substance is subjected to forced decomposition studies and the SIAM is established based on the knowledge of drug degradation behavior. The main focus of validation at this stage is on establishment of specificity/selectivity, followed by other parameters like accuracy, precision, linearity, range, robustness, etc. The limits of detection and quantitation are also determined for degradation products to help in establishment of the mass balance. This validated method finds application in the analysis of stability samples of bulk drug for determination of its retest or expiry period. In the second stage, when the SIAM so developed is extended to formulations or other matrices, the emphasis gets limited to just prove the pertinence of the established validation parameters in the presence of excipients or other formulation constituents. Here only parameters of critical importance like specificity/selectivity, accuracy and precision are revalidated. If the SIAM is being developed directly for a formulation, without involving the bulk drug route, then all validation parameters are necessary to be established.

The specificity/selectivity of a SIAM can be established very simply if degradation chemistry of the drug is known and the standards of the products are available. The only effort involved then is the development of a method that separates components from a physical mixture of drug and the degradation products. At this stage, only peak purity becomes crucial. The peak purity can be established by a variety of techniques, like PDA detection, absorbance ratio method, dual wavelength ratio chromatography, second order derivative spectroscopy, spectral suppression, spectral overlay, etc. [176]. However, not all these are applicable for on-line peak purity testing. The most popular technique is the PDA analysis, the principle of which is the comparison of the spectra of the analyte peak, taken upslope, at the apex and on the downslope. If these spectra do not match then the peak is non-homogeneous. A limitation of the PDA detection for peak homogeneity testing is that this technique is not very sensitive and hence it is unlikely to detect $< 1\%$

of an interfering component in an analyte peak. Further limitation is the high cost of the detector. The normal UV HPLC detectors these days allow for simultaneous measurement at multiple wavelengths, and some of them even give output of ratio plots at two wavelengths. This technique has also been promoted for peak purity testing during development of SIAMs [177]. The technique requires critical selection of measuring wavelengths and is of limited use where the UV spectrum of the co-eluting component is unknown. The second derivative spectroscopy can also be employed to assess peak non-homogeneity, as it amplifies slight deviations from Gaussian peak shape caused by overlapping peaks. Another approach that can be employed is the collection of fractions from the peak and comparing the results with a significantly different chromatographic technique or mass spectrometry. The construction of kinetic plots during drug decomposition [177] is an additional validation step that can be used to confirm specific analysis during establishment of SIAMs.

The accuracy is usually determined by spiking the known amount of drug to either the placeboes or the formulations, and determination of percent recovery of the drug. However, a better method of determining accuracy of a SIAM is by spiking the drug in a mixture of degraded solutions [166]. As far as the precision is concerned, there are no special requirements for stability-indicating methods and the same procedure as advocated for normal assay methods can be applied.

The linearity for SIAMs should be established initially in the range of 0–100%, as the drug may fall to very low concentrations during forced decomposition studies. The final validation range, however, can be narrowed based upon the form in which the drug substance or formulation is dispensed. For example, it may vary from 80 to 120% for solid bulk drug and stable solid formulations. The range may be 50–120% in case of injections or other formulations where the drug is more prone to degradation. Validation range for the degradation products during stability studies usually should vary from 0 to 20% [171].

The detection and quantitation limits are not important for active drug substances, as their concentration is not expected to fall to such a low level in different formulations during their shelf life. However, these limits should be established for the degradation products.

Robustness can also be established for SIAMs in a similar manner as it is done for conventional methods.

7. Some critical issues concerning development of SIAMs and their validation

There are several other issues concerning development of SIAMs on which routinely the questions are asked and clarifications are sought. The important ones are discussed below.

7.1. Definition of 'Specific' and 'Selective' stability-indicating assay methods

The foremost issue is the lack of clarity on the terms used for differentiating the methods that measure quantitatively the component of interest in the sample matrix without separation, and the ones where separation is done of the drug as well all other degradation products. Hong and Shah [17] describe the former as '*stability-specific*', while the discriminating nature of the latter is described as being the combination of '*stability-indicating*' and '*stability-specific*'. Unfortunately, the term *stability-indicating* has been invariably used in the vast number of publications in literature to describe even the so-called 'stability-specific' methods. Here we would rather suggest the use of terms '*Specific stability-indicating*' and '*Selective stability-indicating*' for defining, respectively, the two types of assays. Thus 'Specific stability-indicating assay method (Specific SIAM)' can be defined as 'a method that is able to measure unequivocally the drug(s) in the presence of all degradation products, excipients and additives, expected to be present in the formulation.' The 'Selective stability-indicating assay method (Selective SIAM)' on the other hand can be defined as 'a method that is able to measure unequivocally the drug(s) and all degradation products in the presence of excipients and additives, expected to be present in the formulation.' By this definition, it means that a 'Selective SIAM' is a procedure

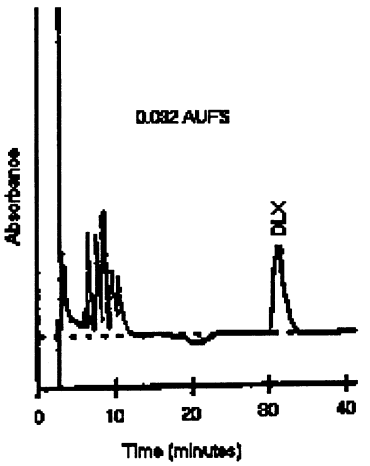


Fig. 6. Example of a 'Specific SIAM' showing degradation of dicloxacillin (DLX) to cluster of products in acid conditions. (Reprinted from Ref. [177] with permission from authors and Elsevier Science).

that is selective to the drug as well its degradation products (separates all of them qualitatively) and is also specific to all the components (measures them quantitatively).

Thus all titrimetric methods employed and reported in literature for the purpose can be classified under 'Specific SIAMs'. The UV methods also fall in this category. In these methods, despite the absence of separation, the analyte of interest is determined quantitatively and specifically.

The chromatographic methods, however, can be of both types—'Specific SIAM' as well as 'Selective SIAM'. In case of the former, the method is not fully separative to all components, but does separate the drug equivocally. This normally is a situation where efforts fail to separate degradation products when they are large in number. Two typical examples of 'Specific SIAM' are shown in Figs. 6 and 7. One of the examples is of dicloxacillin (Fig. 6), and most other penicillins show similar type of behavior [177,178]. Penicillins, as discussed earlier, are very facile undergoing degradation through a complex route (Fig. 1) resulting in multiple products under every condition. The other example (Fig. 7) is of photolytic decomposition of alpha-adrenergic receptor blocking agents where seemingly a free radical mediated reaction results in a series of products. The chromatograph shown in Fig. 5 is a typical example of a 'Selective SIAM' where all degradation products formed under all conditions are well-separated from each other and hence there is possibility of simultaneous and quantitative estimation of the drug as well as the degradation products.

7.2. Does 'Specific SIAM' also has a purpose and is acceptable?

The question thence arises, as to which of the method among the 'Specific' and 'Selective' meets the requirements of ICH and other regulatory guidelines. Certainly, it is the 'Selective SIAM',

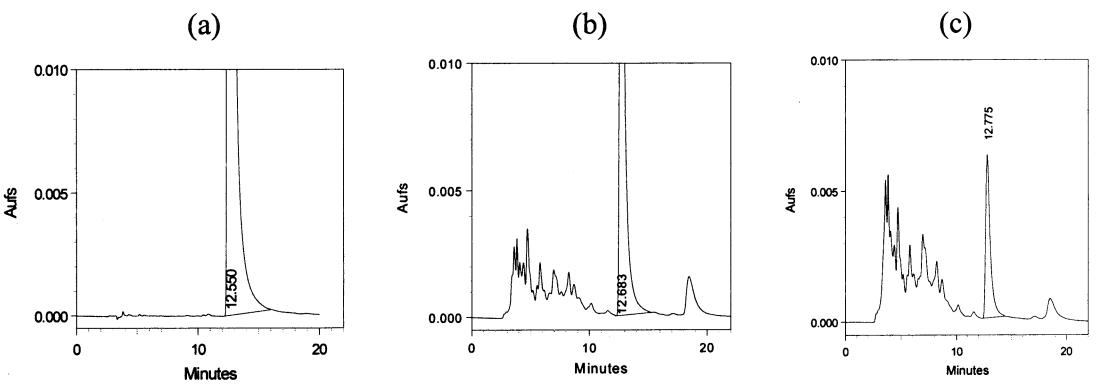


Fig. 7. Another example of 'Specific SIAM' showing formation of cluster of degradation products during photolytic decomposition of prazosin under acid conditions: (a) initial sample (b) 4 day sample (c) 8 day sample.

which the regulatory officers would love to see in the drug registration dossiers. In the opinion of the authors, the 'Selective SIAM' is that of more importance with respect to new drugs, but for old and established drugs, where significant body of information exists; the use of 'Specific SIAM' can do an equally good job during stability sample analysis. This is the reason that the pharmacopoeias, other than USP, have a policy to continue with the titrimetric and spectrophotometric analysis for the assay of drugs, while having control on important degradation products through related substance or impurity tests [179].

The authors suggest that even in the case of new drugs, the use of a 'Selective SIAM' may not be absolutely necessary in all situations. A 'Specific SIAM', if it is less cumbersome and costly, and is proved to give the same results as a 'Selective SIAM' can be used for post-approval follow-up stability testing and also for analysis of market surveillance and returned samples. This can spare chromatographic instrumentation that can be used for other routine and important applications.

The published or pharmacopoeial 'Specific SIAMs' also find use in countries where there is no regulatory requirement of a 'Selective SIAM'. However, preference should be given to those published methods where the specific method has been compared with a selective one [180].

7.3. Is it really necessary to follow the stress-testing route to develop a SIAM?

This is another dilemma that has often been expressed, particularly by the practitioners in industry. Even to the experience of the authors, there are several instances, like shown in Figs. 6 and 7, where large number of degradation products is formed during forced decomposition of drug even in one stress condition. In such situation, it might be truly difficult or impossible to develop a 'Selective SIAM' if degradation products formed under all conditions are simultaneously taken into consideration. Moreover, it has been expressed that some degradation products formed during forced decomposition are never developed in the stability samples. A typical example again is that of ornidazole where several

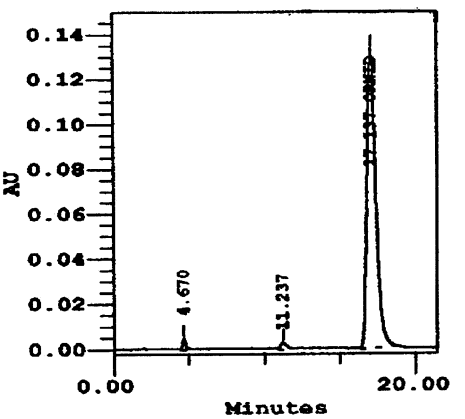


Fig. 8. Chromatogram showing the products formed after subjecting ornidazole infusion to accelerated testing.

products are formed when stressed samples in different stress conditions are mixed together (Fig. 5). Against this, the infusion samples of the drug subjected to stability tests showed only two products (Fig. 8), which are previously identified major degradation products [158]. So a tagged question is 'Should only major degradation products be targeted while developing a SIAM, instead of all degradation products formed under the ICH suggested test conditions?' Interestingly, the dilemma is well answered in the ICH guideline itself [165] where the clarification is provided in the statement '*However, it may not be necessary to examine specifically for certain degradation products if it has been demonstrated that they are not formed under accelerated or long term storage conditions.*'

Therefore, it emerges that a SIAM separating all types of possible degradation products should normally be developed through stress testing under different ICH suggested conditions. In case, however, it is not possible to develop a 'Selective SIAM' due to the complex nature of degradation, one can target for a method that takes into account degradation products only formed under accelerated and long-term storage conditions. In any case, here it would be needed to be proved through sufficient trials that separation of degradation products formed under various conditions is a difficult proposition, and that only a few or major products are found actually in stability samples of long-term and accelerated studies.

The authors at a personal level are also convinced that the regulatory approach of development of a SIAM through stress testing under variety of conditions is a sound approach. Once such a method is established and is validated, its distinct advantage would be that it could be applied to specific determination of drug and degradation products in a broad range of situations. It can eliminate the necessity of modification and revalidation each time when the method is extended from bulk drug analysis to formulation or one formulation to another. The exception might be interference due to excipients and additives, where modification and revalidation may be unavoidable.

7.4. Can formulations instead of drug substance be subjected to stress (forced decomposition) studies for development of a SIAM?

This is another interesting aspect, tagged with the above point. It is whether formulations instead of drug substance can be directly subjected to stress conditions for the development of a SIAM. In true sense, ICH guideline Q1AR and the ICH's Common Technical Document [181] suggest stress testing only of the drug substance. For drug products, however, a definition of 'Stress Testing (drug product)' is provided in Q1AR, which reads as '*Studies undertaken to assess the effect of severe conditions on the drug product. Such studies include photostability testing (see ICH Q1B) and specific testing on certain products, (e.g. metered dose inhalers, creams, emulsions, refrigerated aqueous liquid products).*' [182]. This means there is no suggestion on conduct of stress studies directly on formulations, other than photostability testing.

However, Table 7 shows that there exist a few literature reports where stability-indicating assay has been established by carrying out stress tests directly on pharmaceutical formulations. Looking into it more objectively, this approach seems to be rational for use by the manufacturers involved in production of formulations alone, albeit in specific situations. For new drugs, the information on intrinsic stability behavior of the drug substance and the stability-assay method is usually kept secret by the innovators to protect even this element from exploitation. However, there can be situations where companies in countries where innovator has not sought patent protection manufacture formulations of these drugs, by arranging the drug substance from varied sources. This is practically happening today. Therefore, in such cases if the formulation manufacturer does some sort of stress testing directly on the drug product and uses an analytical method developed on that basis, it can better cover the consumer's risk. The same can even be a case with formulations containing existing and pharmacopoeial drugs, where a bit of stress tests followed by method development can be carried out by the generic formulation manufacturers. This way at least the influence of excipients, additives and package on the degradation behavior can be encompassed directly in the analytical method.

The only hitch in implementation of this approach perhaps is that no stress conditions, except the photostability testing, have been defined for stress testing of formulations in any current guidelines. This might be due to the reason that formulations do not withstand stringent stress conditions. But the positive aspect of this approach is that the formulator can at least consider degradation products formed 'realistically' in the formulation environment during the method development. To the authors, a scientifically sound step in this regard would be to extend ICH recommended stress test conditions for drug substance to the formulations. For example, the recommended [165] dry heat stress condition of 10 °C increments above the accelerated temperature (e.g. 50, 60 °C) can be extended easily to the drug formulations. For those dosage forms, which show severe physical instability at high temperatures, the increments can be reduced to 5 °C or even lower, as applicable. The liquid formulations can be easily tested after changing the pH from low to high. Oxidative stress testing can also be done by purging oxygen or through the addition of hydrogen peroxide. Another stress condition can be 40 °C/90% RH, which can be applied to observe the effect of high humidity in case of solids, semisolids, etc.

It is suggested that scientists involved in drug stability testing research and regulatory authorities should give a serious thought to this approach, and if merit is found, then the above suggested stress conditions can be a starting point for discussions. The approach can be of specific and practical use to manufacturers outside the developed world, where there is a need to provide simple stability test solutions to help improve the quality of drug products in circulation, than available at present.

7.5. The trauma of establishment of SIAMs for combination drug formulations

A large number of formulations around the world are sold as combination preparations of two or more drugs. As the number of drugs in a formulation increases, the complexity per se increases. Therefore, the development of SIAM for formulations that contain more than one drug becomes a real tedious exercise. The effort is compounded, first due to separation of multiple drugs from one another and then from the degradation products, which might be any number. The latter are also supposed to be separated among themselves. This extent of complexity is same whether the combination formulation is a manufactured product or an extemporaneously produced preparation. The development of a 'Selective SIAM' certainly can be a nightmarish experience if it is expected that every drug substance in the formulation (some combinations may contain 6–10 drugs) is subjected to forced decomposition studies in a variety of conditions, and all solutions for all the drugs are pooled and separated on a HPLC column.

Thus for combination formulations, the approaches suggested above of targeting separation of degradation products formed only in long-term and accelerated stability test conditions or limited stressing of formulations, seem to be the best option. This issue again requires thorough discussions and should be settled among the scientists and regulatory agencies for the benefit of manufacturers.

7.6. The aspect of mass balance in development of SIAMs

The mass balance is a process of adding together the assay value and levels of degradation products to see how closely these add up to 100% of the initial value, with due consideration of the margin of analytical error. This is the definition of mass balance given in the ICH parent drug stability guideline [1]. Its establishment hence is a regulatory requirement. The mass balance is very closely linked to development of a SIAM as it acts as an approach to establish its validity. The balance would not be achieved unless all degradation products are separated well.

If a few specified and stable degradation products are formed, which can be separated easily and for which the standards are available, then the establishment of mass balance becomes an easy affair. By the use of the standards, one can easily determine the exact response factors and hence the levels of the products. However, there might be many situations where the mass balance may be difficult to establish. This can happen due to one or more of the following situations [183]:

- Formation of multiple degradation products, involving complex reaction pathways and drug-excipient interaction products
- Incomplete detection due to loss of UV chromophore or lack of universal detection
- Loss of drug/degradation products as volatiles
- Diffusive losses into or through containers
- Elution/resolution problems
- Inappropriate or unknown response factors due to lack of standards
- Errors and variability in the drug content assay

It is common for the drugs to degrade into multiple degradation products. This may be so in a single reaction condition (Figs. 6 and 7) or when samples of different stress conditions are mixed (Fig. 5). Multiple products can be formed through parallel, consecutive or chain reactions, which may occur even together. A typical example is of penicillins, where several complex reactions go together (Fig. 1). Similarly, there can be development of new products due to interaction of drugs with the excipients. Achieving separation and accounting for level of multiple degradation prod-

ucts can be an involved exercise, becoming more tedious, if there are any unstable intermediary products and when the standards of one or more are not available.

A similar complexity is added when the intermediates or final products are not shown up due to their non-UV absorbing nature. The authors observed this during acid decomposition of ornidazole, where there was no corresponding rise of peak with the loss of the drug peak [166]. It has been estimated that $>7\%$ of drugs degrade with the formation of products devoid of or with reduced UV response. At times, the lack of UV transparency of solvents and buffers also limits the detection of compounds with absorptivity only at <220 nm. In all such situations, the use of alternate refractive index (RI), evaporative light scattering, MS, NMR and IR detectors is advocated. However, their disadvantage is that they restrict the types of chromatography, for example, use of gradient elution is incompatible with RI detector, and non-volatile buffers are incompatible with evaporative light scattering and MS detectors. In exactness, no detector is available which is universally applicable. Furthermore, the response factor differences of unknowns as compared to the drug are an issue with absorbance and fluorescence detectors, though it is a smaller problem in case MS, NMR or IR detection is used.

The problems in establishment of mass balance are also encountered if the products are volatile and are lost before completion of analysis. For example, metronidazole decomposes to acetic acid and ammonia, both volatile components, on hydrolysis in alkali [157]. As a matter of fact, almost more than 20% of drugs degrade with production of volatile components. Then there might be physical losses like diffusion into plastic containers e.g. nitroglycerin, diazepam, diltiazem, benzyl alcohol, etc. are all lost to PVC bags. Even there can be loss of volatile component through glass bottles due to exchange of compounds via the closures.

There can even be situations where some of the products are strongly bound to stationary phase and do not elute or elute after very long periods. The other situation can be those where despite best efforts, the products fail to resolve completely. A typical example of the latter is given in Fig. 7. In such situations again, there can be difficulties in achieving the mass balance.

As discussed briefly earlier, the availability of reference standards of degradation products is a very important factor in establishment of mass balance. Even at the global level, the acquisition of reference standards of impurities and degradation products is a difficult proposition. Only a few and major ones are generally available, but in that case too, the costs are high and there is also a problem of long delivery periods. Making one's own standards is again a difficult exercise, requiring facilities for confirmation of structure and purity. This is a time-consuming task, and also a costly affair, requiring expert manpower and sophisticated analytical instrumentation. The lack of standards results in inappropriate or unknown response factors, acting again as a bottleneck in the establishment of mass balance.

Finally, the mass balance may not also be established due to reasons of errors and variability in the drug content assay. The design of analytical method, the calculation approaches, etc. have to be absolute to get the correct results. There might be other issues, like area percent methods have lower sensitivity and overestimate minor components. On the other hand, external standard methods have higher sensitivity and lower errors, but are much more complex to perform.

So there might be a number of situations where one may not be able to attain a mass balance. Keeping all these difficulties into view, a change in stance is taking place at the regulatory level on the mass balance requirements. The changes that have been brought in are as follows:

(i) The original ICH guideline Q1A contained along with the definition of mass balance (already given above), an additional paragraph '*This concept is a useful scientific guide for evaluating data but it is not achievable in all circumstances. The focus may instead be on assuring the specificity of the assay, the completeness of the investigation of routes of degradation, and the use, if necessary, of identified degradants as indicators of the extent of degradation via particular mechanisms.*' [184]. This paragraph has been removed in the revision.

(ii) The original statement in the text of ICH guideline Q1A [185] reads under both 'drug substance' and 'drug product' as '*Any evaluation should consider not only the assay*, ***but the levels of degradation products*** *and other appropriate attributes*.' Under drug product, the following additional statement exists: '*Where appropriate, attention should be paid to reviewing the adequacy of the mass balance, different stability and degradation performance*'. The first sentence has been changed in the revised guideline Q1AR under the drug product to '*Any evaluation should consider not only the assay* ***but also the degradation products*** *and other appropriate attributes*.' [186]. Evidently, there is exclusion of emphasis on 'levels' of the degradation products in the revision for 'drug product'.

(iii) Changes have also been made with respect to mass balance in ICH guideline Q3A entitled 'Impurities in New Drug Substances' where the whole paragraph '*A summation of assay value and impurity levels generally may be used to obtain mass balance for the test sample. The mass balance need not add to exactly 100% because of the analytical error associated with each analytical procedure. The summation of the impurity level plus the assay value may be misleading, e.g. when the assay procedure is non-specific, (e.g. potentiometric titrimetry) and the impurity level is relatively high*.' has been removed [187].

Therefore, the authors can only suggest that if the establishment of mass balance becomes possible during development of a SIAM, it is very fine. Otherwise, it must be kept in mind that a method may be valid for other parameters even when the mass balance is not observed. Hence any efforts towards compulsory establishment of mass balance should not be at the altar of sacrificing basic characteristics, like specificity, precision, ruggedness, etc. The problems encountered in establishment of mass balance must be clearly defined and indicated in the pertinent part in the registration application. The truthful projection of the difficulty can be helpful, as the possible failures in meeting the mass balance requirement have been increasingly realized by the regulators.

7.7. Are pharmacopoeial methods stability-indicating?

This is again a general dilemma. The authors also found a lot of question–answers taking place on this issue in the discussion groups spread over the web. To get an answer to this question, one has to really understand the structure of a pharmacopoeial monograph. As has been discussed briefly above also, the compendial monographs usually control critical decomposition products through separate tests for related substances and impurities, and the purity tests. It is for this reason that assay methods prescribed in the monographs have classically and primarily been designed to be stability-specific ('Specific SIAM') by nature, and not meant to be selective to each decomposition product and other constituents in the drug substance or formulation. The whole class of titrimetric methods is a simple example here as in these methods the target is the drug alone. Specific examples, cited during discussion on the web, are the assay for calcium gluconate that is based on the determination of calcium by complexation with EDTA. This method does not tell anything about the possible chemical transformations of gluconate, hence can not be used for the analysis of stability samples. The principle behind the polarimetric assay of dextrose is based on proportionality of optical rotation to the potency. This assay can not be considered stability-indicating, as the optical rotation being an additive property, remains equal to the original value even when dextrose has degraded to its constituent sugars.

Therefore, it can be said, that pharmacopoeial methods historically were not 'Selective SIAMs' and perhaps, as exemplified above, there are several, which are still not. However, with the advent of technology allowing resolution of an article into its components and introduction of ICH guideline Q1A, in which there is a clear mandate for simultaneous analysis of degradation products, the situation has changed over

the period. The ICH guideline was printed in USP 23 and made official, though it has been removed in USP 24 due to restraint on printed pages. The USP contains a large number (there were > 2000 in USP 23) of assays and tests based on HPLC and several of them supposedly are 'selective' by nature. USP also defines Category II analytical methods that are meant for determination of impurities in bulk drug substances or degradation compounds in finished pharmaceutical products and provides data elements required for validation of these. Thus USP fully recognizes the necessity of compendial methods to be 'Selective'.

Interestingly, the 'Monograph Development: Guidance to Manufacturers' in British Pharmacopoeia, prescribes that '*For bulk drug substances, it has been BP policy generally to use a robust and precise method of assay (such as titration) rather than a specific, but sometimes less precise, stability-indicating method (such as liquid chromatography). Wherever possible, control of potential impurities is provided separately by means of specific impurity tests. It is appreciated, however, that a manufacturer may use, and therefore propose, a chromatographic method for both related substances and assay. In such circumstances, each case is judged on its merits on the basis of the data provided, which must relate to validated methods.*' Separately it is stated that '*The method of assay will not necessarily be that used for the bulk drug substance. For formulations a specific, stability-indicating method is preferred*' [188]. Evidently, there is a slightly different approach in BP as compared to USP, although there is an endeavor for shifting to 'Selective' methods.

The authors, however, feel that even the compendial chromatographic methods should be validated for their 'selectivity' by the end user, as there can be cases where a potentially stability-indicating pharmacopoeial assay does not prove to be so when applied to actual formulations. This is in line with the validation requirements given in the ICH/WHO guideline on GMP of APIs, which clearly states that '*The suitability of all testing methods used should nonetheless be verified under actual conditions of use and documented*' [189].

8. Commercial availability of standards of degradation products

The success of establishment of validated SIAM and also the mass balance depends much on the availability of standards of degradation products. Therefore, a brief discussion on the commercial sources from where one can acquire them would be pertinent here. The standards for old and established degradation products controlled by pharmacopoeial monographs can be procured from the respective pharmacopoeial authorities (www.pheur.org; www.usp.org; www.promochem.com). Also, there are other national and international organizations that supply the standards, and help can be taken in this regard from a comprehensive list of globally available standards published annually by the WHO (WHO/EDM/QSM/2001.2). Apart from this, there are independent agencies that also supply these standards, and information on such sources can be assessed through internet, using search engines like, Altavista, Yahoo, Google, etc. It may be pertinent to add here that the author's lab at NIPER specializes in separation, synthesis and supply of degradation products and those interested can check the institute's web site (www.niper.nic.in) for an updated list.

9. The emerging techniques for analysis of stability samples

As discussed under the instrumental methods employed in literature reports and elsewhere in the text, there is an increasing trend in recent times on involvement of hyphenated techniques (GC-MS, LC-MS or LC-MS-MS, CE-MS, LC-NMR, etc.) at various stages in development of SIAMs. Their use is picking up due to easy availability of bench-top instrumentation and their distinct advantages, like versatility; sensitivity; possibility of profiling, substructural analysis and rapid selective quantitative determination of targeted products even in mixtures. The only limitation yet is the heavy cost of instrumentation, due to which their use is not common and spread

worldwide, like simple GC, HPLC, CE, or NMR systems. These sophisticated techniques as of today are being used mainly for the purpose of monitoring, characterization and identification of impurities, degradation products, metabolites, etc. However, there is a good scope of their use in routine quantitative analysis of stability samples, as their cost per analysis tends to be much lower than conventional techniques. Therefore, developments in the applications and quantitative use of these techniques must be followed and watched with interest.

Other than hyphenated techniques, Fourier-transform near-infrared (FT-NIR) spectroscopy is another emerging technique, which holds lot of promise. The instrument works on the principle of Kubelka–Munk function, and determines the fragment of light reflected from the sample depending upon scattering and absorption of light [190]. It has a distinct advantage that it allows analysis of drugs directly in the dosage forms, without even the need of sample preparation, thus eliminating the use of extraction solvents and hassles involved in their disposal. Being a non-destructive technique, it has capability to analyze drug in tablets, powders, solids, liquids or pastes. Even non-homogenous samples such as multi-layered, coated or cored tablets can be analyzed reliably. It allows measurement in sealed glass containers, and even of sterilized samples without opening. As such, only a small amount of sample is needed for obtaining useful test results. The reproducibility of the test result is ensured. The technique is very fast, with sample time reduced to around 5 s, allowing large number of samples to be analyzed within a short time. Thus the major benefit is that quality standards are maintained while costs and efforts are reduced. It is envisaged that this technique will be very fruitfully employed for the analysis of stability samples, when there are sufficient spectral differences between the drug and the degradation products. Though technically, the assay using FT-NIR would only be 'Specific', but it can be employed selectively if there are only few degradation products formed on storage, which differ structurally among themselves and also from the drug.

10. Making use of computer simulation in development and optimization of SIAMs

As must be realized from the above discussion, the process of development of SIAMs by HPLC is a time consuming and difficult exercise. In general also, there are a large number of interdependent parameters, which exist in the practice of HPLC and the consequent requirement to study these parameters during method development through multiple chromatographic runs makes the situation very difficult overall [191].

A good strategy for development of a SIAM, like any other HPLC method, should require only as many experimental runs as are necessary to achieve the desired result. The manual approach, involving manipulation of experimental variables until the desired separation has been achieved, provides a good understanding of the principles and theory involved and the interaction of the various variables. But unfortunately it is a slow, time consuming and a potentially expensive exercise. These limitations of the manual HPLC method development approach have led to an increased use of computers-based expert systems. These can be used to automate various phases of HPLC process or fit the retention data to various models in order to find the best conditions for a particular separation. The advantages of computer simulation over manual method development are: (i) the computer simulation of chromatographic separations avoids most of the experimental work to be done in chromatographic method development and optimization. Consequently, the cost and time spent in optimization process are dramatically reduced, (ii) once the simulation process begins, it can continue in an unattended manner, and (iii) only the computer is blocked during the optimization process and not the chromatograph, which can be used for other purposes.

Table 9 lists various available expert systems based on their capabilities. More information can be obtained from numerous reports in literature on the subject [192–195].

There are yet not many publications, which have indicated the use of expert systems in the development of SIAMs. However, there are all

the chances that they are being used in the development laboratories in the industry. Due to their distinct advantages, it is worth giving a try and those who are not exploiting them at present, must explore the utility in their set-ups.

11. The SIAM requirements for stability study of biotechnological products

Biotechnological products also undergo degradation during storage. A variety of degradation products arise resulting from deamidation, oxidation, sulfoxidation, aggregation or fragmentation. No single stability-indicating assay or parameter is available that profiles the stability characteristics of biotechnological products, unlike those of chemical drugs. Hence it is a requirement in the ICH guideline Q5C on Stability Testing of Biotechnological Products that the manufacturer should propose a 'stability-indicating profile' that provides assurance that the changes in identity, purity and potency of the product will be detected [4]. Tests for stability should cover those features, which are likely to change during storage and it is required that the tests employed should be product-specific. There are a large number of publications where the use of stability-indicating methodology to determine the shelf life of different biotechnological products has been reported [196–198].

12. Conclusions

As can be seen from the plenty of the examples given in the tables in the text above, the stability-indicating assays have been developed for a large number of drugs for last several decades, starting almost from 1960s. But most of them unfortunately fail to meet the current regulatory requirements of separation and analysis of individual degradation products. Furthermore, there is little guidance provided in the literature on how to establish true 'Selective' stability-indicating methods. In that respect it is hoped that the discussion provided above on development and validation of SIAMs and on several connected issues would be of general and wide interest. It is, however, cautioned that the opinions expressed are purely personal to the authors and do not represent thinking of the regulatory agencies.

Table 9
Selected examples of software employed in HPLC method development

Software	Properties
DryLab	Allows change of one retention variable at a time and predicts separation as a function of that variable. Also predicts separation for any gradient conditions, based upon data for change in gradients
DryLab, ENHANCER	Predicts separation for different chromatographic conditions (column dimensions, particle size, flow rate, etc.)
ICOS, DIAMOND	Allows change of one or more variables at a time and predicts separation as a function of those variables
PESOS	Based on change of one or more conditions, examines experimental chromatograms for best separation
PRISMA model	Software based on correlation of solvent strength and mobile phase selectivity
ELUEX, CHROMSWORD	Expert systems to predict best initial separation conditions on the basis of molecular structure of the sample components
HPLC-METABOLEXPERT, ProDigest-LC, CHROMDREAM	Special purpose programs

The list is only representative and not comprehensive.

References

[1] ICH, Stability Testing of New Drug Substances and Products. International Conference on Harmonisation, IFPMA, Geneva, 1993.
[2] ICH, Impurities in New Drug Products. International Conference on Harmonisation, IFPMA, Geneva, 1996.
[3] ICH, Specifications: Test Procedures and Acceptance Criteria for New Drug Substances and New Drug Products: Chemical Substances. International Conference on Harmonisation, IFPMA, Geneva, 1999.
[4] ICH, Quality of Biotechnological Products: Stability Testing of Biotechnological/Biological Products, International Conference on Harmonisation, IFPMA, Geneva, 1995.
[5] FDA, Guideline for Submitting Documentation for the Stability of Human Drugs and Biologics. Food and Drug Administration, Rockville, MD, 1987.
[6] FDA, Guidance for Industry: Stability Testing of Drug Substances and Drug Products (Draft guidance), Food and Drug Administration, Rockville, MD, 1998.
[7] WHO, Guidelines for Stability Testing of Pharmaceutical Products Containing Well Established Drug Substances in Conventional Dosage Forms, in WHO Expert Committee on Specifications for Pharmaceutical Preparations. Technical Report Series 863, World Health Organization, Geneva, 1996, pp. 65–79.
[8] CPMP, Note for Guidance on Stability Testing of Existing Active Substances and Related Finished Products. Committee for Proprietary Medicinal Products, EMEA, London, 1998.
[9] TPD, Stability Testing of Existing Drug Substances and Products. Therapeutic Products Directorate, Ottawa, 1997.
[10] The United States Pharmacopeia, 24th Revision, Asian Edition, United States Pharmacopeial Convention, Inc., Rockville, MD, 2000.
[11] ICH, Good Manufacturing Practices for Active Pharmaceutical Ingredients. International Conference on Harmonisation, IFPMA, Geneva, 2000.
[12] L. Chafetz, J. Pharm. Sci. 60 (1971) 335–344.
[13] V. Kumar, N. Sunder, East. Pharm. 38 (8) (1989) 47–50.
[14] C. Ho, G.-L. Chen, J. Food Drug Anal. 4 (1996) 271–292.
[15] C. Ho, G.-L. Chen, J. Food Drug Anal. 5 (1997) 1–24.
[16] Q.A. Xu, L.A. Trissel (Eds.), Stability-Indicating HPLC Methods for Drug Analysis, American Pharmaceutical Association, Washington, 1999.
[17] D.D. Hong, M. Shah, Development and Validation of HPLC stability-indicating assays, In: Drug Stability: Principles and Practices, J.T. Carstensen, C.T. Rhodes (Eds.) Marcel Dekker, New York, 2000, pp. 329–384.
[18] T. Radhakrishna, D.S. Rao, K. Vyas, G.O. Reddy, J. Pharm. Biomed. Anal. 22 (2000) 641–650.
[19] Y. Shah, S. Joshi, K.C. Jindal, S. Khanna, Drug Dev. Ind. Pharm. 20 (1994) 85–91.
[20] N. Maron, E.A. Cristi, A.A. Ramos, J. Pharm. Sci. 77 (1988) 638–639.
[21] D.E. Williamson, J. Pharm. Sci. 65 (1976) 138–140.
[22] C. Stubbs, I. Kanfer, Int. J. Pharm. 63 (1990) 113–119.
[23] R.J. Yarwood, W.D. Moore, J.H. Collett, J. Pharm. Sci. 74 (1985) 220–223.
[24] H. Fabre, N. Hussam-Eddine, B. Mandrou, J. Pharm. Sci. 73 (1984) 1706–1709.
[25] T. Radhakrishna, C.L. Narayana, D.S. Rao, K. Vyas, G.O. Reddy, J. Pharm. Biomed. Anal. 22 (2000) 627–639.
[26] R.M. Rao, Y.M. Rao, A.H. Shah, J. Pharm. Biomed. Anal. 20 (1999) 717–722.
[27] S. Hanna, M. Rosen, L. Rasero, L. Lachman, J. Pharm. Sci. 66 (1977) 123–124.
[28] A.E. Bretnall, M.M. Hodgkinson, G.S. Clarke, J. Pharm. Biomed. Anal. 15 (1997) 1071–1075.
[29] J.T. Strode III, L.T. Taylor, A.L. Howard, D. Ip, M.A. Brooks, J. Pharm. Biomed. Anal. 12 (1994) 1003–1014.
[30] M. Emery, J. Kowtko, J. Pharm. Sci. 68 (1979) 1185–1187.
[31] M.B. Evans, P.A. Haywood, D. Johnson, M. Martin-Smith, G. Munro, J.C. Wahlich, J. Pharm. Biomed. Anal. 7 (1989) 1–22.
[32] A.A. Fatmi, E.A. Hickson, J. Pharm. Sci. 77 (1988) 87–89.
[33] K. Brightman, G. Finlay, I. Jarvis, T. Knowlton, C.T. Manktelow, J. Pharm. Biomed. Anal. 20 (1999) 439–447.
[34] S.P. Puthli, P.R. Vavia, J. Pharm. Biomed. Anal. 22 (2000) 673–677.
[35] J.P. Salo, H. Salomies, J. Pharm. Biomed. Anal. 14 (1996) 1261–1266.
[36] A.J. Visalli, D.M. Patel, N.H. Reavey-Cantwell, J. Pharm. Sci. 65 (1976) 1686–1688.
[37] A. Shafaati, B.J. Clark, Drug Dev. Ind. Pharm. 26 (2000) 267–273.
[38] H.R. Bhagat, H.N. Bhargava, D.A. Williams, J. Pharm. Biomed. Anal. 7 (1989) 441–446.
[39] A.-A.M. Wahbi, A.-F.M. El-Walily, E.M. Hassan, F.G. Saliman, A. El-Gendi, J. Pharm. Biomed. Anal. 13 (1995) 777–784.
[40] A. El-Gindy, A. Ashour, L. Abdel-Fattah, M.M. Shabana, J. Pharm. Biomed. Anal. 25 (2001) 913–922.
[41] S.Z. El Khateeb, S.A. Abdel Razek, M.M. Amer, J. Pharm. Biomed. Anal. 17 (1998) 829–840.
[42] F.A. El-Yazbi, H.H. Abdine, R.A. Shaalan, J. Pharm. Biomed. Anal. 19 (1999) 819–827.
[43] Z.R. Zaidi, F.J. Sena, C.P. Basilio, J. Pharm. Sci. 71 (1982) 997–999.
[44] V.D. Gupta, J. Pharm. Sci. 72 (1983) 695–697.
[45] V.D. Gupta, J. Pharm. Sci. 73 (1984) 1331–1333.
[46] F.A. El-Yazbi, H.H. Abdine, R.A. Shaalan, J. Pharm. Biomed. Anal. 20 (1999) 343–350.
[47] J. Frank, L. Chafetz, J. Pharm. Sci. 66 (1977) 439.
[48] S. Rubnov, D. Levy, H. Schneider, J. Pharm. Biomed. Anal. 18 (1999) 939–945.
[49] N.A. El Ragehy, S.S. Abbas, S.Z. El-Khateeb, J. Pharm. Biomed. Anal. 25 (2001) 143–151.

[50] V. Andrisano, R. Gotti, A. Leoni, V. Cavrini, J. Pharm. Biomed. Anal. 21 (1999) 851–857.
[51] E.A. Gad Kariem, M.A. Abounassif, M.E. Hagga, H.A. Al-Khamees, J. Pharm. Biomed. Anal. 23 (2000) 413–420.
[52] E.M. Abdel-Moety, O.A. Al-Deeb, Eur. J. Pharm. Sci. 5 (1997) 1–5.
[53] V.B. Patravale, V.B. Nair, S.P. Gore, J. Pharm. Biomed. Anal. 23 (2000) 623–627.
[54] A.E.-A. El-Bayoumi, A. El-Shanawany, M.E. El-Sadek, A.A. El-Sattar, J. Pharm. Biomed. Anal. 21 (1999) 867–873.
[55] H. Bartsch, A. Eiper, H. Kopelent-Frank, J. Pharm. Biomed. Anal. 20 (1999) 531–541.
[56] D. Mahalaxmi, M.M. Samarth, H.S. Shiravadekar, N.M. Sanghavi, Drug Dev. Ind. Pharm. 22 (1996) 1037–1039.
[57] M.Y. Lye, K.L. Yow, L.Y. Lim, S.Y. Chan, E. Chan, P.C. Ho, Am. J. Health-Syst. Pharm. 54 (1997) 2483–2487.
[58] V.D. Gupta, J. Parasrampuria, Drug Dev. Ind. Pharm. 13 (1987) 2231–2238.
[59] S.O. Thoppil, P.D. Amin, J. Pharm. Biomed. Anal. 22 (2000) 699–703.
[60] M.V. Padval, H.N. Bhargava, J. Pharm. Biomed. Anal. 11 (1993) 1033–1036.
[61] M. Mathew, V.D. Gupta, R.E. Bailey, Drug Dev. Ind. Pharm. 21 (1995) 965–971.
[62] S. Mittal, K.S. Alexander, D. Dollimore, Drug Dev. Ind. Pharm. 26 (2000) 1059–1065.
[63] V.B. Patravale, S. D'Souza, Y. Narkar, J. Pharm. Biomed. Anal. 25 (2001) 685–688.
[64] J.J. Bergh, J.C. Breytenbach, J. Chromatogr. 387 (1987) 528–531.
[65] A.P. Cheung, E. Struble, N. Nguyen, P. Liu, J. Pharm. Biomed. Anal. 24 (2001) 957–966.
[66] C.A. Janicki, W.D. Walkling, R.H. Erlich, C.A. Bainbridge, J. Pharm. Sci. 70 (1981) 778–780.
[67] A.-B. Wu, C.-Y. Chen, S.-D. Chu, Y.-C. Tsai, F.-A. Chen, J. Chromatogr. Sci. 39 (2001) 7–11.
[68] D.N. Tipre, P.R. Vavia, J. Pharm. Biomed. Anal. 24 (2001) 705–714.
[69] P.N. Kotiyan, P.R. Vavia, J. Pharm. Biomed. Anal. 22 (2000) 667–671.
[70] N. Maron, G. Wright, J. Pharm. Biomed. Anal. 8 (1990) 101–105.
[71] C.J. Hinds, D. Johnston, C.J. Thompson, Drug Dev. Ind. Pharm. 10 (1984) 983–990.
[72] D.B. Wiest, W.A. Maish, S.S. Garner, G.M. El-Chaar, Am. J. Hosp. Pharm. 48 (1991) 2430–2433.
[73] S.L. Daniels, A.J. Vanderwielen, J. Pharm. Sci. 70 (1981) 211–215.
[74] S.P. Kulkarni, P.D. Amin, J. Pharm. Biomed. Anal. 23 (2000) 983–987.
[75] J.C. Spell, J.T. Stewart, J. Pharm. Biomed. Anal. 18 (1998) 453–460.
[76] J.H. Bridle, M.T. Brimble, Drug Dev. Ind. Pharm. 19 (1993) 371–381.
[77] Y.-C. Lee, D.M. Baaske, A.S. Alam, J. Pharm. Sci. 73 (1984) 1660–1661.
[78] A.L. Lagu, R. Young, E.J. McGonigle, P.A. Lane, J. Pharm. Sci. 71 (1982) 85–88.
[79] C.M. Shearer, N.J. DeAngelis, J. Pharm. Sci. 68 (1979) 1010–1012.
[80] C.A. Janicki, K.F. Daly, J. Pharm. Sci. 69 (1980) 147–149.
[81] G. Caviglioli, B. Parodi, S. Cafaggi, G. Bignardi, G. Romussi, Drug Dev. Ind. Pharm. 20 (1994) 2395–2408.
[82] S.O. Thoppil, P.D. Amin, J. Pharm. Biomed. Anal. 25 (2001) 191–195.
[83] S.O. Thoppil, R.M. Cardoza, P.D. Amin, J. Pharm. Biomed. Anal. 25 (2001) 15–20.
[84] J. Lambropoulos, G.A. Spanos, N.V. Lazaridis, T.S. Ingallinera, V.K. Rodriguez, J. Pharm. Biomed. Anal. 20 (1999) 705–716.
[85] R.L. Garnick, G.F. Burt, D.A. Long, J.W. Bastian, J.P. Aldred, J. Pharm. Sci. 73 (1984) 75–77.
[86] M.M. Al-Omari, M.K. Abdelah, A.A. Badwan, A.M. Jaber, J. Pharm. Biomed. Anal. 25 (2001) 893–902.
[87] N. Daraghmeh, M. Al-Omari, A.A. Badwan, A.M. Jaber, J. Pharm. Biomed. Anal. 25 (2001) 483–492.
[88] D.M. Baaske, J.F. Demay, C.A. Latona, S. Mirmira, K.W. Sigvardson, Am. J. Health-Syst. Pharm. 53 (1996) 1701–1705.
[89] J. Lambropoulos, G.A. Spanos, N.V. Lazaridis, J. Pharm. Biomed. Anal. 19 (1999) 793–802.
[90] A. Segall, M. Vitale, V. Perez, F. Hormaechea, M. Palacios, M.T. Pizzorno, Drug Dev. Ind. Pharm. 26 (2000) 867–872.
[91] A. Khedr, A. Sakr, J. Chromatogr. Sci. 37 (1999) 462–468.
[92] A.K. Hunt-Fugate, C.K. Hennessey, C.M. Kazarian, Am. J. Hosp. Pharm. 50 (1993) 1186–1187.
[93] S.M. Wintermeyer, M.C. Nahata, Am. J. Health-Syst. Pharm. 53 (1996) 407–409.
[94] D.W. Boulton, J.P. Fawcett, D.J. Woods, Am. J. Health-Syst. Pharm. 53 (1996) 1157–1161.
[95] P.J. Simms, R.W. Towne, C.S. Gross, R.E. Miller, J. Pharm. Biomed. Anal. 17 (1998) 841–849.
[96] N.H. Anaizi, C.F. Swenson, P.J. Dentinger, Am. J. Health-Syst. Pharm. 56 (1999) 1738–1741.
[97] S.M. Abdel-Rahman, M.C. Nahata, Am. J. Health-Syst. Pharm. 54 (1997) 1301–1303.
[98] K.C. Chung, A. Chin, M.A. Gill, Am. J. Health-Syst. Pharm. 52 (1995) 1541–1543.
[99] D.L. Dunn, W.J. Jones, E.D. Dorsey, J. Pharm. Sci. 72 (1983) 277–280.
[100] T. Dine, F. Khalfi, B. Gressier, M. Luyckx, C. Brunet, L. Ballester, F. Goudaliez, J. Kablan, M. Cazin, J.C. Cazin, J. Pharm. Biomed. Anal. 18 (1998) 373–381.
[101] E.R. Montgomery, A.L. Edmanson, S.C. Cook, P.K. Hovsepian, J. Pharm. Biomed. Anal. 25 (2001) 267–284.

[102] M. Kumar, S.K. Singhal, A. Singh, J. Pharm. Biomed. Anal. 25 (2001) 9–14.
[103] E.R. Montgomery, S. Taylor, J. Segretario, E. Engler, D. Sebastian, J. Pharm. Biomed. Anal. 15 (1996) 73–82.
[104] Z. Zhao, Q. Wang, E.W. Tsai, X.-Z. Qin, D. Ip, J. Pharm. Biomed. Anal. 20 (1999) 129–136.
[105] C. Mannucci, J. Bertini, A. Cocchini, A. Perico, F. Salvagnini, A. Triolo, J. Pharm. Sci. 82 (1993) 367–370.
[106] J. Bauer, S. Krogh, J. Pharm. Sci. 72 (1983) 1347–1349.
[107] J. Kirschbaum, S. Perlman, J. Pharm. Sci. 73 (1984) 686–687.
[108] L. Elrod, R.D. Cox, A.C. Plasz, J. Pharm. Sci. 16 (1982) 161–166.
[109] G.N. Menon, B.J. Norris, J. Pharm. Sci. 70 (1981) 697–698.
[110] M.G. Gebauer, A.F. McClure, T.L. Vlahakis, Int. J. Pharm. 223 (2001) 49–54.
[111] F. Belal, I.A. Al-Zaagi, E.A. Gadkariem, M.A. Abounassif, J. Pharm. Biomed. Anal. 24 (2001) 335–342.
[112] S.D. Pearson, L.A. Trissel, Am. J. Hosp. Pharm. 50 (1993) 698–702.
[113] G.N. Menon, L.B. White, J. Pharm. Sci. 70 (1981) 1083–1085.
[114] S.N. Makhija, P.R. Vavia, J. Pharm. Biomed. Anal. 25 (2001) 663–667.
[115] H. Zhang, P. Wang, M.G. Bartlett, J.T. Stewart, J. Pharm. Biomed. Anal. 16 (1998) 1241–1249.
[116] G.W. Schieffer, D.E. Hughes, J. Pharm. Sci. 72 (1983) 55–59.
[117] J.-P. Burm, S.S. Jhee, A. Chin, Y.S.K. Moon, E. Jeong, L. Nii, J.L. Fox, M.A. Gill, Am. J. Hosp. Pharm. 51 (1994) 1201–1204.
[118] L.V. Allen Jr., M.A. Erickson III, Am. J. Health-Syst. Pharm. 53 (1996) 1944–1949.
[119] L.V. Allen Jr., M.A. Erickson III, Am. J. Health-Syst. Pharm. 53 (1996) 2179–2184.
[120] W.E. Wallo, A. D'adamo, J. Pharm. Sci. 71 (1982) 1115–1118.
[121] P.A. Lane, D.O. Mayberry, R.W. Young, J. Pharm. Sci. 76 (1987) 44–47.
[122] V.G. Dabbene, M.C. Brinon, M.M. de Bertorello, J. Pharm. Sci. 83 (1994) 1617–1621.
[123] R.J. Forsyth, D.P. Ip, J. Pharm. Biomed. Anal. 12 (1994) 1243–1248.
[124] M.I. Walash, A.M. el-Brashy, M.A. Sultan, Acta Pharm. Hung. 64 (1994) 5–8.
[125] E.F. Khamis, M. Abdel-Hamid, E.M. Hassan, A. Eshra, M.A. Elsayed, J. Clin. Pharm. Ther. 18 (1993) 97–101.
[126] J. Mezei, J. Pap, Acta Pharm. Hung. 55 (1985) 134–137.
[127] H. Fabre, N.H. Eddine, F. Bressolle, B. Mandrou, Analyst 107 (1982) 61–66.
[128] H. Fabre, N. Hussam-Eddine, J. Pharm. Pharmacol. 34 (1982) 425–428.
[129] S.K. Baveja, S. Singh, J. Chromatogr. 396 (1987) 337–344.
[130] P. Gallagher, S. Jones, Int. J. Pharm. Practice 5 (1997) 101–104.
[131] J.J. Bergh, A.P. Lotter, Drug Dev. Ind. Pharm. 10 (1984) 127–136.
[132] S. Stavchansky, P. Wu, J.E. Wallace, Drug Dev. Ind. Pharm. 9 (1983) 989–998.
[133] L.K. Revelle, D.A. d'Avignon, J.C. Reepmeyer, R.C. Zerfing, J. AOAC Int. 78 (1995) 353–358.
[134] G.M. Hanna, C.A. Lau-Cam, J. AOAC Int. 76 (1993) 526–530.
[135] K.L. House, A.R. Garber, R.B. Dunlap, J.D. Odom, D. Hilvert, Biochemistry 32 (1993) 3468–3473.
[136] H.K. Hansen, S.H. Hansen, M. Kraunse, G.M. Petersen, Eur. J. Pharm. Sci. 9 (1999) 41–46.
[137] K.D. Altria, M.A. Kelly, B.J. Clark, Trends Anal. Chem. 17 (1998) 204–214.
[138] A.K. Lalloo, I. Kanfer, J. Chromatogr. B Biomed. Sci. Appl. 704 (1997) 343–350.
[139] J.A. Visconti, H.G. Anderson, R.D. Williams, C.R. Gibson, ASHP Midyear Clinical Meeting 1996 (1996) P–185R.
[140] M.E. Abdel-Hamid, I.O. Edafiogho, H.M. Hamza, J. Pharm. Biomed. Anal. 27 (2002) 225–234.
[141] W. Feng, H. Liu, G. Chen, R. Malchow, F. Bennett, E. Lin, B. Pramanik, T.M. Chan, J. Pharm. Biomed. Anal. 25 (2001) 545–557.
[142] X. Xu, M.G. Bartlett, J.T. Stewart, J. Pharm. Biomed. Anal. 26 (2001) 367–377.
[143] M. Abdel-Hamid, L. Novotny, H. Hamza, J. Pharm. Biomed. Anal. 22 (2000) 745–755.
[144] J. Ermer, M. Vogel, Biomed. Chromatogr. 14 (2000) 373–383.
[145] M.J. Lovdahl, S.R. Priebe, J. Pharm. Biomed. Anal. 15 (2000) 521–534.
[146] Y. Wu, Biomed. Chromatogr. 14 (2000) 384–396.
[147] S.X. Peng, B. Borah, R.L. Dobson, Y.D. Liu, S. Pikul, J. Pharm. Biomed. Anal. 20 (1999) 75–89.
[148] K.J. Volk, S.E. Klohr, R.A. Rourick, E.H. Kerns, M.S. Lee, J. Pharm. Biomed. Anal. 14 (1996) 1663–1674.
[149] S. Takeda, Y. Tanaka, Y. Nishimura, M. Yamane, Z. Siroma, S. Wakida, J. Chromatogr. A 853 (1999) 503–509.
[150] K. Ensing, T. de Boer, N. Schreuder, R.A. de Zeeuw, J. Chromatogr. B Biomed. Sci. Appl. 727 (1999) 53–61.
[151] K.A. Connors, G.L. Amidon, V.J. Stella (Eds.), Chemical Stability of Pharmaceuticals, Wiley, New York, 1986.
[152] S.W. Hovorka, C. Schoneich, J. Pharm. Sci. 90 (2001) 253–269.
[153] N.H. Anderson (Ed.), Photostability Testing: Design and Interpretation of Tests on Drug Substances and Dosage Forms, Taylor and Francis, London, 1996.
[154] H. Bundgaard, C. Larsen, J. Pharm. Biomed. Anal. 1 (1983) 29–37.
[155] B.J. Wilkins, G.J. Cainsford, D.E. Moore, J. Chem. Soc., Perkin Trans. I (1987) 1817–1820.
[156] A.A. Moustafa, L.I. Bibawy, Spectrosc. Lett. 32 (1999) 1073–1098.
[157] S.K. Baveja, H.K. Khosla, Indian J. Technol. 13 (1975) 528.

[158] J.R. Valdes Santurio, J.A. Martinez Perez, B. Lopez Pelaez, C. Martinez Manchado, STP Pharma. Sci. 5 (1995) 391–395.
[159] K. Florey (Ed.), Analytical Profiles of Drug Substances, Academic Press, London.
[160] L.R. Synder, J.J. Kirkland, J.L. Glajch (Eds.), Practical HPLC Method Development, Wiley, New York, 1997, p. 295.
[161] S.K. Baveja, S. Singh, Indian Drugs 25 (1988) 286–290.
[162] J.T. Carstensen, C.T. Rhodes (Eds.), Drug Stability: Principles and Practices, Marcel Dekker, New York, 2000, p. 339.
[163] Available from http://www.cmcissues.com/stability/stress_stability.htm.
[164] S. Singh, M. Bakshi, Pharm. Tech. On-line 24 (2000) 1–14.
[165] ICH, Stability Testing of New Drug Substances and Products. International Conference on Harmonisation, IFPMA, Geneva, 2000, p. 2.
[166] M. Bakshi, B. Singh, A. Singh, S. Singh, J. Pharm. Biomed. Anal. 26 (2001) 891–897.
[167] ICH, Text on Validation of Analytical Procedures. International Conference on Harmonisation, IFPMA, Geneva, 1994.
[168] ICH, Validation of Analytical Procedures: Methodology. International Conference on Harmonisation, IFPMA, Geneva, 1996.
[169] FDA, Guidance for Industry: Analytical Procedures and Methods Validation (Draft guidance). Food and Drug Administration, Rockville, MD, 2000.
[170] The United States Pharmacopeia, 24th Revision, Asian Edition, United States Pharmacopeial Convention, Inc., Rockville, MD, 2000, pp. 2149–2152.
[171] A.C. Cartwright, B.R. Matthews (Eds.), International Pharmaceutical Product Registration, Ellis Horwood Limited, New York, pp. 246–286.
[172] S. Singh, S. Garg, Pharmatimes 31 (1999) 15–20.
[173] E.W. Ciurczak, Pharm. Tech. 22 (1998) 92–102.
[174] H.G. Brittain, Pharm. Tech. 22 (1998) 82–90.
[175] M.E. Swartz, I.S. Krull, Pharm. Tech. 22 (1998) 104–119.
[176] H. Fabre, A.F. Fell, J. Liq. Chromatogr. 15 (1992) 3031–3043.
[177] M. Grover, M. Gulati, S. Singh, J. Chromatogr. B Biomed. Sci. Appl. 708 (1998) 153–159.
[178] M. Grover, M. Gulati, B. Singh, S. Singh, Pharm. Pharmcol. Comm. 6 (2000) 355–363.
[179] British Pharmacopoeia, British Pharmacopoeial Commission, Middlesex, 2000, p. A411.
[180] S. Singh, R.K. Singla, M. Kumar, R.L. Gupta, Analyst 113 (1988) 1665–1668.
[181] ICH, The Common Technical Document-QUALITY. International Conference on Harmonisation, IFPMA, Geneva, 2000.
[182] ICH, Stability Testing of New Drug Substances and Products. International Conference on Harmonisation, IFPMA, Geneva, 2000, p. 17.
[183] Available from http://www.cmcissues.com/Methods/HPLC/mass_balance.htm.
[184] ICH, Stability Testing of New Drug Substances and Products. International Conference on Harmonisation, IFPMA, Geneva, 1993, p. 10.
[185] ICH, Stability Testing of New Drug Substances and Products. International Conference on Harmonisation, IFPMA, Geneva, 1993, p. 4, 7.
[186] ICH, Stability Testing of New Drug Substances and Products. International Conference on Harmonisation, IFPMA, Geneva, 2000, p. 13.
[187] ICH, Impurities in New Drug Substances. International Conference on Harmonisation, IFPMA, Geneva, 1995, p. 5.
[188] British Pharmacopoeia, British Pharmacopoeial Commission, Middlesex, 2000, p. 412.
[189] ICH, Good Manufacturing Practices for Active Pharmaceutical Ingredients. International Conference on Harmonisation, IFPMA, Geneva, 2000, p. 35.
[190] R.H. Muller, W. Mehnert (Eds.), Particle and Surface Characterization Methods. Medpharm Scientific Publishers, Stuttgart, pp. 129–157.
[191] S.S. Williams, J.F. Karnicky, J.-L. Excoffier, S.R. Abbott, J. Chromatogr. 485 (1989) 267–281.
[192] L.R. Snyder, J.J. Kirkland, J.L. Glach (Eds.), Practical HPLC Method Development, Wiley, New York, 1997, pp. 439–478.
[193] K. Outinen, H. Vuorela, R. Hiltunen, Eur. J. Pharm. Sci. 4 (1996) 199–210.
[194] K. Outinen, H. Haario, P. Vuorela, R. Hiltunen, Eur. J. Pharm. Sci. 6 (1998) 197–205.
[195] Available from www.multisimplex.com.
[196] N.H. Shulman, R.K. Fyfe, J. Clin. Pharm. Ther. 20 (1995) 41–44.
[197] N. Surendran, S.O. Ugwu, E.J. Sterling, J. Blanchard, J. Chromatogr. B Biomed. Appl. 670 (1995) 235–242.
[198] G.L. Hoyera, P.E. Nolan Jr., J.H. LeDouxa, L.A. Moorea, J. Chromatogr. A 699 (1995) 383–388.

Ogenstad Declaration Exhibit C

THE SCIENCE & BUSINESS OF BIOPHARMACEUTICALS
BioPharm INTERNATIONAL.COM

Assessing Shelf Life Using Real-Time and Accelerated Stability Tests

Although Accelerated Tests Are Needed, Real-Time Tests Are the Ultimate Proof
Nov 01, 2003
By Robert T. Magari, Ph.D. [1]
BioPharm International
Volume 16, Issue 11

Biopharmaceutical products in storage change as they age, but they are considered to be stable as long as their characteristics remain within the manufacturer's specifications. The number of days that the product remains stable at the recommended storage conditions is referred to as the shelf life. The experimental protocols commonly used for data collection that serve as the basis for estimation of shelf life are called stability tests.

Shelf life is commonly estimated using two types of stability testing: real-time stability tests and accelerated stability tests. In real-time stability testing, a product is stored at recommended storage conditions and monitored until it fails the specification. In accelerated stability tests, a product is stored at elevated stress conditions (such as temperature, humidity, and pH). Degradation at the recommended storage conditions can be predicted using known relationships between the acceleration factor and the degradation rate.

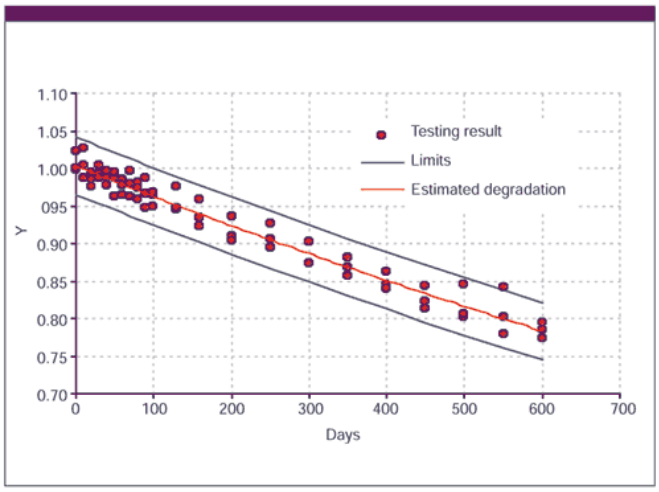


Figure 1. A simulated set of stability results also showing the estimated degradation and 95% confidence limits.

Figure 1. A simulated set of stability results also showing the estimated degradation and 95% confidence limits.

Temperature is the most common acceleration factor used for chemicals, pharmaceuticals, and biological products because its relationship with the degradation rate is characterized by the Arrhenius equation. Several methods of predicting shelf life based on accelerated stability testing are described in the article. Humidity and pH also have acceleration effects but, because they are complex, they will not be discussed in detail here. Also, details on statistical modeling and estimation are outside the scope of the article, but we provide references to computer routines.

Regulations and History The assessment of shelf life has evolved from examining the data and making an educated guess, through plotting, to the application of rigorous physical-chemical laws and statistical techniques. Regulators now insist that adequate stability testing be conducted to provide evidence of the performance of a drug or a biopharmaceutical product at different environmental conditions and to establish the recommended storage conditions and shelf life.[1-3] Recently, Tsong reviewed the latest approaches to statistical modeling of stability tests,[4] and ICH has published some guidelines for advanced testing design and data analysis.[5,6]

rameter	Estimate	Standard Error	Lower Limit	Upper limit
ercept, α	1.003	0.006	0.977	1.029
gradation rate, δ	–0.00041	0.00001	–0.00045	–0.00038
ys	541	9.9	498	584

1. Estimates of the degradation model

	Degradation rate		Days of stability		
mperature	Estimate	Standard Error	Estimate	Lower Limit	Upper Limit
°C	0.0010	7.4E–06	220.1	217.0	223.3
°C	0.0019	5.6E–06	117.9	117.2	118.6
°C	0.0042	1.5E–05	53.6	53.2	54.0
°C	0.0089	5.9%–05	25.1	24.8	25.4

2. Estimates of degradation rates, days of stability and 95% confidence limits.

1. Estimates of the degradation model and Table 2. Estimates of degradation rates, of stability and 95% confidence limits.

Modeling has become easier due to availability of standard statistical software that can perform the calculations. However, an understanding of the general principles of stability testing is necessary to apply these programs correctly and obtain appropriate results. Thus, the purpose of this paper is to provide an outline of the basic approaches to stability testing, as well as to create a foundation for advanced statistical modeling and shelf life prediction.

Stability and Degradation Since degradation is usually defined in terms of loss of activity or performance, a product is considered to be degrading when any characteristic of interest (for example potency or performance) decreases. Degradation usually follows a specific pattern depending on the kinetics of the chemical reaction. The degradation pattern can follow zero-, first-, and second-order reaction mechanisms.[6] In zero-order reactions, degradation is independent of the concentration of remaining intact molecules; in first-order reactions, degradation is proportional to that concentration.[6,7] Zero- and first-order reactions involve only one kind of molecule, and can be described with linear or exponential relationships. Second- and higher-order reactions involve multiple interactions of two or more kinds of molecules and are characteristic of most biological materials that consist of large and complex molecular structures. Although it is common to approximate these reactions with an exponential relationship, sometimes their degradation pattern needs to be modeled more precisely, and no shortcuts will suffice.

The degradation rate depends on the activation energy for the chemical reaction and is product specific. We don't always have to deal with higher-order equations; in many cases, the observed responses of different orders of reactions are indistinguishable for products that degrade slowly.

The degradation rate depends on the conditions where the chemical reaction takes place. Products degrade faster when subjected to acceleration factors such as temperature, humidity, pH, and radiation. Modeling of the degradation pattern and estimation of the degradation rate are important for assessing shelf life. Experimental protocols used for data collection are called stability tests. In practice, evaluators use both real-time stability tests and accelerated stability tests. The real-time stability test is preferable to regulators. However, since it can take up to two years to complete, the accelerated tests are often used as temporary measures to expedite drug introduction.

Real-Time Stability Tests In real-time stability tests, a product is stored at recommended storage conditions and monitored for a period of time (t_{test}). Product will degrade below its specification, at some time, denoted t_s, and we must also assure that t_s is less than or equal to t_{test}. The estimated value of t_s can be obtained by modeling the degradation pattern. Good experimental design and practices are needed to minimize the risk of biases and reduce the amount of random error during data collection. Testing should be performed at time intervals that encompass the target shelf life and must be continued for a period after the product degrades below specification. It is also required that at least three lots of material be used in stability testing to capture lot-to-lot variation, an important source of product variability.[1,2]

The true degradation pattern of a certain product, assuming that it degrades via a first-order reaction, can be described as follows:

The observed result (Y) of each lot has a random component φ associated with that lot, as well as a random experimental error, ε.

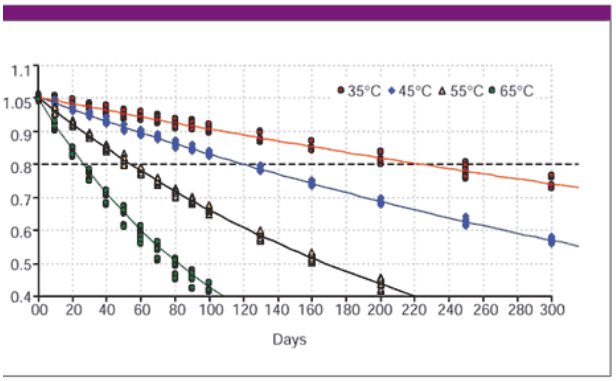
Both α and δ represent the fixed parameters of the model that need to be estimated from the data, while φ and ε are assumed to be normally distributed with mean = 0, and standard deviations of σ_φ and σ_ε respectively. Equation 2 is a nonlinear mixed model. Details on the estimation process are outside the scope of this paper.[8,9]

Let C represent a critical level where the essential performance characteristics of the product are within the specification. A product is considered to be stable when $Y \geq C$.

Product is not stable when Y < C, while Y < C occurs at t_s. The manufacturer determines the value of C. The estimated time that the product is stable is calculated as

$$t_s = \{\ln C - \ln a\}/-d \qquad [3]$$

Here, a and d are the estimated values of the intercept and the degredation rate. The standard error of the estimated time can be obtained from the Taylor series approximation method and is used to calculate confidence limits. The labeled shelf life of the product is the lower confidence limit of the estimated time.[8] Public safety is paramount, that is why we use the lower confidence limit. Lots should be modeled separately when lot-to-lot variability is large. More details on this issue are found in references 9 and 10.



2. A set of simulated data showing degradation of product at four different peratures.

2. A set of simulated data showing degradation of product at four different ratures.

We simulated data for three lots tested for a total period of 600 days (Table 1 and Figure 1). The product loses its activity as it ages, but it is considered to be performing within the specification until it reaches 80% of its activity (C = 0.8). The estimated lot-to-lot standard deviation is 0.000104, and the estimate of experimental error is 0.000262. Therefore, the shelf life of the product was determined to be 498 days. This represents the lower 95% confidence limit corresponding to the estimated time of 541 days.

Parameter	Estimate	Standard Error	Lower Limit	Upper Limit
Degradation rate, δ	0.00039	0.000004	0.00038	0.00040
Activation energy, E_a	15.4	0.060	15.3	15.6
Days	572	5.57	561	583

Table 3. Predictions of parameters at 25°C based on the Arrhenius equation.

Table 3. Predictions of parameters at 25°C based on the Arrhenius equation.

Accelerated Stability Tests In accelerated stability testing, a product is stored at elevated stress conditions. Degradation at recommended storage conditions could be predicted based on the degradation at each stress condition and known relationships between the acceleration factor and the degradation rate. A product may be released based on accelerated stability data, but the real-time testing must be done in parallel to confirm the shelf-life prediction.[1] Sometimes the amount of error of the predicted stability is so large that the prediction itself is not useful. Design your experiments carefully to reduce this error. It is recommended that several production lots should be stored at various acceleration levels to reduce prediction error. Increasing the number of levels is a good strategy for reducing error.

Temperature is probably the most common acceleration factor used for chemicals, pharmaceuticals, and biological products since its relationship with the degradation rate is well characterized by the Arrhenius equation. This equation describes a relationship between temperature and the degradation rate as in Equation 4.

$$\delta = A\exp\left(\frac{-E_a}{RT}\right) \qquad [4]$$

This relationship can be used in accelerated stability studies when the following conditions are met:

A zero- or first-order kinetics reaction takes place at each elevated temperature as well as at the recommended storage temperature.[7]
The same model is used to fit the degradation patterns at each temperature.[7,8]

These requirements do not fully guarantee that the Arrhenius equation can be used to predict the degradation rate at storage temperature, but they are a good start. Do not compromise the analytical accuracy during the course of the study to distinguish between the degradation rates at each temperature.

Select temperature levels based on the nature of the product and the recommended storage temperature. The selected temperatures should stimulate relatively fast degradation and quick testing but not destroy the fundamental characteristics of the product. It is not reasonable to test at very high temperatures for a very short period of time, since the mechanisms of degradation at high temperatures may be very different than those at the recommended storage temperature. Choose the adjacent levels appropriately so that degradation trends are larger than experimental variability. Choosing levels depends on the nature of the product and analytical accuracy, but other practical implications may be considered. Testing should be performed at time intervals that encompass the target stability at each elevated temperature. Acquire some data below C so that the degradation trend can be determined.

Humidity and pH can be used along with temperature to accelerate degradation, but modeling of multi-factor degradation is very complex. A model for parameter estimation and prediction of shelf life when temperature and pH are used as acceleration factors is given by Some et al.[11]

Arrhenius Prediction Assuming that the degradation pattern follows a first-order reaction as described in Equation 2, the Arrhenius equation (Equation 4) can be used to predict the degradation rate at recommended storage temperature. First, an acceleration factor, λ, is calculated as the ratio of the degradation rate at elevated temperature to the degradation rate at storage temperature.[9] This ratio, which can be worked out easily from Equation 4, can be expressed as

$$\lambda = \exp\left[\frac{E_a}{0.00199}\left(\frac{1}{T_s} - \frac{1}{T_e}\right)\right] \quad [5]$$

The true degradation pattern at storage temperature can be expressed as

$$D = \alpha \exp(\delta \lambda t) \quad [6]$$

(Here, λ indicates this was evaluated from accelerated tests.) The testing result (Y) will include random components representing lot-to-lot variability and experimental error. Once the estimates of α and δ are obtained, stability time is calculated in a similar fashion as in real-time stability testing. Shelf life is the lower confidence limit of the estimated time.

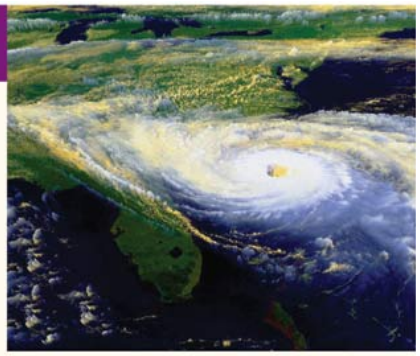
Nomenclature

- A = Arrhenius factor, property/day
- C = Critical level of performance characteristic (minimum)
- D = Degradation pattern of a product, defined in Equation 1
- E_a = Activation energy, kcal/mol
- n = Number of 10 degree K steps
- R = Gas constant, 0.00199 kcal/(mol)(K)
- Q = Temperature proportionality factor, dimensionless
- t = time, days
- t_e = Time product is stable at elevated temperature, days
- t_s = Time to degrade below specification, days
- t_{test} = Monitoring time at recommended storage condition, days
- T = Temperature, K
- T_e = Elevated temperature, K
- T_s = Storage temperature, K
- Y = Observed testing result, dimension not defined
- α = Lot performance at time zero (intercept), dimension not defined
- a = Estimated value of intercept, dimension not defined
- β_0, β_1, β_2 = Coefficients of second-degree polynomial, Equation 8
- δ = Degradation rate, property/day
- d = Estimated value of degradation rate, property/day
- ϵ = Random experimental error, dimension not defined
- λ = Acceleration factor; ratio of (δ at T_e)/(δ at T_s), dimensionless
- ϕ = Random component of a lot, dimension not defined
- σ_ϕ = Standard deviation of ϕ, dimension not defined
- σ_ϵ = Standard deviation of ϵ, dimension not defined

Nomenclature

An example. Simulated data for three lots, each aged at four elevated temperatures for 300 days, are shown in Table 2. The performance of each lot at each time point is measured in three replicates. A critical level of C = 0.8 is the criterion. Data and trends are presented in Figure 2, with the estimates of degradation rates and days given in Table 2. The estimated degradation rate is observed to increase with temperature. The number of days that product performs within the specifications (C = 0.8) is 217 days when stored at 35°C. Stability time will drop down to 24.8 days when product is stored at 65°C. The estimated activation energy is 15.4 kcal/mol. Predictions at 25°C based on the Arrhenius equation are presented in Table 3. The product will perform within specification for an estimated 572 days. Using the lower limit, the recommended shelf life is 561 days.

Unstable Weather Conditions and Shelf-Life



Recent happenings — the massive power outage in the United States and Canada this August and Hurricane Isabel's rampage along the East Coast in September — and the ever-present threat of strong winds and flooding in some regions can create damages that may affect the required storage conditions for biopharmaceutical products.

Most of these products require storage in ventilated dry places, with normal temperatures at around 25°C; sensitive biologicals require refrigeration or frozen storage. The stability of a product will be affected if one or more of the required storage conditions are not at their optimum range, while major deviations from these requirements may affect the basic performance of the product.

In September, the Center for Biologics Evaluation and Research (CBER) issued "Impact of Severe Weather Conditions on Biological Products." In this document, CBER provides information concerning the storage and use of temperature-sensitive biological products that have been involved in a temporary electrical power failure or flood conditions.[1] Shelf life of a product not adversely affected can be reassessed based on the following information:

- Duration of power outage or flood
- Nature of the product and storage requirements
- Age of the product

Manufacturers can perform calculations as well as a risk analysis — in consultation with FDA, as necessary — to assess the continued performance of the involved biopharmaceutical product.

Reference

(1) FDA. Impact of severe weather conditions on biological products. Rockville (MD); 2003 Sep 17.

Unstable Weather Conditions and Shelf-Life

Bracket Method Activation energy is usually estimated from the accelerated stability data. However, when the activation energy is known, the degradation rate at storage temperature may be predicted from data collected at only one elevated temperature. This practice is sometimes preferred in industry since it reduces the size and time of accelerated stability tests. Experience indicates that some pharmaceutical analytes have activation energy in the range of 10 to 20 kcal/mol, but it is unlikely you will have precise information or be able to make assumptions about the activation energy of a certain product.[12]

The bracket method is a straightforward application of the Arrhenius equation that can be used if the value of the activation energy is known.[12] Assuming that stability of a product at 50°C is 32 days, and it will be stored at 25°C, then, t_e = 32 days, T_e = 273 + 50°C = 323K, and T_s = 273 + 25°C = 298K. We know that activation energy is E_a = 10 kcal/mol. Stability at recommended storage temperature is calculated with a modified version of Equation 5 as:

Calculated stability is highly dependent on the value of activation energy. A stability of 435 days results when E_a = 20 kcal/mol.

The bracket method should not be confused with bracketing, which is an experimental design that allows you to test a minimum number of samples at the extremes of certain factors, such as strength, container size, and container fill.[3,4,6] Bracketing assumes that the stability of any intermediate levels is represented by the stability of the extremes and testing at those extremes is performed at all time points.[4]

The Q-Rule The Q-Rule states that the degradation rate decreases by a constant factor when temperature is lowered by certain degrees. The value of Q is typically set at 2, 3, or 4. This factor is proportional to the temperature change as Q^n, where n equals the temperature change in °C divided by 10°C. Since 10°C is the baseline temperature, the Q-Rule is sometimes referred to as Q_{10}.

To illustrate the application of the Q-Rule, let us assume that the stability of a product at 50°C is 32 days. The recommended storage temperature is 25°C and n = (50 - 25)/10 = 2.5. Let us set an intermediate value of Q = 3. Thus, $Q^n = (3)^{2.5}$ = 15.6. The predicted shelf life is 32 days 3 15.6 = 500 days. This approach is more conservative when lower values of Q are used.[12] Both Q-Rule and the bracket methods are rough approximations of stability. They can be effectively used to plan elevated temperature levels as well as the duration of testing in the accelerated stability testing protocol.

High-Order Kinetics Theoretically, the Arrhenius equation does not apply when more than one kind of molecule is involved in reactions. However, if the degradation rate and temperature are linearly related, the prediction of shelf life can be approximated by the Arrhenius equation. Statistics that test the appropriateness of this approximation are presented in literature.

Magari et al. used a polynomial model to fit the degradation of a reagent (HmX PAK) for the Coulter HmX Analyzer.[13] The following degradation pattern was consistent at all elevated temperatures:

$$D = \beta_0 + (\beta_1 + \beta_2 t)t \qquad [8]$$

β_0, β_1, and β_2 are the parameters of the second-degree polynomial and t is time. The degradation rate is a function of time, which is not constant in this case.

$$\delta = \beta_1 + \beta_2 t \qquad [9]$$

Degradation at storage temperature can be predicted from the degradation at elevated temperatures as

$$D = \beta_0 + (\beta_1 + \beta_2 t)t\lambda \qquad [10]$$

The acceleration factor, λ, is based on the Arrhenius equation. Statistical tests indicated that the use of this equation was appropriate in this case. Shelf-life predictions were also verified by real-time stability testing results.

Similar Products When most of the assumptions required to use the Arrhenius equation are not satisfied, comparisons to a product with a known stability is performed to assess shelf life. This approach requires having a similar product with a known shelf life to be used as a control. The new or test product is expected to demonstrate similar behavior to the control since they belong to the same family and have the same kinetics of degradation. Side-by-side testing of the control and test products at different elevated temperatures is then performed. It is necessary to assume that the same model can represent the degradation pattern at each elevated and storage temperature.

If the degradation patterns of the test and control samples at the same elevated temperatures are not statistically different, it can be assumed that they will degrade similarly at the storage temperature. The closer the elevated temperatures are to the storage temperature, the more confident we can be in making this statement. The experimental protocols used are similar to the protocols used with the Arrhenius equation. Degradation patterns of a family of products at certain elevated temperatures can be modeled and used to check the behavior of a new product that belongs to that family.

Data Analysis The complication that was alluded to with Equation 1 is that degradation models are usually nonlinear mixed models, where lot-to-lot variability is the random component. Estimation of the parameters of the models is important for the accuracy of shelf-life predictions. We recommend using the maximum likelihood (ML) approach to estimate these parameters.

Since no closed-form solutions for ML estimates exist, an iterative procedure is performed, starting with some initial values for the parameters and updating them until differences between consecutive iterations are minimal and the estimates converge to their final value. Initial values are usually chosen by experience. The closer these values are to the final values, the faster the model will converge. We used PROC NLMIXED of SAS for data analysis.[14] Values of the real-time stability model (Equation 2) converged relatively quickly, while several initial values for the parameters of the accelerated model (Equation 6) were tried before they converged. Statistical theory and the applicability of ML estimation are common in the literature, and many computer routines are available to facilitate data analysis. However, experience with the modeling and estimation processes is necessary, since any unexpected results must be appropriately interpreted. It is quite easy to get useless numbers from a computer run.

References (1) FDA. Guidelines for submitting documentations for the stability of human drugs and biologics. Rockville (MD); 1987.

(2) ICH. Stability testing for new drug substances and products. Q1A(R). Geneva; 1994.

(3) ICH. Stability testing for new drug substances and products. Q1A(R2). Geneva; 2003.

(4) Tsong Y. Recent issues in stability testing, *J Biopharm Stat* 2003; 13:vii-ix.

(5) ICH. Stability data evaluation. Q1E. Geneva; 2003.

(6) ICH. Bracketing and matrixing designs for stability testing of new drug substances and products. Q1D. Geneva; 2003.

(7) King SP, et al. Statistical prediction of drug stability based on nonlinear parameter estimation. *J Pharm Sci* 1984; 73:2332-2344.

(8) Tydeman MS, Kirkwood TBL. Design and analysis of accelerated degradation tests for the stability of biological standards. 1. Properties of maximum likelihood estimators. *J Biol Stand* 1984; 12:195-206.

(9) Magari R. Estimating degradation in real time and accelerated stability tests with random lot-to-lot variation: a simulation study. *J Pharm Sci* 2002; 91:893-899.

(10) Tsong Y, et al. Shelf life determination based on equivalence assessment. *J Biopharm Stat* 2003; 13:431-449.

(11) Some I, et al. Stability parameter estimation at ambient temperature from studies at elevated temperatures. *J Pharm Sci* 2001; 90:1759-1766.

(12) Anderson G, Scott M. Determination of product shelf life and activation energy for five drugs of abuse. *Clin Chem* 1991; 37:398-429.

(13) Magari R, et al. Accelerated stability model for predicting shelf-life. *J Clin Lab Anal* 2002; 16: 221-226.

(14) SAS Institute. SAS/stat user's guide, vers. 8. Cary (NC); 2000. **BPI**

Source URL: http://www.biopharminternational.com/assessing-shelf-life-using-real-time-and-accelerated-stability-tests

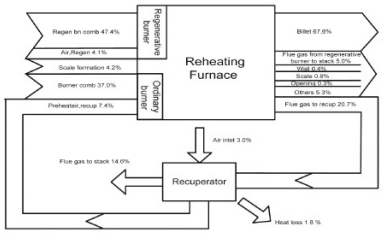
Links:
[1] http://www.biopharminternational.com/robert-t-magari-phd

Ogenstad Declaration Exhibit D

ISSN 1905-9159

Silpakorn University Science and Technology Journal

Volume 4 Number 2 (July - December) 2010



A Study of Specific Energy Consumption in Reheating Furnace using Regenerative Burners Combined with Recuperator

Kanit Manatura and Mingsak Tangtrakul

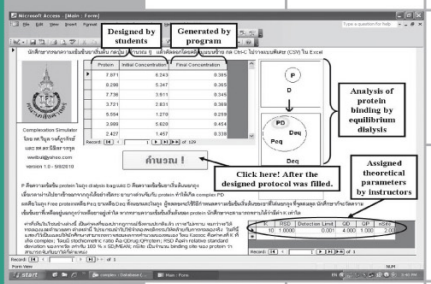
A Simulation Comparison of New Confidence Intervals for the Coefficient of Variation of a Poisson Distribution

Wararit Panichkitkosolkul

$$\left(\bar{X} - Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\bar{X}+0.5}{n}} \;,\; \bar{X} + Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\bar{X}+0.5}{n}} \right).$$

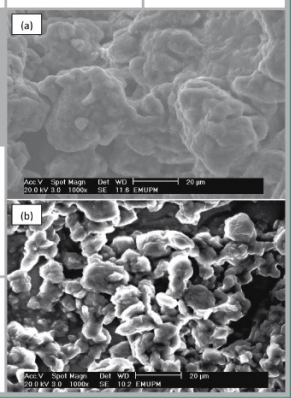
Nutritive Values of Whip Grass (*Hemarthria compressa*) at Different Cutting Intervals Consumed by Thai Indigenous Cattle

Jeerasak Chobtang, Apichat Boonruangkao, Saksan Suankool and Auraiwan Isuwan



Computer Simulation for Studying Complexation between a Model Drug and a Model Protein

Wibul Wongpoowarak, Nimit Worakul, Wiwat Pichayakorn, Payom Wongpoowarak, Prapaporn Boonme



Structural and Morphological Characterization of Chemical Bath Deposition of FeS Thin Films in the Presence of Sodium Tartrate as a Complexing Agent

Anuar Kassim, Ho Soon Min, Loh Yean Yee and Saravanan Nagalingam

SILPAKORN UNIVERSITY
Science and Technology Journal
SUSTJ

Published by Silpakorn University Printing House.
Silpakorn University, Sanamchandra Palace Campus, Nakhon Pathom 73000

© Silpakorn University ISSN 1905-9159

SILPAKORN UNIVERSITY
Science and Technology Journal

Editorial Office

Silpakorn University Research and Development Institute (SURDI), Silpakorn University,
Sanamchandra Palace Campus, Nakhon Pathom, Thailand

Editorial Policy

All articles submitted for publication will be evaluated by a group of distinguished reviewers.
The editorial board claims no responsibility for the contents or opinion
expressed by the authors of individual article.

SILPAKORN UNIVERSITY
Science and Technology Journal

Periodicity
Twice yearly

All correspondence should be addressed to:
Managing Editor, SUSTJ, 44/114 Soi Phaholyothin 52 Phaholyothin Road,
Klongthanon, Saimai, Bangkok 10220, Thailand
Telephone: 080-5996680 Fax: 66-2973-8366
E-mail address: pranee_aon@hotmail.com
Web site: http://www.journal.su.ac.th

Information about the Journal
An electronic journal is provided on the web site (http://www.journal.su.ac.th).
The journal is available at Silpakorn University Book Center. Telephone: 66-2223-7345,
66-2434-1792, 66-3424-4054.

Instructions to Authors

Silpakorn University Science and Technology Journal (SUSTJ) is published twice a year in June and December by the Research and Development Institute of Silpakorn University, Thailand. The journal puts together articles in Science and Technology and aims to promote and distribute peer reviewed articles in the areas of science, health science, animal science, agriculture, engineering, technology and related fields. Articles from local and foreign researchers, invited articles and review from experts are welcome.

Types of contributions

Short communications, Research articles, Review articles

Preparation of manuscripts

1. The text should be double-spaced with line number on A4 and a font Times New Roman size 11 should be used. When using MS Word, insert all symbols by selecting "Insert-Symbol" from the menu and use the "Symbol" font.

2. Manuscripts should be organized in the following order:

Cover page with title and authors' names and affiliations
Abstract (in English and Thai)
Key Words
Introduction
Materials and Methods, Area Descriptions, Techniques
Results
Discussion
Conclusion
Acknowledgements
References
Tables and Figures

Authors' names and affiliations

Full names and affiliations (marked with superscript number) should be provided for all authors on the cover page, separately from the content. The corresponding author (marked with superscript asterisk) should also provide a full postal address, telephone and fax number and an e-mail address as a footnote on the title page.

Abstract

First page of the content starts with Abstract, including title of the article on top of page. Provide a short abstract not more than 200 words, summarizing the question being addressed and the findings.

Key Words

Provide 3-5 key words or short phrases in alphabetical order, suitable for indexing.

References

In text references: Refer to the author's name (without initials) and year of publication, e.g., Feldmann, 2004 (for 1 author), Feldmann and Langer, 2004 (for 2 authors), or Feldmann et al., 2004 (for more than 2 authors).

Article references: References should be listed in alphabetical order of author(s). For journal, list all names of authors.

Book

Feldmann, H. (2004) *Forty Years of FEBS,* Blackwell Publishing Ltd., Oxford.

Chapter in a book

Langer, T. and Neupert, W. (1994) Chaperoning mitochondrial biogenesis. In *The Biology of Heat Shock Proteins and Molecular Chaperones* (Morimoto, R. I., Tissieres, A. and Georgopoulos, C., eds.), pp. 53-83. Cold Spring Harbor Laboratory Press, Plainview, New York.*Article in a journal*

Hammerschlag, F. A., Bauchan, G., and Scorza, R. (1985) Regeneration of peach plants from callus derived from immature embryos. *Journal of Natural Products* 70(3): 248-251.

Hammerschlag, F. A., Bauchan, G., and Scorza, R. Regeneration of peach plants from callus derived from immature embryos. *Journal of Natural Products* (in press).

Article on the web

Lee, K. (1999) Appraising adaptaive management. *Conservation Ecology* 3(2). [Online URL:www.consecolo.org/Journal/vol3/iss2/index.html] accessed on April 13, 2001.

Proceedings

MacKinnon, R. (2003) Modelling water uptake and soluble solids losses by puffed breakfast cereal immersed in water or milk. In *Proceedings of the Seventh International Congress on Engineering and Food*, Brighton, UK.

Patent

Yoshikawa, T. and Kawai, M. (2006) Security robot. U.S. Patent No. 2006079998.

Tables and Figures

Each Table and Figure must be on a separate page of the manuscript.

Tables: Number the tables according to their sequence in the text. The text should include references to all tables. Vertical lines should not be used to separate columns. Leave some extra space instead.

Figures: Figures should be of high quality (not less than 300 dpi JPEG or TIFF format), in black and white only, with the same size as the author would like them to appear in press. Choose the size of symbols and lettering so that the figures can be reduced to fit on a page or in a column.

Submission of Manuscripts

Authors should verify, on the submission form, that the submitted manuscript has not been published or is being considered for publication elsewhere. All information contained in an article is full responsibility of the authors, including the accuracy of the data and resulting conclusion.

Authors are requested to send the manuscript on a CD labeled with the authors' names and file names. The files should be prepared using MS Word only. Three copies of manuscript must be supplied.

The editorial office will acknowledge receipt of the manuscript within 2 weeks of submission. The 'accepted date' that appears in the published article will be the date when the managing editor receives the fully revised version of the manuscript.

The manuscript may be returned to authors for revision. Authors will be given 2 weeks after receipt of the reviewers' comments to revise the article.

Please submit the manuscript with a CD and a submission form to the following address:
Pranee Vichansvakul
44/114 Soi Phaholyothin 52 Phaholyothin Road,
Klongthanon, Saimai, Bangkok 10220

Proofs

Proofs will be sent to the corresponding author by e-mail (as PDF file) or regular mail. Author is requested to check the proofs and return any corrections within 2 weeks.

SUSTJ

Silpakorn University Science and Technology Journal

Volume 4 Number 2 (July - December) 2010

Contents

Research Articles

SUSTJ is now available on the following databases:
Chemical Abstract Service (CAS),
International Information System for the Agricultural Sciences and Technology (AGRIS),
AGRICultural Online Access (AGRICOLA)
Food Science and Technology Abstracts (FSTA),
Directory of Open Access Journals (DOAJ),
Google Scholar,
Thai Journal Citation Index Centre (TCI Centre).

Research Article

A Study of Specific Energy Consumption in Reheating Furnace Using Regenerative Burners Combined with Recuperator

Kanit Manatura[1*] and Mingsak Tangtrakul[2]

[1]*Department of Mechanical Engineering, Faculty of Engineering Kamphaeng Saen, Kasetsart University, Kamphaeng Saen Campus, Nakhon Pathom, Thailand*
[2]*Department of Mechanical Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok, Thailand*
[*]*Corresponding author. E-mail address: kanitman@hotmail.com, fengknm@ku.ac.th*

Received March 9, 2010; Accepted December 9, 2010

Abstract

The steel industry is one of high energy consumption industries. In order to mill steel bar into steel rod, the steel bar is heated to 1,100 - 1,250 ˚C. The objective of this work is to investigate energy utilization in reheating furnace using regenerative burners combined with recuperator. The furnace capacity is 30 tonne per hour, pusher type and the natural gas is used as fuel. Billet sizes 120 x 120 x 4000 mm are used for reheating. Waste heat recovery in recuperator system can preheat combustion air to 300 ˚C due to material temperature limitations. In order that preheating combustion air temperature near furnace temperature so regenerative burner system provides for substituting that it can preheat combustion air up to 1000 ˚C. The results from measurements and energy balance analysis indicate that the regenerative burners combined with recuperator system consume energy approximately 43% less than the case study of conventional recuperative system

***Key Words*:** Energy consumption; Reheating furnace; Regenerative burners; Recuperator; Energy balance

Introduction

At the present time, the cost of fuel using as an energy source has constantly increased due to limiting of natural resource. Each of countries around the world has realized on its higher cost because it is one of the main capital costs of production, makes their products more expensive and consequently can not compete in market. One of the industries that use a lot of heat is the steel industry. The steel industry is mainly basic in development of destination industries in which are principal industry of each country, such as construction, automotive and electric appliance industries. The reheating furnace use fuel for heating billets or slabs for rolling process. In Thailand, have waste heat recovery in preheating combustion air by recuperator. The most of combustion air can be preheated the maximum temperature of 300 ˚C when the temperature efficiency only is 30%. The application of regenerative burner technology instead of conventional burner and recuperator are able to preheat combustion air nearly 1,000 ˚C and the temperature efficiency up to 90%. These can save energy of 10-20% (O'Connor et al., 2006)

compare with conventional recuperator system in reheating process. Generally, the regenerative burner consists of regenerators such as alumina, ceramic ball or honeycomb as regenerative media. The principle of regenerative burner using the regenerator that recovers the heat from the flue gas and use it to increase the temperature of combustion air. Normally, regenerative burner is suitable for installation in furnace capacity 500 kW at least. The problems of low capacity furnace are area for installation and high cost in installation that affect to in late payback period. However, regenerative burners have been developing for applying in low capacity furnace that can save energy more than 35% (Wuenning, 2008).

Methodology

Reheating Furnace Description

A capacity of the reheating furnace is 30 tonne/hour, pusher type. It consists of 3 zones: preheating zone, heating zone and soaking zone. The heating zone equipped with 3 pairs of regenerative burner, capacity of each pair is 2326 kW, the switching time is 30 s. The soaking zone equipped with 4 ordinary burners, capacity of each burner is 872 kW. Size of Billet is 120mm x 120mm x 4000mm used for reheating, each piece weighs 444 kg. The diagram of the reheating furnace system is shown as Figure 1

Compositions and properties of fuel

The natural gas is used as fuel in this research for combustion. Mixtures of natural gas varied with resource. In this research, Compositions are shown in volumetric percentage of natural gas from Ratchaburi gas station,Thailand (Table 1).

Table 1 Natural gas components (PTT, 2002)

COMPONENTS	PERCENT BY VOLUME
CH_4	72.6
C_2H_6	3.5
C_3H_8	1.1
C_4H_{10}	0.4
C_5H_{12}	0.2
C_6H_{14}	0.1
CO_2	6.1
N_2	16

The higher heating value of the heat of combustion calculated by assuming that all of the water in products has condensed to liquid is always used in calculation. However, in practical events the flue gas temperature from combustion has valued in high level around 500-800 ˚C effect on water vapor in flue gas is still prior condition. Moreover, it does not have latent heat value from condition transferring. Therefore, the heating value have been used in calculation should be the lower heating value (Table 2).

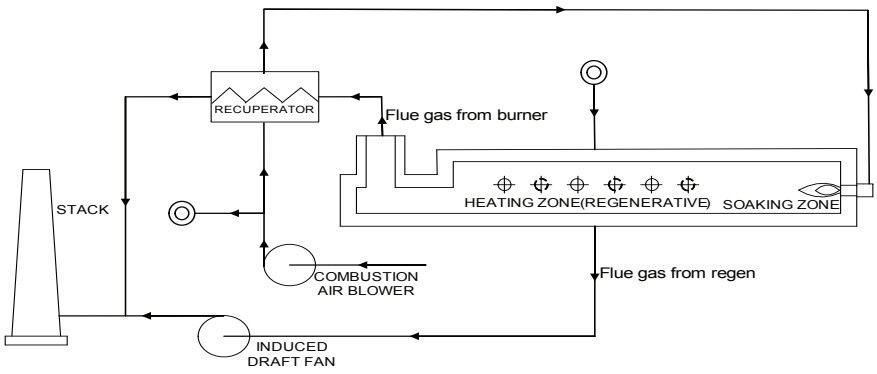


Figure 1 Diagram of the reheating furnace

Table 2 Properties of natural gas (PTT, 2002)

PROPERTIES	VALUE	UNIT
Molecular weight	20.7	kg/kmol
Specific gravity	0.7	-
Higher heating value	31.7	MJ/Nm^3
Lower heating value	26.4	MJ/Nm^3

Analysis of air-fuel ratio in combustion

The hydrocarbon such as the natural gas, the stoichiometric combustion equation can be expressed as (Turns, 2006)

$$(0.726CH_4 + 0.035C_2H_6 + 0.011C_3H_8 + 0.005C_4H_{10} + 0.001C_5H_{12} + 0.001C_6H_{14} + 0.061CO_2 + 0.16N_2) + a_{th}(O_2+3.76N_2) \rightarrow 0.858CO_2 + 0.061CO_2 + 1.64H_2O + a_{th}(3.76N_2) + 0.16N_2 \quad (1)$$

where a_{th} is the stoichiometry ratio of oxygen mole per natural gas mole, the number atoms on the L.H.S of the equation must exactly balance the number on the R.H.S because the combustion process does not create or destroy atoms. Solving equations for oxygen mole getting, $a_{th} = 1.68\ kmol_{O_2}/kmol_{fuel}$

Stoichiometric Air-Fuel ratio (AF_{stoic}) is necessary to achieve complete combustion of the fuel and no more. It can be written as

$$AF_{stoic} = \dot{m}_{air}/\dot{m}_{fuel} = 16.96\ kg_{air}/kg_{fuel} \quad (2)$$

In practical situations, more than the stoichiometric quantity the excess of oxidizer required for completely in combustion so Eq. (1) can be rewritten as

$$(0.726CH_4 + 0.035C_2H_6 + 0.011C_3H_8 + 0.004C_4H_{10} + 0.002C_5H_{12} + 0.001C_6H_{14} + 0.061CO_2 + 0.16N_2) + ma_{th}(O_2+3.76N_2) \rightarrow 0.858CO_2 + 0.061CO_2 + 1.64H_2O + (m-1)\,a_{th}O_2 + ma_{th}(3.76N_2) + 0.16N_2 \quad (3)$$

Where m is the correction factor of excess air

From above equations, found by oxygen mole in excess air in products (flue gas) are equal to $(m-1)a_{th}O_2$. In practice, the flue gas analyzer is used to measure it. The results from the measure are showed in percentage of dry-basis, water vapor and humidity in flue gas is blown off before the apparatus analysis. The remaining gases are then expressed as a percentage (by volume) of the total dry gas constituents - in this case CO_2, O_2, N_2. Therefore, percentage of oxygen from measurement can be adapted to correction factor of excess air as

$$\frac{\%O_2}{100} = \frac{(m-1)a_{th}}{0.858 + 0.061 + (m-1)a_{th} + 3.76ma_{th} + 0.16} \quad (4)$$

Eq. (4) is arranged in the form of m as

$$m = \frac{1.68 - 0.6\left(\frac{\%O_2}{100}\right)}{1.68 - 8\left(\frac{\%O_2}{100}\right)} \quad (5)$$

When knows value of m from Eq. (5) so the Actual Air-Fuel ratio, AF_{actual} is

$$AF_{actual} = m.AF_{stoic} \quad (6)$$

The methodology and analysis of energy balance in reheating furnace

Analysis of energy balance is divided to mass and energy (heat) balance. The conservation of mass and energy are applied in this research; if there is no mass/heat accumulation what goes into process (reheating) must come out in continuous operation by giving the reheating furnace is control volume. A steady state, thermal equilibrium is considered in type of continuous steel reheating furnace. Energy balance analysis point out in energy consumption, furnace efficiency, SEC (specific energy consumption) and other performance parameters.

Mass balance of the reheating furnace

Given the reheating furnace is valuable in negative pressure during operation. The mass flow rate into and out from reheating furnace can be expressed as Eq. (7) and the details of mass balance (Cengel and Boles, 2007) are shown as Table 3 and Table 4, respectively.

Table 3 Mass flow rate into the reheating furnace

INPUT	Formula	kg/s	%
1) Fuel flow rate into regenerative burner	$\dot{m}_{fuel,regen} = \rho_{fuel}\dot{\forall}_{fuel,regen}$	0.2	2.3
2) Combustion air flow rate into regenerative burner	$\dot{m}_{air,regen} = AF_{actual} \times \dot{m}_{fuel,regen}$	3.8	53.4
3) Fuel flow rate into ordinary burner	$\dot{m}_{fuel,burner} = \rho_{fuel}\dot{\forall}_{fuel,burner}$	0.1	1.8
4) Preheated air flow rate into ordinary burner	$\dot{m}_{preheatair,burner} = AF_{actual} \times \dot{m}_{fuel,burner}$	3.1	42.5
Total mass input		7.2	100

Table 4 Mass flow rate out from the reheating furnace

OUTPUT	Formula	kg/s	%
1) Flue gas flow rate from regenerative burner to stack	$\dot{m}_{flue,stack} = \dot{m}_{fuel,regen} + \dot{m}_{air,regen}$	4.0	55.7
2) Flue gas flow rate from the reheating furnace to recuperator	$\dot{m}_{flue,recup} = \dot{m}_{fuel,burner} + \dot{m}_{preheatair,burner}$	3.2	44.3
Total mass output		7.2	100

From above tables, it is clear for applying diagram of mass balance of a reheating furnace as Figure 2

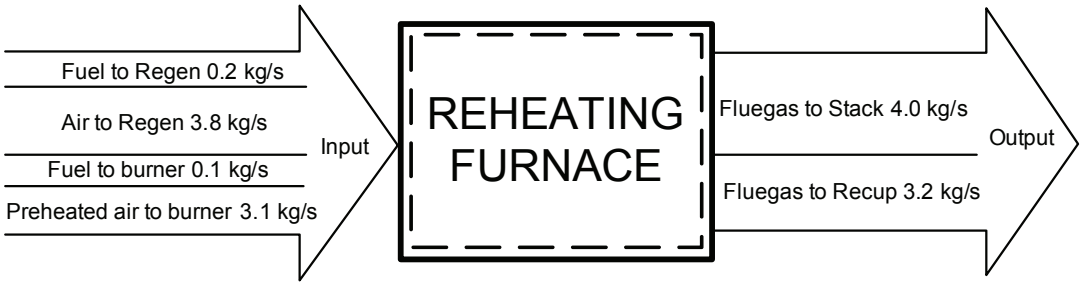


Figure 2 Mass balance diagram for reheating furnace

Figure 2 shows mass balance diagram leads to know about that the mass flow rate of fuel flow rate in regenerative burner system which is twice to burner system according to the law of conservation of mass.

Energy balance of the reheating furnace

The energy balance (Cengel and Boles, 2007) are consist of heat input and heat output of the case study shown in Table 5

Table 5 Energy balance of reheating furnace

Heat input	Heat output
1) Combustion from fuel at regenerative burners 2) Combustion from fuel at ordinary burner 3) Preheated air by recuperator 4) Sensible heat of air inlet 5) Sensible heat of scale formation	1) Sensible heat into billet 2) Sensible heat of flue gas from regenerative burners 3) Sensible heat of flue gas of ordinary burner 4) Heat loss in wall 5) Heat loss from opening 6) Sensible heat into scale 7) Other loss

Assigned to temperature (T_0) and pressure (P_0) are 25 ˚C and 1 atm, respectively as references in calculation.

Furnace efficiency

The efficiency of the reheating furnace is the ratio of the sensible heat of billet ($\dot{Q}_{billet}$) to the heat of combustion in regenerative burner combined with recuperator ($\dot{Q}_{comb}$) is defined as

$$\eta_{furnace} = \frac{\dot{Q}_{billet}}{\dot{Q}_{comb}} \times 100\% \quad (7)$$

The Specific Energy Consumption (SEC)

The Specific Energy Consumption (SEC) is defined as total of energy consumption to total of quantity material processed. In this case study, the reheating furnace using regenerative burner combined with recuperator is control volume studied. The unit of SEC is MJ/ton in this case study.

$$SEC = \frac{\text{Quantity of Energy Consumption}}{\text{Quantity of material processed}} \times 100\% \quad (8)$$

From calculation the SEC of this case study is 1042 MJ/ton or equal to 26.2 litre/ton of fuel oil, 28 Nm^3/ton of natural gas.

Results and Discussion

Data measurement

Measurement results of the reheating furnace are tabulated in Table 6 below.

Table 6 Measurement results of the reheating furnace

Measured parameter	Unit	Values (After)	Values (Before)
Average fuel consumption at regenerative burners	m^3/s	0.192	-
Average % oxygen in flue gas of regenerative burners	%	5.93	-
Preheat air temperature from regenerator	˚C	931.77	-
Flue gases temperature into regenerator	˚C	1003.7	-
Flue gases temperature from regenerator to stack	˚C	145.79	-
Average fuel consumption at ordinary burner	m^3/s	0.15	0.0002
Average % oxygen in flue gas of ordinary burner	%	6.2	15
Preheat air temperature by recuperator into reheating furnace	˚C	262.57	68.05
Flue gas temperature into recuperator	˚C	625.63	716.4
Quantity of produced billet	kg/s	8.64	5.65
Billet temperature inlet to reheating furnace	˚C	39.08	34.12
Billet temperature outlet from reheating furnace	˚C	1035.38	928.2
Average furnace wall temperature	˚C	105.1	300.1
Temperature in reheating furnace	˚C	1085.23	716.4

The Sankey Diagram of energy balance in case study is presented in Figure 3.

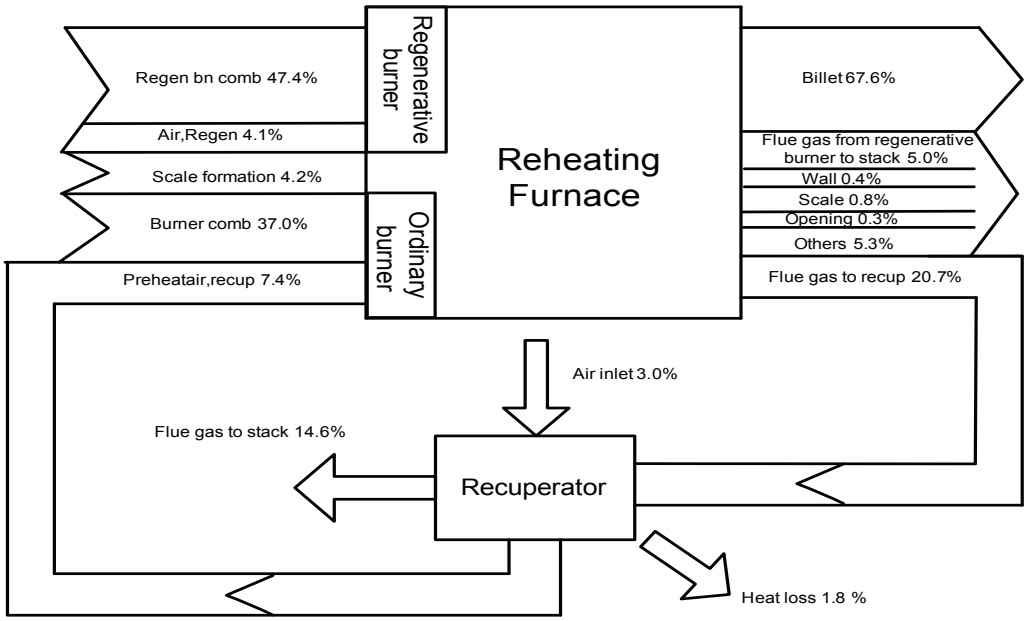


Figure 3 Sankey Diagram of the reheating furnace

From Sankey diagram in Figure 3, shown by the total heat input into the case study are consist of sensible heat from combustion of regenerative burners by 5055 kW (47.4%), sensible heat from combustion of ordinary burner by 3948.7 kW (37.0%), sensible heat from preheated air at recuperator by 784.9 kW (7.4%), sensible heat of fresh air into regenerative burner by 439 kW (4.1%), sensible heat from scale formation by 445.9 kW (4.2%).

Moreover, the total heat output are composed of sensible heat into billet, sensible heat of flue gas from regenerative burners, sensible heat of flue gas from the reheating furnace to recuperator, heat loss in wall, heat loss from opening, heat loss into scale and other loss by 7213.3 kW (67.6%), 528.8 kW (5.0%), 2210.7 kW (20.7%), 42.8 kW (0.4%), 28.5 kW (0.3%), 83.2 kW (0.8%) and 566.3 kW (5.3%), respectively.

Conclusion

In energy balance analysis, the Specific Energy Consumption (SEC) of the reheating furnace using regenerative burners combined with recuperator (furnace capacity = 30 tonne/hr) is 1042 MJ/ton. Moreover, the furnace efficiency of this case study is 80.1% when compare with the reheating furnace using only recuperator (former furnace capacity = 15 tonne/hr). The case study designates in sensible heat of flue gas from regenerative burner which goes down by 25% or 1/4 of sensible heat of flue gas from the reheating furnace to recuperator. The energy saving found by 43.3% of this case study, comparing to the reheating furnace using only recuperator and shows significantly of efficiency in regenerative burner performance.

Acknowledgements

The authors gratefully acknowledge the contribution of Kasemsakdi Trading CO., LTD (Thailand), Iron and steel institute of Thailand, Energy Policy and Planing Office (EPPO) Ministry of Energy (Thailand) and Energy research institute Chulalongkorn University (Thailand).

References

Cengel, Y.A. and Boles, M.A. (2007) *Thermodynamics, an Engineering Approach*, McGraw-Hill, Singapore.

O'Connor,S.J., Konziela J., Yoo ln and Kim Byung Gi. (2006) Regenerative burner in the INI large structural mill furnace, In *AIS Tech 2005 proceeding,The Iron & Steel Technology*.

PTT Public Company Limited. (2002) [Online URL:www.http://pttinternet.pttplc.com/csc_gas/csc_ind/onlinegas/9_3.asp] accessed on April 23, 2009.

Turns, S. (2006) *An Introduction to combustion*, McGraw-Hill. Singapore.

Wuenning J. G. (2008) Regenerative Burners for Heat Treating Furnaces, In *8th European Conference on Industrial Furnaces and Boilers*, Vilamoura, Portugal.

Research Article

A Simulation Comparison of New Confidence Intervals for the Coefficient of Variation of a Poisson Distribution

Wararit Panichkitkosolkul*

Department of Mathematics and Statistics, Faculty of Science and Technology, Thammasat University, Phathumthani, Thailand

**Corresponding author. E-mail address: wararit@mathstat.sci.tu.ac.th*

Received July 6, 2010; Accepted December 13, 2010

Abstract

This paper proposes four new confidence intervals for the coefficient of variation of a Poisson distribution based on obtaining confidence intervals for the Poisson mean. The following confidence intervals are considered: confidence intervals for the coefficient of variation of a Poisson distribution based on Wald (W), Wald with continuity correction (WCC), Scores (S) and Variance stabilizing (VS) confidence interval. Using Monte Carlo simulations, the coverage probabilities and lengths of these confidence intervals are compared. Simulation results have shown that the confidence interval based on WCC has desired closeness coverage probabilities of 0.95 and 0.90. Additionally, the lengths of newly proposed confidence intervals are slightly different. Therefore, the confidence interval based on WCC is more suitable than the other three confidence intervals in terms of the coverage probability.

Key Words: Coefficient of variation; Confidence interval; Coverage probability; Length; Poisson distribution

Introduction

The coefficient of variation is a dimensionless number that quantifies the degree of variability relative to the mean (Kelley, 2007). The population coefficient of variation is defined as

$$\kappa = \frac{\sigma}{\mu}, \qquad (1)$$

where σ is the population standard deviation and μ is the population mean. The typical sample estimate of κ is given as

$$\hat{\kappa} = \frac{S}{\bar{X}}, \qquad (2)$$

where S is the sample standard deviation, the square root of the unbiased estimator of the variance, and $\bar{X}$ is the sample mean.

The coefficient of variation has long been a widely used descriptive and inferential quantity in many applications of science, economics and others. In chemical experiments, the coefficient of variation is often used as a yardstick of precision for measurements. For example, two measurement methods may be used to compare precision on the

basis of their respective coefficients of variation. The coefficient of variation can be used to measure relative risks (Miller and Karson, 1977) in finance and actuarial science. Furthermore, testing the equality of the coefficients of variation for two stocks can help determine whether the two stocks possess the same risk or not. In physiology, the coefficient of variation can also be applied to assess homogeneity of bone test samples (Hamer et al., 1995). In the field of safety engineering, the coefficient of variation is used as a tool in the uncertainty of fault trees analysis (Ahn, 1995). Additionally, the coefficient of variation is also employed in assessing the strength of ceramics (Gong and Li, 1999).

Although the point estimator of coefficient of variation can be a useful measure, the greatest use of it is to construct a confidence interval of coefficient of variation for the quantity of interest. (Mahmoudvand and Hassani, 2009), since a confidence interval provides much more information about the population value of the quantity of interest than does a point estimate (e.g., Smithson, 2001; Thompson, 2002; Steiger, 2004).

An approximate $(1-\alpha)100\%$ confidence interval for the coefficient of variation (see e.g. Vangel, 1996) is given by

$$CI=\left\{\frac{\hat{\kappa}}{\sqrt{t_1(\theta_1\hat{\kappa}^2+1)-\hat{\kappa}^2}},\frac{\hat{\kappa}}{\sqrt{t_2(\theta_2\hat{\kappa}^2+1)-\hat{\kappa}^2}}\right\}, \quad (3)$$

where $\nu=n-1$, $t_1\equiv\chi^2_{\nu,1-\alpha/2}/\nu$, $t_2\equiv\chi^2_{\nu,\alpha/2}/\nu$ and $\theta=\theta(\nu,\alpha)$ is a known function selected so that a random variable $W_\nu\equiv Y_\nu/\nu$, where Y_ν has a χ^2_ν distribution, has approximately the same distribution as a pivotal quantity $Q\equiv\frac{\hat{\kappa}^2(1+\kappa^2)}{(1+\theta\hat{\kappa}^2)\kappa^2}$. This pivotal quantity can be used to construct hypothesis tests and a confidence interval for κ.

McKay (1932) proposed that the choice $\theta=\frac{\nu}{\nu+1}$ gives a good approximation for the confidence interval in equation (3), but he was unable to investigate the small-sample distribution of Q. McKay's approximate confidence interval is

$$CI_{01}=\left\{\hat{\kappa}\left[\left(\frac{\chi^2_{\nu,1-\alpha/2}}{\nu+1}-1\right)\hat{\kappa}^2+\frac{\chi^2_{\nu,1-\alpha/2}}{\nu}\right]^{-1/2},\right.$$

$$\left.\hat{\kappa}\left[\left(\frac{\chi^2_{\nu,\alpha/2}}{\nu+1}-1\right)\hat{\kappa}^2+\frac{\chi^2_{\nu,\alpha/2}}{\nu}\right]^{-1/2}\right\}, \quad (4)$$

where $\nu=n-1$ is the degrees of freedom of the χ^2 distribution. Several authors have carried out numerical investigations of the accuracy of McKay's confidence interval. For instance, Iglewicz and Myers (1970) had compared McKay's confidence interval with the exact confidence interval based on the noncentral t distribution and they found that McKay's confidence interval is efficient for $n\geq 10$ and $0<\kappa<0.3$.

Vangel (1996) proposed a new confidence interval for the coefficient of variation which he called the modified McKay's confidence interval. He proposed the use of the function θ where $\theta=\frac{\nu}{\nu+1}\left[\frac{2}{\chi^2_{\nu,\alpha}}+1\right]$. He also suggested that the modified McKay method gave confidence intervals for the coefficient of variation that are closely related to the McKay's confidence interval but they are usually more accurate. The modified McKay's confidence interval for a coefficient of variation is given by

$$CI_{02}=\left\{\hat{\kappa}\left[\left(\frac{\chi^2_{\nu,1-\alpha/2}+2}{\nu+1}-1\right)\hat{\kappa}^2+\frac{\chi^2_{\nu,1-\alpha/2}}{\nu}\right]^{-1/2},\right.$$

$$\hat{\kappa}\left[\left(\frac{\chi^2_{\nu,\alpha/2}+2}{\nu+1}-1\right)\hat{\kappa}^2+\frac{\chi^2_{\nu,\alpha/2}}{\nu}\right]^{-1/2}\Bigg\}. \quad (5)$$

When data are normally distributed, McKay's confidence interval and the modified McKay's confidence interval, CI_{01} and CI_{02}, can be used very well in terms of coverage probability and length. However, for non-normal data, these confidence intervals cannot be used practically. The aim in this paper is to construct the new confidence intervals for the coefficient of variation of the Poisson distribution. The modified confidence intervals for the coefficient of variation are obtained from applying confidence intervals for the Poisson mean. Additionally, the coverage probabilities and the lengths of new confidence intervals for a coefficient of variation are compared through a Monte Carlo simulation study.

The paper is organized as follows. In the next section, new confidence intervals for the coefficient of variation of a Poisson distribution are presented. Simulation results obtained from the Monte Carlo simulation and discussions are shown in the third section. The conclusions are presented in the final section.

New Confidence Intervals for the Coefficient of Variation of a Poisson Distribution

In this section the new confidence intervals for the coefficient of variation of a Poisson distribution are presented. Newly proposed confidence intervals are based on confidence intervals for the Poisson mean. Suppose $X_i \sim Poi(\lambda)$, $i=1,2,...,n$. Hence, the population coefficient of variation for a Poisson distribution is given by

$$\kappa = \frac{\sigma}{\mu} = \frac{\sqrt{\lambda}}{\lambda} = \frac{1}{\sqrt{\lambda}}.$$

In order to construct new confidence intervals, there are first mentioned confidence intervals for the Poisson mean. These confidence intervals considered are: (Barker, 2002)

(1) Wald (W) confidence interval. The W confidence interval is derived from the asymptotic standard normal distribution of $(\bar{X}-\lambda)/\sqrt{\bar{X}/n}$. This quantity can be inverted to provide the confidence interval

$$\left(\bar{X}-Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{X}}{n}},\ \bar{X}+Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{X}}{n}}\right). \quad (6)$$

(2) Wald with continuity correction (WCC) confidence interval. The W confidence interval uses a continuous distribution (normal) to approximate a discrete distribution (Poisson). A continuity correction might make this approximation more accurate. The WCC confidence interval is given by

$$\left(\bar{X}-Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{X}+0.5}{n}},\ \bar{X}+Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{X}+0.5}{n}}\right). \quad (7)$$

(3) Scores (S) confidence interval. The S confidence interval is derived from the asymptotic standard normality of $(\bar{X}-\lambda)/\sqrt{\lambda/n}$. This quantity can be inverted to provide the S confidence interval

$$\left(\bar{X}+\frac{(Z_{1-\alpha/2})^2}{2n}-Z_{1-\frac{\alpha}{2}}\sqrt{\frac{4\bar{X}+\frac{(Z_{1-\alpha/2})^2}{n}}{4n}},\right.$$

$$\left.\bar{X}+\frac{(Z_{1-\alpha/2})^2}{2n}+Z_{1-\frac{\alpha}{2}}\sqrt{\frac{4\bar{X}+\frac{(Z_{1-\alpha/2})^2}{n}}{4n}}\right). \quad (8)$$

(4) Variance stabilizing (VS) confidence interval. The quantity $(\sqrt{X}-\sqrt{\lambda})/\sqrt{1/4n}$ is the asymptotically standard normal. This can be inverted into the confidence interval

$$\left(\bar{X} + \frac{(Z_{1-\alpha/2})^2}{4n} - Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\bar{X}}{n}}, \right.$$

$$\left. \bar{X} + \frac{(Z_{1-\alpha/2})^2}{4n} + Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\bar{X}}{n}} \right), \qquad (9)$$

where $\bar{X} = n^{-1} \sum_{i=1}^{n} X_i$ and $Z_{1-\frac{\alpha}{2}}$ is a $\left(1-\frac{\alpha}{2}\right)th$ quantile of the standard normal distribution. From Equations (6)-(9), we therefore can derive the confidence intervals for a coefficient of variation of a Poisson distribution based on the above confidence intervals for the Poisson mean as follows:

$$\begin{aligned} 1-\alpha &= P\left(L_i < \lambda < U_i\right) \\ &= P\left(\sqrt{L_i} < \sqrt{\lambda} < \sqrt{U_i}\right) \\ &= P\left(\frac{1}{\sqrt{U_i}} < \frac{1}{\sqrt{\lambda}} < \frac{1}{\sqrt{L_i}}\right) \\ &= P\left(\frac{1}{\sqrt{U_i}} < \kappa < \frac{1}{\sqrt{L_i}}\right). \end{aligned} \qquad (10)$$

where L_i and U_i, $i = 1,2,3,4$ denote the lower and upper limit of confidence intervals for the Poisson mean based on W, WCC, S, and VS, respectively.

Hence, we obtain $(1-\alpha)100\%$ four new confidence intervals for the coefficient of variation of a Poisson distribution which are

$$CI_1 = \left\{ \left(\sqrt{\bar{X} + Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\bar{X}}{n}}} \right)^{-1}, \left(\sqrt{\bar{X} - Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\bar{X}}{n}}} \right)^{-1} \right\}, \qquad (11)$$

$$CI_2 = \left\{ \left(\sqrt{\bar{X} + Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\bar{X}+0.5}{n}}} \right)^{-1}, \left(\sqrt{\bar{X} - Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\bar{X}+0.5}{n}}} \right)^{-1} \right\}, \qquad (12)$$

$$CI_3 = \left\{ \left(\sqrt{\bar{X} + \frac{(Z_{1-\alpha/2})^2}{2n} + Z_{1-\frac{\alpha}{2}} \sqrt{\frac{4\bar{X} + \frac{(Z_{1-\alpha/2})^2}{n}}{4n}}} \right)^{-1}, \left(\sqrt{\bar{X} + \frac{(Z_{1-\alpha/2})^2}{2n} - Z_{1-\frac{\alpha}{2}} \sqrt{\frac{4\bar{X} + \frac{(Z_{1-\alpha/2})^2}{n}}{4n}}} \right)^{-1} \right\}, \qquad (13)$$

$$CI_4 = \left\{ \left(\sqrt{\bar{X} + \frac{(Z_{1-\alpha/2})^2}{4n} + Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\bar{X}}{n}}} \right)^{-1}, \left(\sqrt{\bar{X} + \frac{(Z_{1-\alpha/2})^2}{4n} - Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\bar{X}}{n}}} \right)^{-1} \right\}, \qquad (14)$$

where $CI_i, i = 1,2,3,4$ denote the confidence intervals for the coefficient of variation of a Poisson distribution based on W, WCC, S, and VS confidence interval, respectively.

To study the different confidence intervals, we consider their coverage probability and length. For each of the methods considered, we obtain a $(1-\alpha)100\%$ confidence interval denoted by

(L,U) (based on M replicates) and estimated the coverage probability and the length, respectively, by

$$\widehat{1-\alpha} = \frac{\#(L \le \kappa \le U)}{M},$$

and $$\widehat{Length} = \frac{\sum_{j=1}^{M}(U_j - L_j)}{M}.$$

Results and Discussions

In this section, the performance of the estimated coverage probabilities of the new asymptotic confidence intervals (11), (12), (13) and (14) and their lengths was examined via Monte Carlo simulations. Data are generated from the Poisson distribution with κ = 0.1, 0.2 and 0.3 sample sizes; n = 10, 15, 25, 50 and 100. All simulations were performed using programs written in the R statistical software (The R Development Core Team, 2009a, 2009b) with the number of simulation runs, M = 50,000 at the level of significance $\alpha = 0.05$ and 0.10. The simulation results are shown in Tables 1 and 2. The following information is presented here: the estimated coverage probabilities of the confidence intervals, CI_1, CI_2, CI_3 and CI_4, and their lengths for a Poisson distribution at α = 0.05 and 0.10, respectively. As can be seen from Tables 1 and 2, the confidence interval based on WCC, CI_2, has a closeness coverage probability of $1-\alpha$ for all sample sizes and values of κ except when $n = 10$, $\kappa = 0.2$ and $\alpha = 0.10$. The other three confidence

Table 1 The estimated coverage probabilities and lengths of a 95% confidence interval in (11), (12), (13) and (14) for a Poisson distribution.

n	κ	Coverage probabilities				Lengths			
		CI_1	CI_2	CI_3	CI_4	CI_1	CI_2	CI_3	CI_4
10	0.1	0.9506	0.9521	0.9488	0.9508	0.0062	0.0062	0.0062	0.0062
	0.2	0.9502	0.9528	0.9468	0.9510	0.0251	0.0254	0.0249	0.0250
	0.3	0.9444	0.9564	0.9481	0.9471	0.0576	0.0590	0.0563	0.0568
15	0.1	0.9505	0.9521	0.9491	0.9491	0.0051	0.0051	0.0051	0.0051
	0.2	0.9506	0.9529	0.9491	0.9481	0.0204	0.0206	0.0203	0.0204
	0.3	0.9473	0.9559	0.9484	0.9488	0.0465	0.0476	0.0458	0.0461
25	0.1	0.9494	0.9506	0.9481	0.9493	0.0039	0.0039	0.0039	0.0039
	0.2	0.9499	0.9518	0.9471	0.9495	0.0158	0.0159	0.0157	0.0157
	0.3	0.9512	0.9537	0.9475	0.9478	0.0357	0.0365	0.0354	0.0355
50	0.1	0.9491	0.9499	0.9490	0.9490	0.0028	0.0028	0.0028	0.0028
	0.2	0.9511	0.9530	0.9512	0.9514	0.0111	0.0112	0.0111	0.0111
	0.3	0.9470	0.9515	0.9478	0.9470	0.0251	0.0257	0.0250	0.0250
100	0.1	0.9503	0.9508	0.9494	0.9501	0.0020	0.0020	0.0020	0.0020
	0.2	0.9507	0.9533	0.9498	0.9510	0.0079	0.0079	0.0078	0.0078
	0.3	0.9515	0.9546	0.9511	0.9512	0.0177	0.0181	0.0177	0.0177

intervals, CI_1, CI_3, and CI_4, give slightly lower coverage probabilities than $1-\alpha$. The estimated coverage probabilities of the CI_1 and CI_2 increase as the values of κ get larger (i.e. for CI_2, n=10 and $\alpha = 0.05$, 0.9521 for κ = 0.1; 0.9528 for κ = 0.2; 0.9564 for κ = 0.3). The lengths of all confidence intervals are slightly different. Further, the lengths increase as the values of κ get larger (i.e. for CI_2, n=10 and $\alpha = 0.05$, 0.0062 for κ = 0.1; 0.0254 for κ = 0.2; 0.0590 for κ = 0.3). Moreover, when the sample sizes increase, the lengths are shorter (i.e. for CI_2, κ=0.1 and $\alpha = 0.05$, 0.0062 for n=10; 0.0051 for n=15; 0.0039 for n=25; 0.0028 for n =50; 0.0020 for n=100).

Conclusions

Four new confidence intervals for the coefficient of variation of the Poisson distribution have been developed. The proposed confidence intervals are compared through a Monte Carlo simulation study. The new confidence intervals are based on a confidence interval for the Poisson mean. The confidence interval based on WCC has closeness coverage probabilities $1-\alpha$. In addition, the lengths of all of the confidence intervals are slightly different. Therefore, if a confidence interval with a closeness coverage probability equal to a pre-specified value is preferred, the confidence interval based on WCC is preferable to the other three confidence intervals.

Table 2 The estimated coverage probabilities and lengths of a 90% confidence interval in (11), (12), (13) and (14) for a Poisson distribution.

n	κ	Coverage probabilities				Lengths			
		CI_1	CI_2	CI_3	CI_4	CI_1	CI_2	CI_3	CI_4
10	0.1	0.9003	0.9003	0.9037	0.9006	0.0052	0.0052	0.0052	0.0052
	0.2	0.8980	0.8980	0.9052	0.8984	0.0210	0.0213	0.0209	0.0210
	0.3	0.8895	0.9104	0.9041	0.9023	0.0480	0.0491	0.0472	0.0475
15	0.1	0.9007	0.9007	0.8985	0.9013	0.0043	0.0043	0.0042	0.0043
	0.2	0.9009	0.9009	0.8967	0.9018	0.0171	0.0173	0.0170	0.0171
	0.3	0.9033	0.9033	0.8967	0.8958	0.0388	0.0397	0.0384	0.0386
25	0.1	0.9000	0.9024	0.9026	0.9004	0.0033	0.0033	0.0033	0.0033
	0.2	0.8961	0.9001	0.9011	0.8963	0.0132	0.0133	0.0132	0.0132
	0.3	0.8999	0.9048	0.9010	0.8946	0.0299	0.0306	0.0297	0.0298
50	0.1	0.8986	0.9002	0.9007	0.8991	0.0023	0.0023	0.0023	0.0023
	0.2	0.9000	0.9058	0.9030	0.9003	0.0093	0.0094	0.0093	0.0093
	0.3	0.8958	0.9042	0.9007	0.9003	0.0210	0.0215	0.0210	0.0210
100	0.1	0.8992	0.9004	0.8992	0.8992	0.0016	0.0016	0.0016	0.0016
	0.2	0.8981	0.9024	0.9007	0.8987	0.0066	0.0067	0.0066	0.0066
	0.3	0.9004	0.9065	0.8977	0.9008	0.0148	0.0152	0.0148	0.0148

Acknowledgements

The author is very appreciate of valuable insights and comments provided by two anonymous referees, leading to many improvements in the paper. The author also wishes to thank Sheldon Daniels for his careful reading.

References

Ahn, K. (1995) Use of coefficient of variation for uncertainty analysis in fault tree analysis. *Reliability Engineering and System Safety* 47: 229-230.

Barker, L. (2002) A comparison of nine confidence intervals for a Poisson parameter when the expected number of events is ≤ 5. *American Statistician* 56: 85-89

Gong, J. and Li, Y. (1999) Relationship between the estimated Weibull modulus and the coefficient of variation of the measured strength for ceramics. *Journal of the American Ceramic Society* 82: 449-452.

Hamer, A.J., Strachan, J.R., Black, M.M., Ibbotson, C., and Elson, R.A. (1995) A new method of comparative bone strength measurement. *Journal of Medical Engineering and Technology* 19: 1-5.

Iglewicz, B. and Myers, R.H. (1970) Comparisons of approximations to the percentage points of the sample coefficient of variation. *Technometrics* 12: 166-169.

Kelley, K. (2007) Sample size planning for the coefficient of variation from the accuracy in parameter estimation approach. *Behavior Research Methods* 39: 755-766.

Mahmoudvand, R. and Hassani, H. (2009) Two new confidence intervals for the coefficient of variation in a normal distribution. *Journal of Applied Statistics* 36: 429-442.

McKay, A.T. (1932) Distribution of the coefficient of variation and the extended t distribution. *Journal of the Royal Statistics Society Series B* 95: 695-698.

Miller, E.G. and Karson, M.J. (1977) Testing the equality of two coefficients of variation. *American Statistical Association: Proceedings of the Business and Economics Section* Part I, 278-283.

Smithson, M. (2001) Correct confidence intervals for various regression effect sizes and parameters: The importance of noncentral distributions in computing intervals. *Educational and Psychological Measurement* 61: 605-632.

Steiger, J.H. (2004) Beyond the F test: Effect size confidence intervals and tests of close fit in the analysis of variance and contrast analysis. *Psychological Methods* 9: 164-182.

The R Development Core Team. (2009a) An introduction to R. Vienna: R Foundation for Statistical Computing. [Online://cran.r-project.org/doc/manuals/R-intro.pdf] accessed on October 16, 2009.

The R Development Core Team. (2009b) R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. [Online: URL:// cran.r-project.org/doc/manuals/refman.pdf] accessed on October, 16, 2009.

Thompson, B. (2002) What future quantitative social science research could look like: Confidence intervals for effect sizes. *Educational Researcher* 31: 25-32.

Vangel, M.G. (1996) Confidence intervals for a normal coefficient of variation. *American Statistician* 50: 21-26.

Research Article

Nutritive Values of Whip Grass (*Hemarthria compressa*) at Different Cutting Intervals Consumed by Thai Indigenous Cattle

Jeerasak Chobtang[1], Apichat Boonruangkao[2], Saksan Suankool[2] and Auraiwan Isuwan[3*]

[1]*Animal Nutrition Division, Department of Livestock Development, Ratchathewi, Bangkok, Thailand*
[2]*Suratthani Animal Nutrition Research and Development Center, Tha Chang, Suratthani, Thailand*
[3]*Faculty of Animal Science and Agricultural Technology, Silpakorn University, Petchaburi IT Campus, Cha-Am, Petchaburi, Thailand*
[*]*Corresponding author. E-mail address: auraiwan_i@hotmail.com*

Received July 6, 2010; Accepted December 17, 2010

Abstract

This study aimed to evaluate the effect of cutting intervals of *Hemarthria compressa* on feed intake, nutrient digestibility, and energy concentration of the grass consumed by Thai indigenous cattle. A 3 × 4 incomplete Latin Square Design was used. Treatments were 3 cutting intervals (30, 45 and 60 days) of grass fed to the bulls as green forage. Four Thai indigenous bulls were allocated into the experimental treatments. Total collection method was used. Feed intake, apparent nutrient digestibility and energy concentration of grass were subsequently assessed.

The result showed that it was a linear decline ($p<0.05$) in crude protein (CP) and ash contents as cutting intervals increased while there were linearly increased ($p<0.05$) in dry matter (DM), organic matter, neutral detergent fiber, and acid detergent lignin contents. As grass maturation increase, CP intake was decreased but not for DM intake (79.11-81.90 g/$BW^{0.75}$/d). There was no significant difference of the digestibility coefficient of any nutrient among cutting intervals except gross energy. Nevertheless, the metabolizable energy concentration of the grass were significantly decreased (linear, $p<0.05$) in accordance with increasing cutting intervals. In conclusion, harvesting the grass at 30 days of regrowth gave better nutritive values and suitable to use as livestock feed.

***Key Words*:** Hemarthria compressa; Maturation; Nutritive value; Thai indigenous cattle

Introduction

Whip grass (*Hemarthria compressa*) is a perennial plant having culms decumbent to long-stoloniferous roots at the lower nodes, up to 1 m or more. As reported by Bixing and Phillips (2006), the grass is branches from base with the conspicuous, dark and glabrous nodes. Leaf sheaths are loose, compressed, keeled, glabrous and hairy along mouth. Leaf blades are linear, 2–15 x 0.2–0.5 cm, base rounded and apex sub-acute. Racemes are solitary and lightly compressed with the length of 2–10 cm. Its shape is articulation line oblique and tardily disarticulating. Sessile spikelet is slightly longer than adjacent internodes with the length of 3–5 mm. Callus are broadly triangular (with the size of 0.5–1 mm) with lower glume narrowly

oblong, leathery, flat or sub-convex on back, abruptly constricted into obtuse and emarginated apex. The upper glume is adnate to rachis and thin. This grass has a wide range of habitat, including its capability to thrive in the marshes and wet land areas.

In Thailand, this grass was mostly found in the low-land area in the southern region. Insung et al. (2005) reported that *Hemarthria compressa* was the most popular grass that the farmer used to feed the fighting bulls. Waipanya et al. (2005) reported that the grass yielded about 9 to 10 tons $ha^{-1}year^{-1}$ of dry matter with a range of 8 to 10% of crude protein (CP). When the ruminal degradability of this grass was studied using the nylon bag technique, it was found that the potential degradability (PD) of the dry matter (DM) varied from 72.5–87.8% (Insung et al., 2005). Additionally, this grass was palatable to goats (Baoli and Shilin, 1997).

In term of nutrition, forage quality and utilization could be influenced not only by forage species but also by the plant maturity or growth stage of the grass (Skerman and Riveros, 1990). There are reports which showed that increasing dry matter yield, nutritive compositions and feed quality of the grass had decreased as the grass had undergone maturation (Chobtang et al., 2008; Arthington and Brown, 2005; Kamalak et al., 2005).

In term of cultural practice and pasture management, maintaining the appropriate stage of forage is a good option for improving ruminant animal productivity. Although there is a general convention that optimal nutritive values of grass suitable to feed the animals are dependent upon growth stage of the grass at harvesting, there is no scientific data on the nutritive values of whip grass in accordance with its stages of growth.

Therefore, the objective of this study aims to evaluate the relationship of nutritive values of the whip grass and its cutting intervals.

Materials and Methods

Location and climate data

The study was conducted at Suratthani Animal Nutrition Research and Development Center, Tha Chang, Thailand. The soil characteristics were loamy-skeletal, mixed, semiactive and isohyperthermic Typic Hapludults. In general, the temperature during the experiment varied between 26–34 ˚C, with 68% relative humidity and 1,110 mm precipitation.

Experimental plan

A 3 × 4 Incomplete Latin Square Design (LSD) was used in this experiment. Sources of variance were the cutting intervals (3 stages of cutting at 30, 45 and 60 days of regrowth), the periods of the experiment (3 periods) and the animal (4 Thai indigenous bulls). The experiment was started from July 2008 and ended on September 2008.

Animal management

Four Thai-indigenous bulls, approximately 1.5 years of age, average body weight of 202±21.89 kg, were used. All bulls were vaccinated against foot and mount disease and hemorrhagic septicemia. Vitamin A, D_3, and E were injected to ensure that the animals had sufficient vitamins and anthelmintic and ascaric druge were administered to make the bulls free of external parasite, two weeks before the experiment started. The bulls were then fed with Ruzi grass (*Brachiaria ruziziensis*) hay (9% CP) ad libitum prior to the experiment started and also at the resting period (21 days) between each experimental period. The animals were kept in an individual confinement of 2 × 2.5 m and randomly assigned to the treatments. Clean water and mineral blocks were also given to the animals during the experiment.

Pasture management and treatments

A 1.6-ha pasture plot was carefully chosen from a 7-ha of 2-year old stand of *Hemarthria compressa* pasture. The chosen field was cut using a drum mower machinery at 3-cm stable height for a uniformly growth, the residue was removed and

the pasture was fertilized with N-P-K fertilizer (15-15-15) at 8 $kgha^{-1}$. The pasture was irrigated using sprinkle every 7 days and allowed to regrowth for 45 days.

Consequently, the pasture was divided into 4 main plots for each experimental period (green grass from one plot used for one animal); each plot was then divided into 21 subplots. Areas of the subplots were 80, 60 and 40 m^2 for 30, 45 and 60 days plot, respectively. The first subplot of 30, 45 and 60 days of regrowth age was cut at 31, 46 and 61 days, respectively before the beginning of the experiment. As a result, there were a grass with 30, 45 and 60 days of regrowth at the beginning of the experiment and other subplots would be reached the expected date at the next day of the experiment.

Intake and total tract digestion study

In this study, total collection method was used. Each experimental period, dry matter and nutrient intake, and total tract digestion of nutrient measurements were conducted for 21 days in which the first 14-day period was a preliminary period and the last 7 days were the collection period. Grass was cut in the morning, kept under shed and fed to the animal ad libitum according to the treatments twice a day, 08.30 and 16.00 h. In order to eliminate the carrying effects, the animals were initially fed with a Ruzi grass (*Brachiaria ruziziensis*) hay for 21 days prior to the beginning of the next experimental period. At the collection period, feeds, feed refusals and feces were weighed and then, 10% of feed refusals and feces were sampled for each day. Composites of feeds, feed refusals and feces were sub-sampled at approximately 1,000 g from each bull. Samples were dried at 65 ˚C for 72 h and ground to pass a 1 mm screen using Wiley mill. The samples were then kept at -21˚C freezer for subsequent chemical analysis.

Chemical analysis

Dry matter (DM), organic matter (OM), crude protein (CP), ether extract (EE) and ash components of composited samples of feeds, feed refusals and feces from each bull and period were analyzed using method described by AOAC (1990). Cell wall components (neutral detergent fiber (NDF) and acid detergent fiber (ADF) and acid detergent lignin (ADL)) were analyzed using method described by Van Soest et al. (1991). Gross energy of all samples was determined using adiabatic bomb calorimeter. Metabolizable energy (ME) content of feeds was estimated using the equation of ME = 0.82*DE as proposed by NRC (1996).

Statistical analysis

The nutritive values and chemical composition of the grass at different cutting intervals, and intake and apparent nutrient digestibility of grass consumed by Thai indigenous bulls were analyzed as a 3 × 4 Incomplete Latin Square Design. Pre-planned orthogonal polynomial (linear and quadratic) was then statistically tested (Muller and Fetterman, 2003).

Results and Discussion

Chemical composition of grass

Chemical compositions of the grass are presented in Table 1. There was a linear increase ($p<0.05$) in DM, OM, NDF, and ADL concentrations, with increasing maturity of grass. However, the concentrations of CP, ash and P were decreased linearly ($p<0.05$), with increasing maturity. Crude protein content decreased 17.11 and 33.69% when cutting intervals increases from 30 days to 45 and to 60 days, respectively. This result was similar to the report of Arthington and Brown (2005) who found that crude protein of tropical grass was reduced 37.80% when regrowth age was extended from 28 to 70 days. Even though CP content of the grass in this study was linearly declined throughout the period of regrowth, the final concentration was not lower than the lowest level which affected the ruminal microbial activity (Minson and Wilson, 1980).

Table 1 Least square means of chemical compositions (dry matter basis) of *Hemarthria compressa*.

Chemical composition	Cutting intervals (days)			SEM	Contrast	
	30	45	60		linear	quadratic
DM (%)	14.05	15.78	17.37	0.53	**	NS
OM	90.78	92.33	92.89	0.35	**	NS
CP	13.21	10.95	8.76	0.63	***	NS
EE	1.60	1.46	1.36	0.12	NS	NS
Ash	9.22	7.67	7.11	0.35	**	NS
NDF	69.90	72.46	72.38	0.41	**	NS
ADF	35.50	36.56	37.14	0.50	NS	NS
ADL	4.57	5.07	5.70	0.16	***	NS
GE ($kcalkg^{-1}DM$)	4,618	4,507	4,403	37	*	NS

-p<0.01, *-p<0.0001, NS-non significant difference (p>0.05).

Even there was a significant change (p<0.05) of NDF content by grass maturity but it was narrowly varied from 69.90 to 72.46% DM**.** The ADF content was not affected (p>0.05) by cutting intervals and its value varied marginally from 35.50 to 37.14% DM. This finding was in agreement to the reports of Chobtang et al. (2008), Boval et al. (2007), Isuwan et al. (2007), Arthington and Brown (2005) and Achemede et al. (2000). These authors summarized that the early maturation of tropical grass was normally found. The high temperature condition and excess of sun light might be the main factors affecting this aspect (Wilson and Minson, 1980).

Table 2 Least square means of intake characteristics of Thai indigenous bulls consumed *Hemarthria compressa.*

Intake		Cutting intervals (days)			SEM	Contrast	
		30	45	60		linear	quadratic
Dry	Matter						
	%BW	2.12	2.08	2.05	0.05	NS	NS
	Kgd^{-1}	4.73	4.70	4.53	0.12	NS	NS
	$gMBW^{-1}d^{-1}$	81.90	80.61	79.11	2.03	NS	NS
Crude	Protein						
	Kgd^{-1}	0.64	0.54	0.43	0.04	**	NS
	$gMWB^{-1}d^{-1}$	11.07	9.20	7.52	0.64	**	NS

** - p<0.01, NS-non significant difference (p>0.05).

Dry matter and crude protein intake

The values of dry matter and crude protein intake of Thai indigenous bulls fed with green *Hemarthria compressa* are shown in Table 2. There was no significant difference ($p>0.05$) in dry matter intake (DMI) among three cutting intervals (averaging of 2.08 %BW, 4.65 kgd^{-1} or 80.54 g metabolic body weight $(MBW)^{-1}d^{-1}$). Even though the bulls were ad libitum fed, the bulls had consumed green grass (DM basis) in average of 2.08% BW. This value, however, was higher when comparing to the dry matter intake of cattle receiving the other tropical grasses (Burns and Fisher, 2007). Jung and Allen (1995), Van Soest (1994) and Lippke (1980) reported that high fiber content was a major factor affecting DMI of ruminant animals.

Crude protein intake was affected by cutting intervals. It was linearly declined ($p<0.05$) as grass maturity increases. This was caused by the decline in CP content of grass with increasing maturity (Table 2).

Nutrient digestibility and energy utilization efficiency

The nutrient digestibility of the grass is presented in Table 3. There was no significant relationship ($p>0.05$) between the nutrient digestibility and cutting intervals of the grass, except the gross energy. Digestibility of gross energy decreased linearly ($p<0.05$) as cutting intervals increased. This may be caused by the increasing content of cell wall components and ADL value, and also the decreasing level of CP and GE content of grass (Table 1). Archimede et al. (2000) also reported that even harvested at 14 days of regrowth, *Digitaria decumbens* was digested at the acceptable level, leading to the availability of nutrient for the animals.

The values of energy concentration are shown in Table 4. Gross energy (GE), digestible energy (DE) and metabolizable energy (ME) concentrations of the grass decreased linearly ($p<0.05$) as the cutting intervals increased. The DE content of grass harvested at 30 days of regrowth was considered at stage of a good quality (Pond et al., 1995). At the energy values ranging from 2.53 to 2.65 $Mcalkg^{-1}DM$, the cattle can consume the grass at the level of 2.0 to 2.5% of body weight (BW). The ME of 30 days of regrowth was comparable to the temperate grass reported by NRC (1996).

Table 3 Least square means of nutrient digestibility of *Hemarthria compressa* consumed by Thai indigenous bulls.

Digestibility	Cutting intervals (days)				Contrast	
(%)	30	45	60	SEM	linear	quadratic
DM	68.41	67.16	65.69	1.38	NS	NS
OM	69.44	67.98	67.04	1.44	NS	NS
CP	69.50	69.09	64.02	2.03	NS	NS
EE	64.67	66.45	65.83	4.09	NS	NS
NDF	69.18	68.56	66.23	1.51	NS	NS
ADF	54.14	54.69	52.08	2.30	NS	NS
GE	55.20	50.88	46.22	2.28	*	NS

*-$p<0.05$, NS-non significant difference ($p>0.05$).

Table 4 Least square means of energy concentration of *Hemarthria compressa* consumed by Thai indigenous bulls.

Energy concentration	Cutting intervals (days)				Contrast	
	30	45	60	SEM	linear	quadratic
GE (kcalkg^{-1}DM)	4,618	4,507	4,403	37	*	NS
DE (kcalkg^{-1}DM)	2,591	2,311	2,047	116	*	NS
ME (kcalkg^{-1}DM)[1]	2,125	1,895	1,679	96	*	NS

[1]ME = 0.82*DE; *-$p<0.05$; NS-non significant difference ($p>0.05$).

Conclusion

The results of this study showed that even though whip grass (*Hemarthria compressa*) grown in tropical condition (Thailand) had declining values of nutritive composition such as organic matter and crude protein content, crude and digestible protein intake, this decline did not affect the nutritive values, such as dry matter intake and nutrient digestibility were not affected as the cutting intervals increased. However, the gross, digestible and metabolizable energy of the grass decreased as the cutting intervals increased. In conclusion, grass harvested at 30 days of regrowth had a quality which was good enough for feeding the animals.

Acknowledgments

This project was funded by National Research Council of Thailand. The authors would like to thank the Director of Suratthani Animal Nutrition Research and Development Center for all convenient providing. The staffs of Narathiwat Animal Nutrition Research and Development are also greatly acknowledged for their help in chemical analysis.

References

AOAC. (1990) *Official Methods of Analysis*. 15th ed. Association of Official Analytical Chemists, Washington, D. C.

Archimede, H., Boval, M., Alexandre, G., Xande, A., Aumont, G., and Poncet, C. (2000) Effect of regrowth age on intake and digestion of *Digitaria decumbens* consumed by Black-belly sheep. *Animal Feed Science and Technology* 87, 153–162.

Arthington, J. D. and Brown, W. F. (2005) Estimation of feeding value of four tropical forage species at two stages of maturity. *Journal of Animal Science* 83: 1726–1731.

Baoli, W. and Shilin, W. (1997) Goat production from forage at Mengongshan farm, Hunan. ***In*:** Forages for the Red Soils Area of China. J.M. Scott, D.A. MacLeod, Minggang Xu and A.J. Casanova, eds. *Proceedings of an International Workshop*, Jianyang, Fujian Province, P. R. China, October 6–9, 1997. 169-171 pp.

Bixing, S. and Phillips, S. M. (2006) Hemarthria. *Flora of China* 22: 640–642.

Boval, M., Archimede, H., Cruz, P., and Duru, M. (2007) Intake and digestibility in heifers grazing a *Dichanthium* spp. dominated pasture at 14 and 28 days of regrowth. *Animal Feed Science and Technology* 134: 18–31.

Burns, J. C. and Fisher, D. S. (2007) Dry matter intake and digestibility of 'Coastal', 'Tifton 44', and 'Tifton 85' Bermudagrass hays grown in the U.S. upper South. *Crop Science* 47: 795–810.

Chobtang, J., Prajakboonjetsada, S., Watananawin, S., and Isuwan, A. (2008) Change in dry matter and nutritive composition of *Brachiaria humidicola* grown in Ban Thon soil series. *Maejo International Journal of Science and Technology* 2: 551-558.

Insung, O., Vearasilp, T., and Meulen, U. T. (2005) Species diversity and the ruminal dry matter degradability of grasses fed to fighting bulls in southern Thailand. ***In*: *Tropentag 2005 International Research on Food Security***, Natural Resource Management and Rural Development the Global Food & Product Chain. E. Tielkes, C. Hulsebusch, I. Hauser, A. Deininger and K. Becker, eds. University of Hohenheim, Stuttgart, October 11-13, 2005. 179 pp.

Isuwan, A., Saelim, J., and Poathong, S. (2007) Effects of levels of sulfur fertilizer on growth of *Digitaria eriantha* grass. *Silpakorn University Science and Technology Journal* 1:13-19.

Jung, H. G. and Allen, M. S. (1995) Characteristics of plant cell wall affecting intake and digestibility of forages by ruminants. *Journal of Animal Science* 73: 2774–2790.

Kamalak, A., Canbolat, O., Gurbuz, Y., Erol, A., and Ozay, O. (2005) Effect of maturity stage on chemical composition, in vitro and in situ dry matter degradation of tumbleweed hay (*Gundelia tournefortii* L.). *Small Ruminant Research* 58: 149-156.

Lippke, H. (1980) Forage characteristics related to intake, digestibility and gain by ruminants. *Journal of Animal Science* 50: 223–230.

Minson, D. J. and Wilson, J. R. (1980) Comparative digestibility of tropical and temperate forage – a contrast between grasses and legumes. *Journal of the Australian Institute of Agricultural Science* 46: 247-249.

Muller, K. E. and Fetterman, B. A. (2003) *Regression and ANOVA: An integrated approach using SAS software*. Joinly-copublished by John Wiley and Sons Inc. and SAS Institute Inc. Cary, NC, USA. 566 pp.

NRC. (1996) *Nutrient Requirements of Beef Cattle*. 7th ed. National Academic Press, Washington, D. C.

Pond, W. G., Church, D. C., and Pond, K. R. (1995) *Basic Animal Nutrition and Feeding*. 4th ed. John Wiley and Sons Inc. NY. 615 pp.

Skerman, P. J. and Riveros, F. (1990) *Tropical Grasses*. Food and Agriculture Organization of the United Nations, Rome.

Van Soest, P. J. (1994) *The Nutritional Ecology of the Ruminant*. 2nd ed., Cornell University Press, Ithaca, N.Y.

Van Soest, P. J., Robertson, J. B., and Lewis, B. A. (1991) Methods for dietary fiber, neutral detergent fiber and nonstarch polysaccharides in relation to animal nutrition. *Journal of Dairy Science* 74: 3583-3597.

Waipanya, S., Klum-em, K., and Suankoon, S. (2005) Whip grass (*Hemarthria compressa*) for fighting bulls in Nakornsrithumarat area. (1) Effect of cutting intervals and rates of nitrogen fertilizer on forage yield and nutritive value of whip grass (*Hemarthria compressa*). *Annual Research Report 2005*, Animal Nutrition Division, Department of Livestock Development, Ministry of Agriculture and Cooperative, pp. 100-116 (in Thai with English abstract).

Wilson, J. R. and Minson, D. J. (1980) Prospects for improving the digestibility and intake of tropical grasses. *Tropical Grasslands* 14: 253-259

Research Article

Computer Simulation for Studying Complexation between a Model Drug and a Model Protein

Wibul Wongpoowarak[1], Nimit Worakul[1], Wiwat Pichayakorn[1],
Payom Wongpoowarak[2], Prapaporn Boonme[1*]

[1]Department of Pharmaceutical Technology, [2]Department of Clinical Pharmacy,
Faculty of Pharmaceutical Sciences, Prince of Songkla University, Hat-Yai, Songkhla, Thailand
[*]Corresponding author. E-mail address: prapaporn.b@psu.ac.th

Received November 2, 2010; Accepted January 11, 2011

Abstract

Computer simulation is one of effective tools for instructors to illustrate complexation or binding interaction between a model drug and a model serum protein in its entire intricacy since the students can be economically exposed to a large variety of results of laboratory design within a relative short period of time. The program of computer simulation was created with Microsoft Access™. In this simulation program, theoretical parameters such as stoichiometric ratio and binding constants were assigned. After users defined initial concentrations for drug and protein, the program would calculate free drug after complexation and adding noise with zero mean and standard deviation according to the user-defined relative standard deviation. The noise added would make the dataset to be more realistic. Users could use this obtained data to further create a Scatchard plot. The fourth-year pharmaceutical care students of the Faculty of Pharmaceutical Sciences, Prince of Songkla University used this program in studying "complexation" topic. Satisfaction of the students on the instruction using this computer simulation program was determined using a five-choice questionnaire. The results indicated that this learning method was useful and satisfactory. Most responses on the satisfaction with the study via this simulation program were averagely rated above 3 from 5.

***Key Words*:** Computer simulation; Complexation; Protein binding

Introduction

Formation of complexes or coordination compounds is occurred via donor-acceptor mechanism or Lewis acid-base reaction between two or more different chemical constituents. Protein binding with drug is a good example of most commonly found pharmaceutical complexes. The binding of a drug to plasma proteins, e.g., albumin and α_1-acid glycoprotein, can influence the inactivation or retardation of the activity of the entrapped drug. Moreover, complexation may increase the drug activity due to competitive binding of two drugs or more when administered together (Martin, 1993). Therefore, it is necessary for pharmaceutical care students to clearly understand about complexation. In order to study binding constant of complexation, useful data analyzed from this tool includes Scatchard plot and stoichiometric ratio of drug-to-protein. However, due to the delicacy of analytical methods, the experiment consumes both time and resources, and the specific assay techniques are required.

Practicing in data analysis may be practically cost-prohibitive.

In order to set-up a laboratory session for data analysis of binding constant in complexation, computer simulation is an interesting tool of choice since it can reduce time and resources. Students can fully concentrate on the data analysis methodology instead of performing very lengthy experiments in order to obtain a few datasets that may not be suitable to data analysis methodology at all. Moreover, the data analysis resulted from simulation could be validated against the theoretical values used in the software. This should enable the learning process to be performed easily within the resource-constrained teaching environment. Many reports have shown that computer simulations were used as study tools successfully in various topics such as in pharmaceutical industry management (Nelson and Gagnon, 1975), pharmacokinetics (Hayton and Collins, 1991), pharmacotherapy (Chiholm et al., 1996), pharmaceutical calculation (Ramanathan et al., 1997), and pharmaceutical technology (Mezei et al., 1990; Wongpoowarak and Boonme, 2005; Wongpoowarak et al., 2008).

A computer simulation program for studying complexation between a model drug and a model protein is demonstrated here. The software was constructed in Microsoft Access™ using automated tools for generating user-interfaces and embedded the queries. There are limitations in the reliability of simulation methodology due to the unrealistic assumptions, i.e., all binding sites are equivalent thermodynamically, relative probability of vacancy in binding process would be independent from the concentrations of both proteins and drug molecules and the average binding site could be used for representing the actual distribution of binding sites. With these limitations, the simulation should be used for a simple scenario that is not affected by the aforementioned assumptions. Hence, this program was expected to be used as a studying tool for skill improvement in data analysis of the students in a very short period. The fourth-year pharmaceutical care students of the Faculty of Pharmaceutical Sciences, Prince of Songkla University were assigned to study the complexation between a model drug and a model protein via the simulation program from "581-401 Pharmaceutical Preparation III" course. The satisfaction of the students with the program was also determined using five-choice questionnaires.

The aim of this study was to evaluate the satisfaction of the generated computer simulation program on the students' learning of the subject "complexation".

Materials and Methods

Theoretical aspects

Theoretical aspects of complexation are presented as follows (Martin, 1993). The interaction between a drug molecule *D* and a free receptor *P* in a protein can be written as Eq. 1.

$$P + D \rightleftharpoons PD \quad \text{(Eq. 1)}$$

The binding or association constant, assuming the equivalent between activities and concentrations, can be calculated by Eq. 2.

$$K = \frac{[PD]}{[P][D_f]} \quad \text{(Eq. 2)}$$

where K is the binding constant, $[P]$ is the concentration of the protein in the term of free binding sites, $[D_f]$ is the free drug, and $[PD]$ is the concentration of the drug-protein complex.

If the total protein concentration is appointed as $[P_t]$, $[P]$ can be replaced with Eq. 3 and Eq. 2 can be rewritten as Eq. 4.

$$[P] = [P_t] - [PD] \quad \text{(Eq. 3)}$$

$$\frac{[PD]}{[P_t]} = \frac{K[D_f]}{1 + K[D_f]} \quad \text{(Eq. 4)}$$

The stoichiometric ratio, *r*, or number of moles of drug bound $[PD]$ per mole of total protein $[P_t]$ can be found with Eq. 5.

$$r = \frac{[PD]}{[P_t]} = \frac{K[D_f]}{1 + K[D_f]} \quad \text{(Eq. 5)}$$

If independent binding sites, υ, referred to as *nSite* here, are accessible, *r* can be performed as Eq. 6 and then rearranged as exhibited in Eq. 7 providing a graph called a Scatchard plot.

$$r = \upsilon \frac{K[D_f]}{1 + K[D_f]} \quad \text{(Eq. 6)}$$

$$\frac{r}{[D_f]} = \upsilon K - rK \quad \text{(Eq. 7)}$$

Simulation methodology for average binding site calculation

In order to simulate the behavior of Eq.1 – Eq.7, we had to start from guessing the complex formation first and compute all related values from the complex formation. If all related values could provide exact value of binding constants, it is equivalent to using binding constants to calculate all species involved.

Suppose that a molecule of *P* could bind to a maximum of *n* molecules of *D*. Under an assumption that all binding sites are all equivalent thermodynamically, occupation of *P* by *D* would distribute with equal probability for all binding sites.

Considering a molecule of protein *P*, the relative probability of vacancy is supposed to be 1. With *n* sites of binding, the relative chance that 1 molecule of *D* would bind to this site will be 1/*n*. The relative chance that 2 molecules of *D* bind to *P* simultaneously would be 1/*n* multiplied with 1/*n*, i.e., $1/n^2$. The chance that 3 molecules of *D* bind to *P* simultaneously would be $1/n^3$, and so on. For full occupation, the occupancy chance will be the sum for all binding situation. The relative chance that a molecule of *P* will be in bound state is thus equal to $1/n + 1/n^2 + 1/n^3 + \ldots + 1/n^n$, while the relative chance of staying in the vacancy state equal to 1. The average occupied binding site should thus be equal to such relative chance multiplied with total binding site, *n*. This provides an average binding site (*nSite*) formula as presented in Eq. 8.

$$nSite = \sum_{i=1}^{n} \frac{1}{n^{i-1}} \quad \text{(Eq. 8)}$$

where *n* is the maximum binding site.

If one molecule of drug could bind to several proteins, the value of average binding site will be reversed from the Eq. 8.

Simulation methodology for equilibrium concentration

Due to unbalance concentration of drug and protein, the maximum amount of complex formation will be determined from the minimum value between {drug amount/number of drug molecules involved in the complex formation} and {protein amount/number of protein molecules involved in the complex formation}. Moreover, the actual complex will be formed only partially in relative to this maximum amount of complex formation. We could express the actual complex formed by using formation efficiency (*f*), which always between 0 and 1, multiply with this maximum amount of complex formation. For each value of *f*, all chemical species involved could be calculated explicitly, i.e., by assigning *QD* for stoichiometric ratio for drug and *QP* for stoichiometric ratio for protein as illustrated in Eq. 9 – Eq. 12.

$$[PD] = (f)(Maximum\ complex\ formation) \quad \text{(Eq. 9)}$$

$$[D_f] = [D] - (QD)[PD] \quad \text{(Eq. 10)}$$

$$[P] = [P_t] - (QP)[PD]/nSite \quad \text{(Eq. 11)}$$

$$K = \frac{[PD]}{[D_f]^{QD}[P]^{QP}} \quad \text{(Eq. 12)}$$

One could clearly see that *K* is monotonically depends on *f*. At very low value of *f*, *K* would be

very small. At very high value of *f*, *K* would be very high. If we use trial-and-error method to find the condition that the computed *K* (as a function of *f*) is exactly equal to the assigned value of *K*, we will know all species involved with such *K*. By using binary search for *f* that could provide the desired *K*, we could limit the trial-and-error test to within 60 iterations in order to obtain 16-significant digits of accuracy ($2^{60} \approx 1.2*10^{18}$, this mean that by using 60 trials, we would obtain accuracy to within a fraction of 10^{-18}) and all species could be computed directly.

In brief, by varying *f* between 0 and 1 (using binary-search), all species related to that *f* could be known and *K* could be computed. The best value of *f* providing the desired *K* could give all related chemical species of interest. This algorithm could be implemented as a user-defined function in Visual Basic Code with maximum of 60 iterations to provide the result that is reliable to 16 significant digits.

Description of computer simulation program

The program of computer simulation for studying complexation between a model drug and a model protein was created with Microsoft Access™. The source code for implementing in the software was exhibited in the Appendix. The Microsoft Access™ was preferable since it has been easier to make the computation process invisible to the user's attention, and the user can control recalculation to be made manually only when needed. In addition, it was proved to be suitable program for simulation tools for teaching (Wongpoowarak and Boonme, 2005; Wongpoowarak et al., 2008). In this study, the 2003 version of Microsoft Access™ was used, and the program could be upgraded to other versions since it was created with minimal user-interfaces and involved a few queries.

The simulation intends to mimic equilibrium dialysis data. According to the equilibrium dialysis method, a model protein was placed in a number of dialyzing-membrane vesicles. The vesicles were tied firmly and suspended in beakers containing the drug in various concentrations. Finally, the concentrations of the drug in free form were determined by analytical assay. This program simulates instead of actual chemical assay, by allowing users to define their own experimental conditions at will.

By assigning theoretical parameters such as stoichiometric ratio (*QD/QP*) and binding constants (*K*), users could define initial concentrations for drug ([*D*]) and protein ([*P*]) for the software to generate simulated data according to the protocol according to the theoretical concept. The generated noise, using log-and-trig formula (Daykin et al., 1994), normally distributing around zero mean with an assigned relative standard deviation (RSD) was added to the generated data. This will make the dataset to be more realistic. The students could use these simulated data to practice data analysis process. They could also learn that different research designs would also provide different quality of the obtained parameters from a Scatchard plot.

Laboratory methodology

The 39 fourth-year pharmacy students of the Faculty of Pharmaceutical Sciences, Prince of Songkla University attended the lecture of 'complexation' topic in the course of "581-401 Pharmaceutical Preparation III" in Semester 1/2010. In this lecture, the students had been introduced about the concept of complexation. Classification of complexes, applications of complexation in pharmaceutics, method of analysis, protein binding, and thermodynamic treatment of stability constants had been informed. Subsequently, the students were assigned to plan an experimental design for studying protein binding between a model drug and a model protein. The instructors suggested the students to produce the concentration profile of a drug and protein with randomized

and fixed maximum values in a spreadsheet of Microsoft Excel™. For example, if the students would like to use drug concentrations from 0 to 10 moles, they could create the concentrations with "=10*RAND()". Afterwards, the data were copied and then pasted on the spreadsheet of Microsoft Access™ to generate concentrations of free drug according to the assigned theoretical parameters. The students could design their own experiment with the computer simulation program posted on http://mail.pharmacy.psu.ac.th/~wwibul/complex.mdb. Moreover, they could study and review the topic of complexation by themselves at any times.

Evaluation of satisfaction of the students

All students received a questionnaire asking about the degree of their satisfaction with studying by the computer simulation program. The questionnaire was five-choice of satisfaction level with the statement, i.e., excellence (score = 5), good (score = 4), fair (score = 3), poor (score = 2), and very poor (score = 1). The students were asked to return the completed questionnaire after their first using of the program in the computer center of the faculty.

Results and Discussion

Figure 1 shows an example of spreadsheet in Microsoft Access™ simulation program. The designed protocol, i.e., [*P*] and [*D*], was filled in columns of "Protein" and "Initial Concentration", respectively. The generated concentrations of free drug ([D_f]) after complexation according to the assigned theoretical parameters appeared in column of "Final Concentration". Afterwards, the students could use the obtained data for further calculation in computer spreadsheet software, e.g., Microsoft Excel™. The values of [D_f]/[*D*], [*PD*], *r* and $r/[D_f]$ could be computed from the simulated dataset, where [*PD*] and *r* are defined in Eq. 3 and Eq. 5, respectively. The data after equilibrium or saturation point obtained from the plot between [*P*] and [D_f]/[*D*] were selected for further creating a Scatchard plot. The equilibrium or saturation point could be observed when [D_f]/[*D*] was near zero. A Scatchard plot could be created using *r* as x-axis and $r/[D_f]$ as y-axis. The linear equation was then estimated by simple linear regression. For the interpretation according to Eq. 7, *K* was equal to minus slope and *nSite* was equal to y-intercept over *K*. If the students used appropriate concentrations of drug and protein, they could obtain the experimental parameters similar to the assigned theoretical parameters. If not, the generated data would not provide a Scatchard plot, due to the reason that any extreme values of drug or protein used would be very sensitive to noise effect in the Scatchard transformation. However, the students could repeat all over again with different experimental conditions.

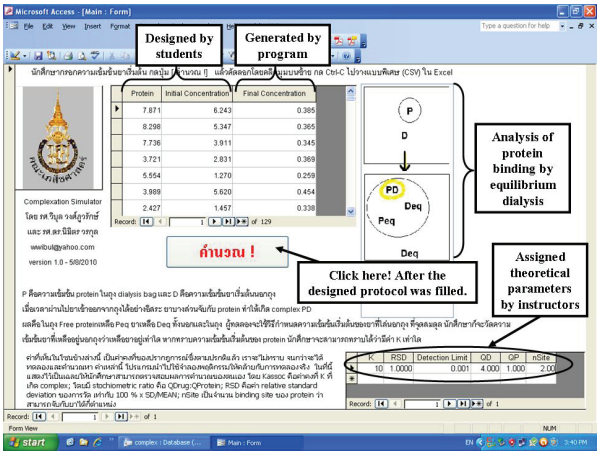


Figure 1 A spreadsheet of Microsoft Access™ computer simulation.

The total of 38 questionnaires (97.44%) were answered and returned from 39 students of one class. The numbers of each answer for each question were presented in Table 1. The results exhibited that the method of learning was satisfactory. The students averagely rated above 3 from 5 for their satisfaction on this simulation tool. The majority of the students thought that the learning method was useful, helped them to study by themselves, and reduced the time of studying. Computer simulation experiments offered possibility to practice data analysis in complexation topic within a very short period of time.

Table 1 The numbers of the answers and average scores on the questionnaires about the opinion of 38 students on the studying with the computer simulation program of complexation.

No.	Question	Satisfaction level*					Average Scores
		5	4	3	2	1	
1	The computer simulation was remarkable.	13	18	4	1	2	4.03
2	The studying with computer simulation supported the student to understand the subject in a relative short period of time.	8	12	11	6	1	3.53
3	The computer simulation could be applied to use in studying of other topics.	9	17	10	2	0	3.87
4	Using computer simulation in this study provided the student to comprehend the subject of "complexation".	8	11	14	5	0	3.58
5	Listening to the lecture in this study provided the student to understand the subject of "complexation".	5	18	12	3	0	3.66
6	Searching information by your own in this study provided the student to understand the subject of "complexation".	1	11	19	6	1	3.13
7	The software program of computer simulation using in this study was effortless to apply.	5	16	14	1	2	3.55
8	Using computer simulation helped the student to understand the subject of "complexation" more than only listening to the lecture.	9	15	11	1	2	3.74

Note: *Score: 5 = excellence, 4 = good, 3 = fair, 2 = poor, 1 = very poor.

Conclusions

Studying requires a proactive approach by both the instructors and the students. Developing new tools such as computer simulations allows the instructors to offer the students new avenues to study. The students not only develop their knowledge but also develop their ability to study by themselves. The results suggested that the proposed simulation program for complexation was a satisfactory studying tool.

References

Chiholm M.A., Dehoney J., and Poirier S. (1996) Development and evaluation of a computer assisted instructional program in an advanced pharmcotherapeutics course. *American Journal of Pharmaceutical Education* 60: 365-369.

Daykin C.D., Pentikainen T., and Pesonen M. (1994) *Practical Risk Theory for Actuaries*, pp. 469. London: Chapman & Hall.

Hayton W.L. and Collins P.L. (1991) STELLA: Simulation software for pharmacokinetic software. *American Journal of Pharmaceutical Education* 55: 131-134.

Martin A. (1993) *Physical Pharmacy,* 4th ed., pp. 251-283. Philadelphia: Lea & Febiger.

Mezei J., Kuttel S., and Rácz I. (1990) Computerassisted instruction: how to solve drug formulation problems. *American Journal of Pharmaceutical Education* 54: 30-34.

Nelson A.A. and Gagnon J.P. (1975) PHARMASIM: a computer simulation game of the pharmaceutical industry. *Med Mark Media* 10: 54-60.

Ramanathan M., Chau R.I., and Straubinger R.M. (1997) Integration of Internet-based technologies as a learning tool in a pharmaceutical calculations course. *American Journal of Pharmaceutical Education* 61: 141-148.

Wongpoowarak W, Wongpoowarak P, and Boonme P. (2008) Simulation tool for teaching multivariate experimental design on paracetamol syrup formulation. *Silpakorn University Science and Technology Journal* 2(1): 45-52.

Wongpoowarak W. and Boonme P. (2005) Computer simulation for studying effects of laboratory design on results of accelerated stability test. *Silpakorn University International Journal* 5(1-2): 108-117.

Appendix

The source code for implemented in the software was a user-defined function that could be used to calculate the equilibrium drug concentration from assigned values of binding constants and the input values of total protein concentration and initial drug concentration.

UPDATE Drug, [Default] SET Drug.Dfree =ConcEquilibrium(Drug!P,Drug!D,Default!K,Default!RSD,Default!QD,Default!QP,Default!nSite) WHERE (((Drug.P) Is Not Null) AND ((Drug.D) Is Not Null));

The default value for K, QD, QP, nSite and RSD of noise to be simulated are given in a table [Default] and initial value of P and D are given in a table [Drug].

By execute such SQL code, the new value of equilibrium drug will be updated in a column *Dfree* in table [Drug]. This is the simulated dataset that can be further processed in spreadsheet software.

```
'Source code start here
Function NormalDistribution(xbar, sd) As Double
    'This is log-and-trig formula (Daykin et al., 1994)
    Do
    r1 = Rnd(Timer)
    r2 = Rnd(Timer)
    Loop Until r1 > 0
    x1 = (Cos(2 * Pi * r2)) * Sqr(-2 * Log(r1))
```

```
    If Rnd(Timer) > 0.5 Then
      NormalDistribution = xbar + x1 * sd
    Else
      NormalDistribution = xbar - x1 * sd
    End If
End Function
Function AverageBindingSite(maxsite) As Double
    If maxsite > 1 Then
      N = CInt(maxsite)
    ElseIf maxsite = 1 Then
      N = 1
    Else
      N = CInt(1 / maxsite)
    End If
    sump = 0
    For i = 1 To N
      sump = sump + 1 / N ^ i
    Next i
    If maxsite >= 1 Then
      AverageBindingSite = 1 / sump
    Else
      AverageBindingSite = sump
    End If
End Function
Function ConcEquilibrium(P, D, K, RSD, QD,
QP, nSite)
On Error GoTo xx
    If P / QP < D / QD Then
       UpperboundComplex = P / QP
    Else
       UpperboundComplex = D / QD
       Minf = 0
    Maxf = 1
    For i = 1 To 60
      f = (Minf + Maxf) / 2
      cpx = f * UpperboundComplex
      myK = cpx / (((D - cpx * QD) ^ QD) * ((P
- cpx * QP / AverageBindingSite(nSite)) ^ QP))
      If myK > K Then Maxf = f Else If myK <
K Then Minf = f Else Exit For
    Next i
    De = D - cpx * QD
    'Now add error term to the result
    ConcEquilibrium = NormalDistribution(De,
RSD * De / 100)
    Exit Function
xx:
Resume Next
End Function
```

Research Article

Structural and Morphological Characterization of Chemical Bath Deposition of FeS Thin Films in the Presence of Sodium Tartrate as a Complexing Agent

Anuar Kassim[1*], Ho Soon Min[1], Loh Yean Yee[1] and Saravanan Nagalingam[2]

[1]*Department of Chemistry, Faculty of Science, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia*
[2]*Faculty of Science, Universiti Tunku Abdul Rahman, 31900 Kampar, Perak, Malaysia*
[*]*Corresponding author. E-mail address: soonminho@gmail.com*

Received July 20, 2010; Accepted February 1, 2011

Abstract

In this paper, we presented the results of X-ray diffraction and scanning electron microscopy of the iron sulphide thin films prepared using a simple and cost effective chemical bath deposition method. The effects of solution concentration and pH on the structural and morphological properties of thin films were studied in the presence of sodium tartrate as a complexing agent. The thin films deposited using higher solution concentration showed higher number of FeS peaks and larger grain size according to X-ray diffraction and scanning electron microscopy results, respectively as compared with other solution concentrations. On the other hand, when the thin films were deposited at higher pH, the number of FeS peaks reduced to two peaks and the films showed incomplete coverage of material over the surface of the substrate with the smaller grain size.

Key Words: Chemical bath deposition; Iron sulphide; Scanning electron microscopy; Thin films

Introduction

Iron sulphide thin films are very attractive materials for a wide variety of technological applications such as photoelectrochemical and photovoltaic applications. Various methods are used for the preparation of iron sulphide thin films such as chemical vapor transport (Willeke et al., 1992), metal-organic chemical vapour deposition (Thomas et al., 1997), sputtering (Birkholz et al., 1992), molecular beam deposition (Bronold et al., 1997), flash evaporation (Ferrer et al., 1990), electrodeposition (Nakamura and Yamamoto, 2001) and chemical bath deposition (Anuar et al., 2010). Among various other methods, the chemical bath deposition method is found to be a cheap and simple way to deposit large area polycrystalline metal chalcogenide thin films. The preparations of various thin films using chemical bath deposition technique such as CdS (Moualkia et al., 2009), As_2S_3 (Mane et al., 2004), MnS (Gumus et al., 2007), PbS (Larramendi et al., 2001), ZnS (Ubale et al., 2007), $Cd_{0.5}Zn_{0.5}Se$ (Kale et al., 2007) and Cu_4SnS_4 (Anuar et al., 2009) have reported by several authors.

However, there is no attempt made on the chemical bath deposition of the iron sulphide thin films, using sodium tartrate as a complexing agent.

In view of this, the synthesis of iron sulphide thin films was performed at different pH values and solution concentrations. The crystal structure and surface morphology of iron sulphide thin films were then investigated.

Materials and methods

All the chemicals used for the deposition were analytical grade reagents and all the solutions were prepared in deionised water (Alpha-Q Millipore). The iron sulphide thin films were prepared from an acidic bath containing aqueous solutions of iron nitrate, sodium thiosulfate and sodium tartrate. The microscope glass slide was used as the substrate for the chemical bath deposition of iron sulphide thin films. Before deposition, the microscope glass slide was degreased with ethanol for 15 min, then, ultrasonically cleaned with distilled water for another 15 min and dried in desiccators. Deposition of iron sulphide thin films was carried out using following procedure. 20 mL of iron nitrate was complexed with 20 mL of 0.2 M sodium tartrate. Then, 20 mL of sodium thiosulfate was added slowly to the mixture. The cleaned glass slide was immersed vertically into the solution. The deposition process was carried out by varying solution concentrations (0.1, 0.15 and 0.2 M) and pH values (2 and 2.5). During deposition process, the beaker was kept undisturbed. After the completion of deposition (2 h), the glass slide was removed, washed several times with distilled water and dried in desiccators for further characterization.

In order to investigate the crystallographic properties of the iron sulphide thin films, the X-ray diffraction analyses were carried out using Philips PM 11730 diffractometer with CuK_{α} (λ=1.5418 Å) radiation for the 2θ ranging from 20 to 65°. The surface morphology was observed by a scanning electron microscopy (JEOL, JSM-6400). All the samples were taken at 20 kV with a 1000 X magnification.

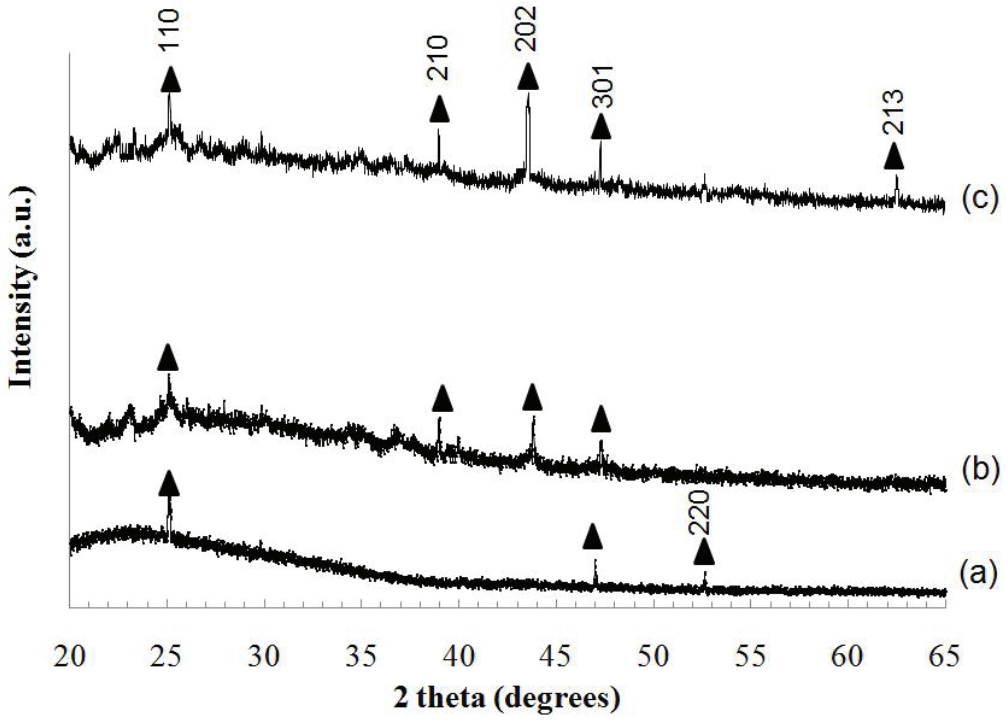


Figure 1 X-ray diffraction patterns for iron sulphide thin films deposited at various solution concentrations at pH 2. (a) 0.1 M (b) 0.15 M (c) 0.2 M

Table 1 Comparison of the JCPDS *d*-spacing data for iron sulphide thin films to experimentally observed values for the sample deposited at various solution concentrations at pH 2

Solution concentration (M)	2θ (°)	*hkl*	*d*-spacing (Å)	
			Observed value	JCPDS value
0.1	25.1	110	3.5	3.5
	46.9	301	2.0	1.9
	52.6	220	1.8	1.7
0.15	25.2	110	3.5	3.5
	38.9	210	2.3	2.3
	43.7	202	2.1	2.1
	47.0	301	1.9	1.9
0.2	25.2	110	3.5	3.5
	38.6	210	2.4	2.3
	43.7	202	2.1	2.1
	47.3	301	1.9	1.9
	62.5	213	1.5	1.5

Results and discussion

Figure 1 and Table 1 show the X-ray diffraction (XRD) patterns and data for the thin films deposited at various solution concentrations at pH 2, respectively. For the thin films prepared using 0.1 M iron nitrate and sodium thiosulfate, three peaks at 2θ= 25.1°, 46.9° and 52.6° are observed, which referred to the (110), (301) and (220) planes of FeS, respectively. However, the number of peaks increased to four (Figure 1b) and finally five peaks (Figure 1c) when the concentration is increased to 0.15 and 0.2 M, respectively. The position of several peaks is used to determine the iron sulphide as shown in Table 1. These peaks are well matched with the Joint Committee on Powder Diffraction Standard (JCPDS) data for FeS (JCPDS reference code: 01-080-1028) (Keller-Besrest and Collin, 1990). The lattice parameter values are a=b=6.958 Å, c=5.824 Å, α=β=90°, γ=120°.

Figure 2 shows the scanning electron microscopy (SEM) micrographs of the FeS thin films prepared using different solution concentrations at pH 2. Based on the Figure 2a, the films prepared using 0.1 M $Fe(NO_3)_3$ and $Na_2S_2O_3$ show incomplete coverage of material over the surface of the substrate. This may be caused by insufficient amount of iron and sulfide ions in the mixture. The thin films deposition process on a substrate depends mainly on the formation of nucleation sites and subsequent growth of the thin films from this centre. However, further increment in the solution concentration to 0.15 M $Fe(NO_3)_3$ and $Na_2S_2O_3$ indicates almost complete coverage of the FeS material over the substrate compared to the films prepared at lower concentration. At higher concentration (0.2 M), the material is found to cover the surface of the substrate completely. Formation of granules, which is uniformly distributed over the deposit layer, can

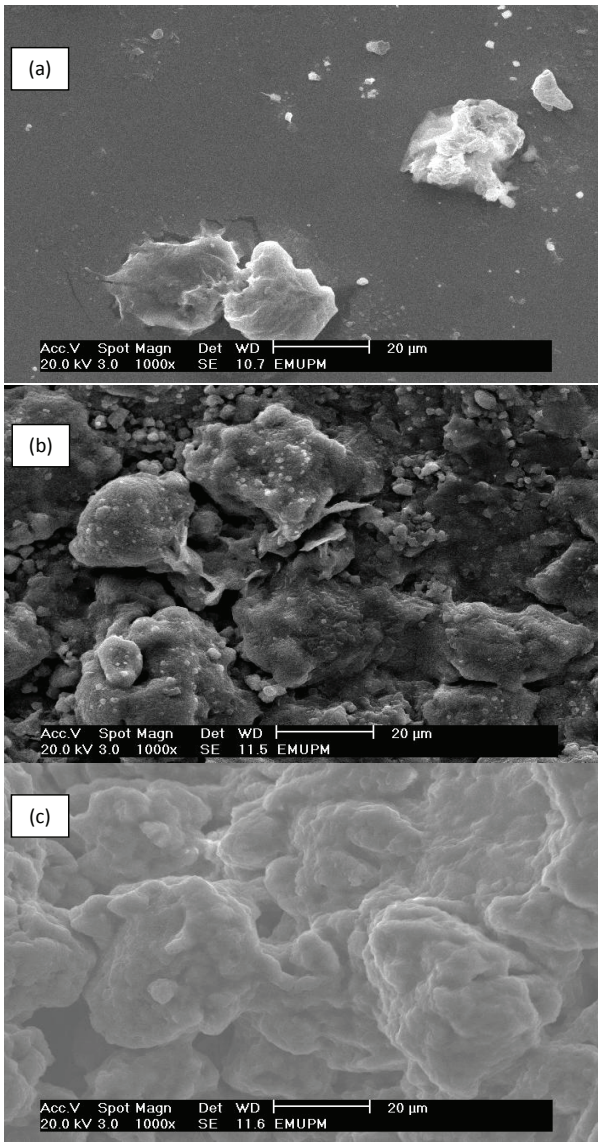


Figure 2 Scanning electron microscopy micrographs of the FeS thin films prepared using different solution concentrations at pH 2. (a) 0.1 M (b) 0.15 M (c) 0.2 M

be seen in Figure 2c. The grain sizes (4-6 µm) were almost similar to each other. Based on the SEM micrograph, the grains structures are formed in an agglomerated morphology (average size is about 20 µm).

Figure 3 and Table 2 show the X-ray diffraction (XRD) patterns and XRD data for the iron sulphide thin films prepared at various pH values using 0.2 M iron nitrate and sodium thiosulfate, respectively. Comparison between the thin films deposited at pH 2 and 2.5 reveals that the number of FeS peaks increased, indicating better crystalline phase for the films prepared at lower pH. The films deposited at pH 2 show five peaks and the d-spacing values obtained match with the standard JCPDS data (Table 2). The positions of the peaks obtained indicate that hexagonal FeS structure with (110), (210), (202), (301) and (213) planes have been deposited. On the other hand, we observed that the intensity of the peaks were much better for the films deposited at pH 2. At lower pH value, the peak intensities were increasing which showed the improvement in the crystallinity of the films. As the pH was decreased from pH 2.5 to 2, the intensity of the peaks corresponding to (110) and (202) planes increased. These planes seem dominant at this stage of experiment.

The scanning electron microscopy (SEM) micrographs of the iron sulphide thin films prepared at different pH solutions using 0.2 M $Fe(NO_3)_3$ and $Na_2S_2O_3$ are shown in Figure 4.

The SEM micrograph of the thin films deposited at pH 2 shows distribution of grains, which covers the surface of the substrate completely (Figure 4a). However, as the pH is increased to 2.5, the distribution of grains has been reduced and resulted in a lower surface coverage. These films have smaller grains compared to the other films (Figure 4b). The pinholes can be observed on the surface of these films. The pinholes are areas which were not covered by thin films.

Conclusions

FeS thin films have been successfully deposited by chemical bath deposition method. XRD study revealed polycrystalline nature of the films with hexagonal phase. Based on the XRD data, the films prepared at lower pH and higher solution concentration indicated higher number of FeS peaks. The surface morphology of these films was observed quite uniform and well covered on the substrate than

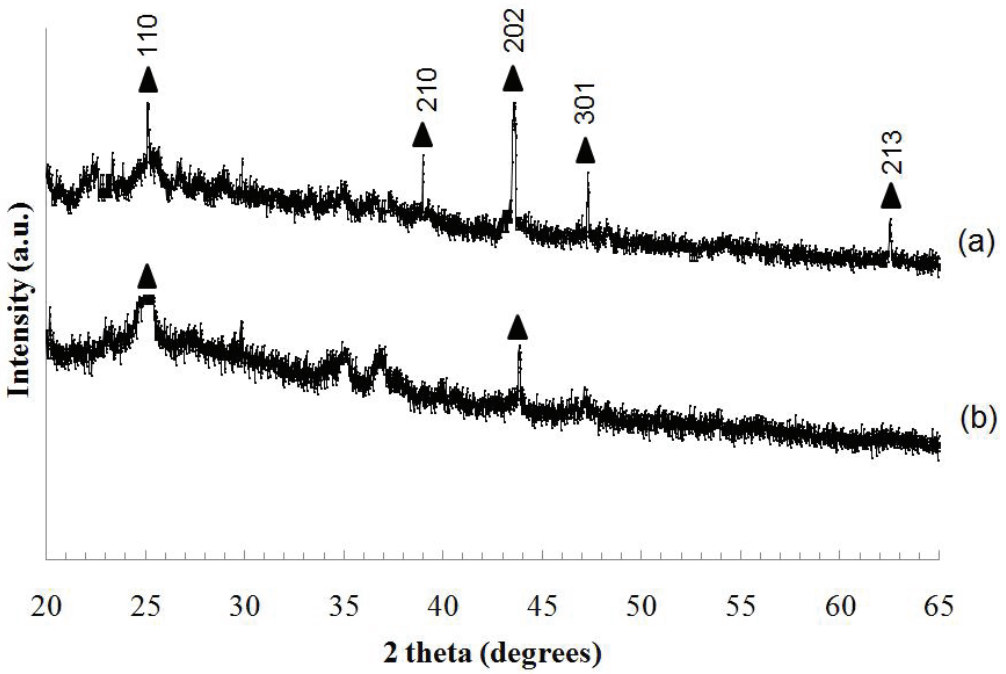


Figure 3 X-ray diffraction patterns for iron sulphide thin films deposited at various pH values using 0.2 M iron nitrate and sodium thiosulfate (a) pH 2 (b) pH 2.5

Table 2 Comparison of the JCPDS *d*-spacing data for iron sulphide thin films to experimentally observed values for the sample deposited at various pH values using 0.2 M iron nitrate and sodium thiosulfate

pH	2θ (°)	*hkl*	*d*-spacing (Å)	
			Observed value	JCPDS value
2	25.2	110	3.5	3.5
	38.6	210	2.4	2.3
	43.7	202	2.1	2.1
	47.3	301	1.9	1.9
	62.5	213	1.5	1.5
2.5	25.2	110	3.5	3.5
	43.7	202	2.1	2.1

other samples. Experimental results indicated that the deposition at pH 2 using 0.2 M iron nitrate and sodium thiosulphate was the optimum condition for the preparation of FeS films.

Acknowledgements

The authors would like to thank the Department of Chemistry, University Putra Malaysia for the provision of laboratory facilities and MOSTI for the National Science Fellowship.

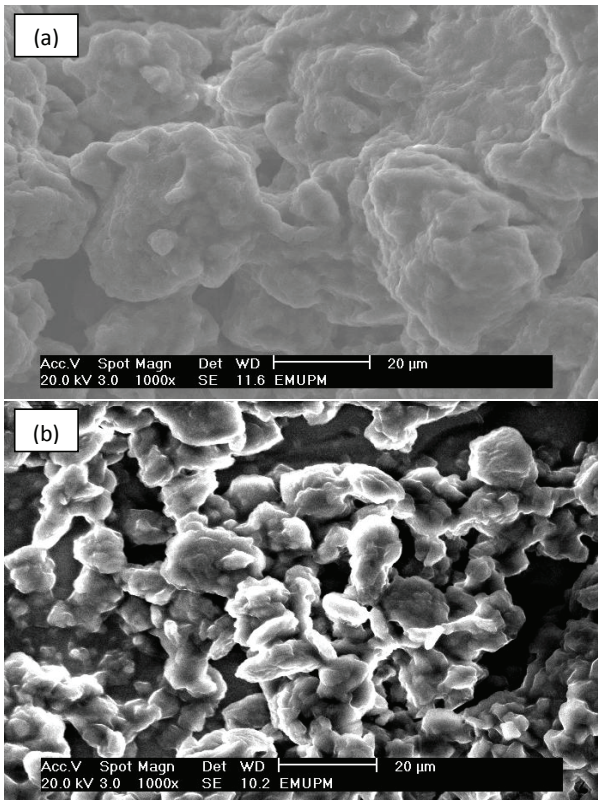


Figure 4 Scanning electron microscopy micrographs of the iron sulphide thin films prepared at different pH solutions using 0.2 M $Fe(NO_3)_3$ and $Na_2S_2O_3$. (a) pH 2 (b) pH 2.5

References

Anuar, K., Saravanan, N., Tan, W.T., Atan, M.S., Dzulkefly, K.A., Elas, M.J. and Ho, S.M. (2009) Effect of deposition period and pH on chemical bath deposited Cu_4SnS_4 thin films. *Philippine Journal Science* 138: 161-168.

Anuar, K., Tan, W.T., Dzulkefly, K.A., Atan, M.S., Ho, S.M., Gwee, S.Y. and Saravanan, N. (2010) Preparation and characterization of FeS_2 thin films by chemical bath deposition method. *Indonesian Journal of Chemistry* 10: 8-11.

Birkholz, M., Lichtenberger, D., Hopfner, C. and Fiechter, S. (1992) Sputtering of thin pyrite films. *Solar Energy Materials and Solar Cells* 27: 243-251.

Bronold, M., Kubala, S., Pettenkofer, C. and Jaegermann, W. (1997) Thin pyrite (FeS_2) films by molecular beam deposition. *Thin Solid Films* 304: 178-182.

Ferrer, I.J., Nevskaia, D.M., Heras, C. and Sanchez, C. (1990) About the band gap nature of FeS as determined from optical and photoelectrochemical measurements. *Solid State Communications* 74: 913-916.

Gumus, C., Ulutas, C. and Ufuktepe, Y. (2007) Optical and structural properties of manganese sulfide thin films. *Optical Materials* 29: 1183-1187.

Kale, R.B., Lokhande, C.D., Mane, R.S. and Han, S.H. (2007) $Cd_{0.5}Zn_{0.5}Se$ wide range composite thin films for solar cell buffer layer application. *Applied Surface Science* 253: 3109-3112.

Keller-Besrest, F. and Collin, G. (1990) Structural aspects of the α transition in stoichiometric FeS: Identification of the high temperature phase. *Journal of Solid State Chemistry.* 84: 194-210.

Larramendi, E.M., Calzadilla, O., Arias, A.G., Hernandez, E. and Garcia, J.R. (2001) Effect of surface structure on photosensitivity in chemically deposited PbS thin films. *Thin Solid Films* 389: 301-306.

Mane, R.S., Todkar, V.V. and Lokhande, C.D. (2004) Low temperature synthesis of nanocrystalline As_2S_3 thin films using novel chemical bath deposition route. *Applied Surface Science* 227: 48-55.

Moualkia, H., Hariech, S. and Aida, M.S. (2009) Properties of CdS thin films grown by chemical bath deposition as a function of bath temperature. *Materials Science Forum* 609: 243-247.

Nakamura, S. and Yamamoto, A. (2001) Electrodeposition of pyrite (FeS_2) thin films for photovoltaic cells. *Solar Energy Materials and Solar Cells* 65: 79-85.

Thomas, B., Ellmer, K., Muller, M., Hopfner, C., Fiechter, S. and Tributsch, H. (1997) Structural and photoelectrical properties of FeS_2 (Pyrite) thin films grown by MOCVD. *Journal of Crystal Growth* 170: 808-812.

Ubale, A.U., Sangawar, V.S. and Kulkarni, D.K. (2007) Size dependent optical characteristics of chemically deposited nanostructured ZnS thin films. *Bulletin of Materials Science* 30: 147-151.

Willeke, G., Blenk, O., Kloc, C.H. and Bucher, E. (1992) Preparation and electrical transport properties of pyrite (FeS_2) single crystals. *Journal of Alloys and Compounds* 178: 181-191.

Acknowledgement to Referees 2009 – 2010

Between 2009 and 2010 (4 issues), 32 manuscripts were submitted to SUSTJ. Nineteen manuscripts were accepted for publication, 9 were rejected, and 4 manuscripts the authors have withdrawn their submission.

The members of the Editorial Advisory Board and the Editorial Board wish to express their grateful appreciation to the reviewers named below for their competent and critical evaluation of submitted manuscripts during 2009 and 2010.

Suchitra Adulkasem
Department of Computing, Faculty of Science, Silpakorn University, Thailand.
Chokechai Aekatasanawan
National Corn and Sorghum Research Center, Kasetsart University, Thailand.
Thanaporn Amnuaikit
Southern Chemical & Pharmaceutical Science Research Network, Prince of Songkla University, Thailand.
Titinun Auamnoy
Faculty of Pharmaceutical Sciences, Chulalongkorn University, Thailand.
Sujin Bureerat
Department of Mechanical Engineering, Faculty of Engineering, Khon Kaen University, Thailand.
Chantana Chantrapornchai
Department of Computing, Faculty of Science, Silpakorn University, Thailand.
Chatchai Chinpaisal
Department of Pharmacology and Toxicology, Faculty of Pharmacy, Silpakorn University, Thailand.
Sethavidh Gertphol
Department of Computer Science, Faculty of Science, Kasetsart University, Thailand.
Daranee Hormdee
Department of Computer Engineering, Faculty of Engineering, Khon Kaen University, Thailand.
Junya Intaranongpai
Faculty of Pharmaceutical Sciences, Ubon Rajathanee University, Thailand.
Verayuth Lertnattee
Department of Health-Related Informatics, Faculty of Pharmacy, Silpakorn University, Thailand.
Vorrada Loryuenyong
Department of Materials Science and Engineering, Silpakorn University, Thailand.
Sathit Niratisai
Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Silpakorn University, Thailand.
Boonsri Ongpipattanakul
Faculty of Pharmaceutical Sciences, Chulalongkorn University, Thailand.
Vimolvan Pimpan
Department of Materials Science, Faculty of Science, Chulalongkorn University, Thailand.
Veeranan Pongsapukdee
Department of Statistics, Faculty of Science, Silpakorn University, Thailand.
Nalinee Poolsap
Department of Pharmacy, Faculty of Pharmacy Silpakorn University, Thailand.

Sommai Priprem
Department of Mechanical Engineering, Khon Kaen University, Thailand.
Paiboon Ratanaprasert
Department of Statistics, Faculty of Science, Silpakorn University, Thailand.
Rungpetch Sakulbumrungsil
Pharmacy Administration Unit, Faculty of Pharmaceutical Sciences, Chulalongkorn University, Thailand.
Rachada Settavongsin
Agriculture Technology, Songkhla Rajabhat University, Thailand.
Denpong Soodphakdee
Department of Mechanical Engineering, Khon Kaen University, Thailand.
Thongchai Sooksawate
Department of Physiology, Faculty of Pharmaceutical Sciences, Chulalongkorn University, Thailand.
Uthai Sotanaphun
Department of Pharmacognosy, Faculty of Pharmacy, Silpakorn University, Thailand.
Lawan Sratthaphut
Department of Health-Related Informatics, Faculty of Pharmacy, Silpakorn University, Thailand.
Dilok Sriprapai
Tool & Materials Engineering, Faculty of Engineering, King Mongkut's University of Technology Thonburi, Thailand.
Nunthaluxna Sthapornnanon
Department of Pharmacy, Faculty of Pharmacy, Silpakorn University, Thailand.
Supachai Supalaknari
Department of Chemistry, Faculty of Science, Silpakorn University, Thailand.
Wisit Tangkeangsirisin
Department of Biopharmacy, Faculty of Pharmacy, Silpakorn University, Thailand.
Linna Tongyonk
Faculty of Pharmaceutical Sciences, Chulalongkorn University, Thailand.
Sayan Tudsri
Department of Agronomy, Faculty of Agriculture, Kasetsart University, Thailand.
Mont Kumpugdee Vollrath
Faculty of Pharmaceutical and Chemical Engineering, University of Applied Sciences Berlin, Germany.
Chalong Wachirapakorn
Department of Animal Science, Faculty of Agriculture, Khon Kaen University, Thailand.

Ogenstad Declaration Exhibit E

molecular pharmaceutics

Article

pubs.acs.org/molecularpharmaceutics

In Silico Prediction of Pharmaceutical Degradation Pathways: A Benchmarking Study

Mark H. Kleinman,*[,†] Steven W. Baertschi,[‡] Karen M. Alsante,[§] Darren L. Reid,[∥] Mark D. Mowery,[⊥] Roman Shimanovich,[#] Chris Foti,[§,●] William K. Smith,[‡] Dan W. Reynolds,[▽] Marcela Nefliu,[⊥] and Martin A. Ott[○]

[†]Projects, Clinical Platforms and Sciences, GlaxoSmithKline, 2301 Renaissance Boulevard, RN0320, King of Prussia, Pennsylvania 19406, United States

[‡]Eli Lilly and Company, Lilly Research Laboratories, Small Molecule Design and Development, Indianapolis, Indiana 46285, United States

[§]Pfizer Worldwide Research and Development Pharmatherapeutics Pharmaceutical Sciences, Groton, Connecticut 06340, United States

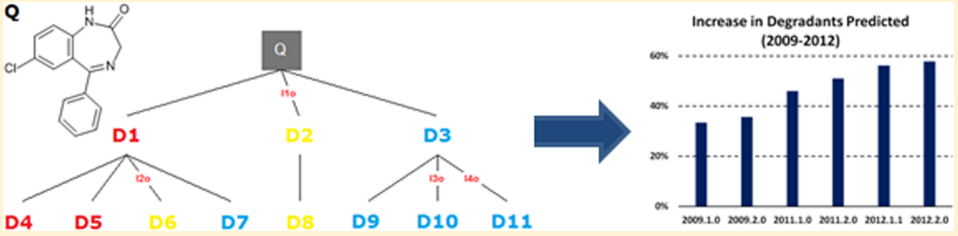
[∥]Oral Delivery, Product and Process Development, Drug Product Technologies, Amgen, Inc., One Amgen Center Drive, Thousand Oaks, California 91320, United States

[⊥]Merck and Co., Inc., Pharmaceutical Sciences and Clinical Supplies, West Point, Pennsylvania 19486, United States

[#]Oral Delivery, Product and Process Development, Drug Product Technologies, Amgen, Inc., 360 Binney Street, Cambridge, Massachusetts 02142, United States

[▽]GlaxoSmithKline, 5 Moore Drive, RTP, Research Triangle Park, North Carolina 27709, United States

[○]Lhasa Limited, Granary Wharf House, 2 Canal Wharf, Holbeck, Leeds LS11 5PS, U.K.



ABSTRACT: Zeneth is a new software application capable of predicting degradation products derived from small molecule active pharmaceutical ingredients. This study was aimed at understanding the current status of Zeneth's predictive capabilities and assessing gaps in predictivity. Using data from 27 small molecule drug substances from five pharmaceutical companies, the evolution of Zeneth predictions through knowledge base development since 2009 was evaluated. The experimentally observed degradation products from forced degradation, accelerated, and long-term stability studies were compared to Zeneth predictions. Steady progress in predictive performance was observed as the knowledge bases grew and were refined. Over the course of the development covered within this evaluation, the ability of Zeneth to predict experimentally observed degradants increased from 31% to 54%. In particular, gaps in predictivity were noted in the areas of epimerizations, N-dealkylation of *N*-alkylheteroaromatic compounds, photochemical decarboxylations, and electrocyclic reactions. The results of this study show that knowledge base development efforts have increased the ability of Zeneth to predict relevant degradation products and aid pharmaceutical research. This study has also provided valuable information to help guide further improvements to Zeneth and its knowledge base.

KEYWORDS: *pharmaceutical forced degradation, stress testing, degradation pathways, expert system, knowledge base, prediction, Zeneth benchmark*

INTRODUCTION

Stability in Drug Development. Stability has long been recognized as critically important in the drug development process, affecting both the safety and efficacy of drugs. Stability or degradation studies are the main tool utilized to predict stability problems, develop analytical methods, identify degradation products and pathways, as well as define product shelf life. Reviews[1−13] and multiple books[14−16] have been published summarizing the chemistry of drug degradation and related strategies for assessing drug stability. Efforts to improve and streamline processes related to early identification of potential impurity problems are important to the goal of providing new, safe medicines more quickly to patients.

Stress testing (often mentioned interchangeably with the term "forced degradation") is distinguished from other stability studies

Received: June 2, 2014
Revised: September 7, 2014
Accepted: September 12, 2014
Published: October 7, 2014

(*e.g.*, accelerated and long-term stability studies) by both the severity of the conditions and the purpose. As defined by ICH Q1A,[17] stress testing includes the effect of temperatures (at least 10 °C above that for accelerated testing (*e.g.*, 50 °C, 60 °C)), humidity (*e.g.*, 75% RH or greater), and oxidative conditions on the drug substance. It is recommended that the testing should evaluate the susceptibility of the drug substance to hydrolysis across a wide range of pH values when in solution or suspension. In addition, photostability testing is an integral part of stress testing. The standard conditions for photostability testing are described in ICH Q1B.[18] The examination of degradation products under stress conditions is useful in establishing degradation pathways and developing and validating suitable analytical procedures. However, as described in ICH Q1A, "it may not be necessary to assess the formation of certain degradation products if it has been demonstrated that they are not formed under accelerated or long-term storage conditions. Results from these studies can form an integral part of the information provided to regulatory authorities."

Need for a Predictive Tool. The field of pharmaceutical degradation chemistry is relatively undeveloped and generates a limited amount of publicly available, easily accessible data. Even though a significant amount of degradation chemistry has been summarized, more comprehensive and detailed information is still highly dispersed over a number of pharmaceutical journals, none of which are dedicated to (forced) degradation. Knowledge of chemical pathways usually belongs to the domain of the practicing synthetic or organic chemist and may be obtained from chemistry textbooks.[19] However, it is not trivial to apply this knowledge to scientific problems in forced degradation, and in particular structure determination of degradation products or pathway elucidation. Therefore, a tool to support, simplify, and streamline these tasks is desirable.

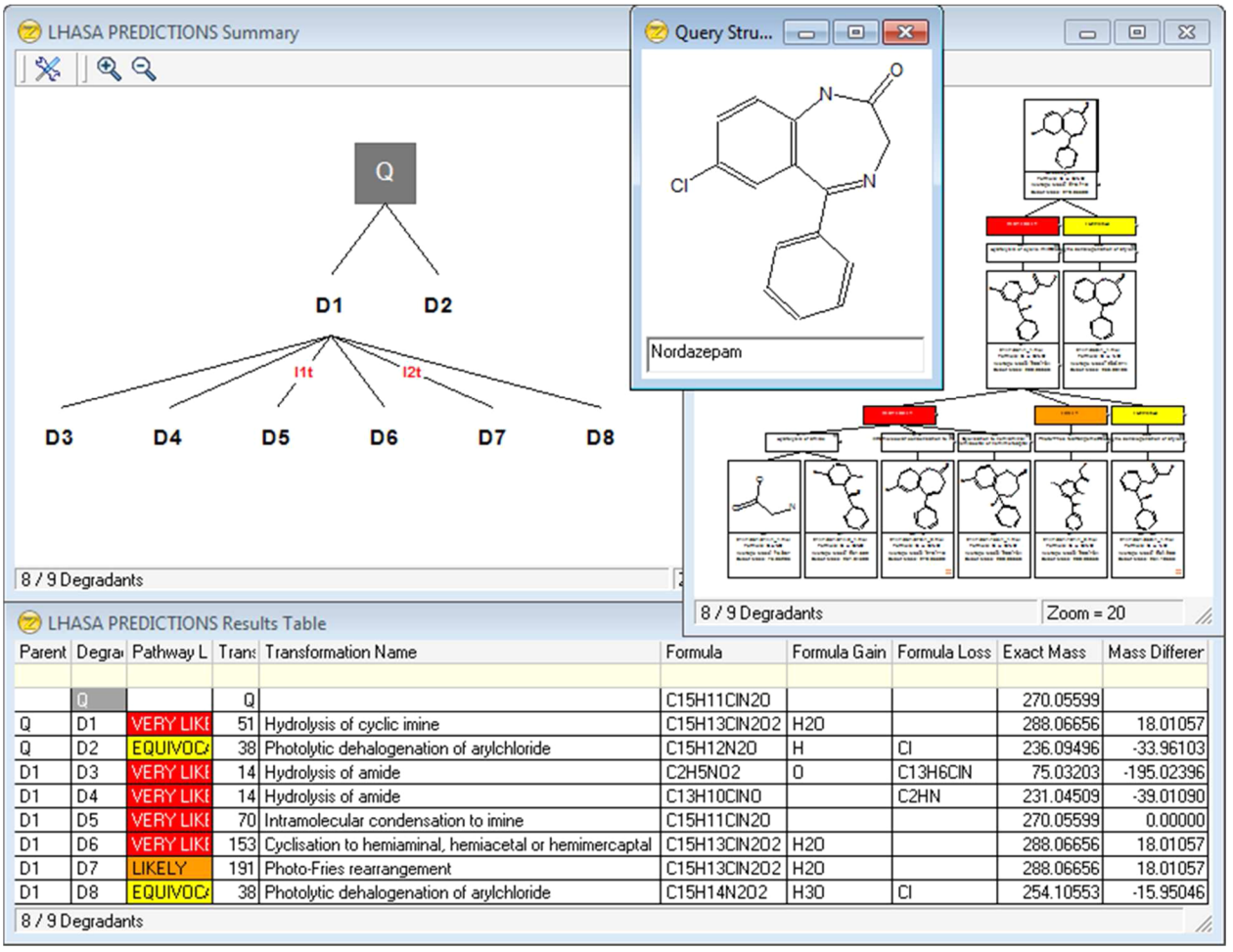
Current Status of Drug Degradation Prediction. Degradation prediction enables understanding of labile functionalities critical in designing less reactive, more stable analogs. Degradation studies conducted by a chemistry-guided[7] predictive stability approach enable analysts to deliver stability indicating methodology more efficiently. In a 2003 Stress Testing Benchmarking Survey the respondents indicated that stress testing was often first initiated in Discovery (*i.e.*, preclinical) and typically repeated at least once during the clinical phase of development.[8] Given the diverse backgrounds of typical pharmaceutical scientists, it is important that they have the proper tools to interpret the stress testing information that has been generated. To this end, significant developments have also been made in the area of degradation databases and expert knowledge systems for pharmaceutical degradation.

Recent years have seen the development of the Pharma Drug Degradation Database (Pharma D3).[20] This publicly available database allows searches (of both the parent drug compound as well as the associated degradation products) by compound name (common, generic, trade), compound number, condition, and, most importantly, functional group and change in molecular weight. These last two options make this database an excellent resource when trying to understand the degradation of apparently unrelated structures.

Advanced molecular modeling programs designed to perform quantum mechanics or force field modeling are also slowly finding applications in pharmaceutical development.[21] There have been some investigations on pharmaceutical molecules aimed at defining suitable procedures for identifying likely degradation pathways. For example, quantum mechanical procedures for determining the location of the highest occupied molecular orbital and for calculating the X−H bond dissociation energies of a typical new chemical entity have been presented.[22,23] An excellent review of the use of computational methods in stress testing is provided by Boyd and Sharp;[21] however, molecular modeling is not generally applied to stress testing, especially in the very early stages of development, due in part to the significant resources and expertise required to achieve accurate results.

Other computer-aided prediction programs[24] or expert knowledge bases have also begun to find wider application. These programs generally predict chemical degradation pathways and degradants based on the identification of functional groups in the active pharmaceutical ingredient and the known chemistry of these groups kept in a knowledge base. Typically, the only input required by the user is the chemical structure and the degradation conditions (*e.g.*, peroxide, acid, base, oxygen). CAMEO was a software tool developed by the Jorgensen group to predict the outcome of synthetic reactions.[25] This program does not generally rely on a large database of set reactions for given functional groups but instead reduces reactivity to discrete fundamental steps, such as proton transfer, nucleophilic and electrophilic additions, and eliminations to evaluate whole reactions. Unfortunately, development of this software is no longer being continued. A more recent software example is DELPHI, which was developed at Pfizer.[26] This program builds up degradation pathways by identifying reactive functional groups and successively applying known reactions from an internal knowledge base. This particular expert knowledge base was not publicly available, but the approach had a number of advantages. Since the degradation pathways were based on a series of known reactions with the potential to be easily accessed from the knowledge base with references and comments, the software could have become a teaching tool, educating the user about the known chemistry of a given molecule. In addition, the possibility existed to easily build in institutional knowledge. For example, if a molecular scaffold of interest to an organization underwent a specific rearrangement, that reaction could be added to the knowledge base. Unfortunately, software experts were required to perform knowledge base updates and DELPHI required additional compatibility updates. As a result, Pfizer has since retired the DELPHI program.

Zeneth is the only commercially available program designed to predict degradation pathways of pharmaceutical compounds.[27] It is built on the Meteor metabolite prediction platform.[28] Zeneth employs a reasoning engine with a high-quality knowledge base that has been developed largely from public sources. The first commercial version of this software was released at the end of 2010. In addition to the core knowledge base, Zeneth includes options for building private institutional knowledge. There are detailed tree depictions showing chemical degradation pathways that have been implemented, including chemical formula, exact mass, degradation pathway description, as well as the predicted likelihood of occurrence. A reasoning engine uses the rules present in the knowledge base to estimate the likelihood of the reactions with due consideration of the reaction conditions (such as pH). Predicted degradant structures are displayed along with predicted likelihood, the transformation name, empirical formula, and exact mass. In addition, several result filters (which are applied subsequent to the prediction) allow the refinement of the degradation tree by highlighting reaction pathways that lead to degradants reported with specific criteria such as an observed or theoretical exact mass or molecular formula (*cf.* Figure 1).



Parent	Degra	Pathway L	Trans	Transformation Name	Formula	Formula Gain	Formula Loss	Exact Mass	Mass Differer
	Q		Q		C15H11ClN2O			270.05599	
Q	D1	VERY LIKE	51	Hydrolysis of cyclic imine	C15H13ClN2O2	H2O		288.06656	18.01057
Q	D2	EQUIVOCA	38	Photolytic dehalogenation of arylchloride	C15H12N2O	H	Cl	236.09496	-33.96103
D1	D3	VERY LIKE	14	Hydrolysis of amide	C2H5NO2	O	C13H6ClN	75.03203	-195.02396
D1	D4	VERY LIKE	14	Hydrolysis of amide	C13H10ClNO		C2HN	231.04509	-39.01090
D1	D5	VERY LIKE	70	Intramolecular condensation to imine	C15H11ClN2O			270.05599	0.00000
D1	D6	VERY LIKE	153	Cyclisation to hemiaminal, hemiacetal or hemimercaptal	C15H13ClN2O2	H2O		288.06656	18.01057
D1	D7	LIKELY	191	Photo-Fries rearrangement	C15H13ClN2O2	H2O		288.06656	18.01057
D1	D8	EQUIVOCA	38	Photolytic dehalogenation of arylchloride	C15H14N2O2	H3O	Cl	254.10553	-15.95046

Figure 1. Zeneth summary and detailed results tables for nordazepam as displayed by Zeneth for pH 1, water and light conditions. Q is the query compound and D_n are the degradants.

Chemistry Coverage of Zeneth. The development of Zeneth's knowledge base started with general degradation chemistry as described in several books.[14,16,28−32] It was then expanded to include primary literature sources, general chemistry textbooks, the web-based Pharmaceutical Drug Degradation Database (Pharma D3),[20] and pathways suggested by Zeneth users. New knowledge is incorporated on an ongoing basis, with the initial focus on the most general and frequently encountered chemistry. The knowledge base now contains a diverse range of chemistries for many common reaction types (Table 1).

Table 2 provides an overview of the historical knowledge base releases, including those used in the present study.

The knowledge base has now reached some degree of maturity, as many degradation products are now correctly predicted. Continuous improvement takes place with the implementation of new, specialized pathways and through scope expansion/refinement of existing transformations.

Study Objectives and Performance Measures. The objective of the present study is to (1) evaluate Zeneth in terms of its predictive power compared to experimentally conducted stability studies (long-term, accelerated and forced degradation); (2) assess the practical growth of the knowledge base; and (3) provide guidance toward further improvement of the system. To this end, five pharmaceutical companies have shared confidential experimental data from Zeneth predictions as well as corresponding forced degradation, accelerated, and long-term stability studies, with Lhasa Limited, which then collated and summarized the data.

Table 1. Constitution of the Zeneth Knowledge Base According to the Type of Reaction[a]

	Number of transformations present in knowledge base versions		
Transformation category	2010.2.0	2011.2.0	2012.2.0
Oxidations	41	56	88
Hydrolyses	44	55	61
Condensations and additions	27	42	52
Isomerizations and rearrangements	19	23	27
Eliminations and fragmentations	17	20	25
Photochemical reactions	14	19	24
Total	**162**	**215**	**277**

[a]Reactions with singlet oxygen are classified as oxidations rather than as photochemical reactions.

The present study is primarily concerned with sensitivity (*i.e.*, the number of experimentally observed degradants that are predicted by the system). Specificity (*i.e.*, the number of predicted degradants observed experimentally) is not addressed here for the reason outlined below. In general, Zeneth predicts many more potential degradants than are observed in practice, and the number grows substantially as a function of the number of pathway "steps". Many of these Zeneth-predicted degradants

dx.doi.org/10.1021/mp5003976 | *Mol. Pharmaceutics* 2014, 11, 4179−4188

Table 2. List of the Knowledge Base Versions Used in the Present Study[a]

Version	Program Version	Release Date	Number of Transformations	Used in this Study
2009.1.0	**1**	June 2009	80	Yes
2009.2.0	**2**	December 2009	109	Yes
2010.2.0	**3**	December 2010	162	No
2011.1.0	**3**	June 2011	183	Yes
2011.2.0	**4**	December 2011	215	Yes
2012.1.1	**4**	July 2012	243	Yes
2012.1.3/2012.2.0	**5**	October 2012/ December 2012	277	Yes

[a]Version 2010.2.0 was not used in the present study, as version 2011.1.0 was only slightly larger. Versions 2012.1.3 and 2012.2.0 are functionally equivalent (2012.1.3 is the beta version of 2012.2.0).

may actually be formed experimentally, but not in sufficient quantity to enable routine detection or justify their identification (below typical ICH Q3A[33]/Q3B[34] thresholds). Therefore, any determined specificity would be influenced by the experimental design for the stress test and stability studies performed by each company. Additionally, the current primary use of the software is to aid in the elucidation of structures and pathways that are observed empirically. To achieve this goal, a low specificity is not a significant impediment. As a high sensitivity is important, this is the main criterion considered in the present study.

One of the main benefits of a benchmarking study such as this one is that it allows the progress of the predictions to be tracked. This progress can be the result of expansion of the knowledge base and/or that of program functionality. For instance, the first version of Zeneth (Zeneth 1) could not perform predictions of intermolecular reactions. The capability to predict dimerization, reactions with excipients, or reactions with excipient impurities was only added with Zeneth 4 (cf. Table 2). This study can also reveal areas where predictivity has not improved during development, providing a focus for future improvements. In general, there are five possible reasons why Zeneth would fail to make a correct prediction of a particular degradation reaction (assuming a particular minimum likelihood and a particular maximum number of reaction steps):

1. The relevant transformation is absent from the knowledge base.
2. The transformation is present in the knowledge base, but the reaction is outside its scope (transformations are written with specific atomic inclusions and exclusions; see below).
3. The transformation is present in the knowledge base and the reaction is within its scope, but the predicted likelihood is too low and is hence filtered out of the results.
4. The degradation reactions are expressed as a sequence of transformations, and there are too many steps in the pathway.
5. There are too many predicted degradants leading to the degradation tree being truncated, and the degradation pathway is discarded.

Consequently, the goal of this study was to assess the sensitivity of Zeneth relative to the outputs of experimentally conducted forced degradation, accelerated, and long-term stability studies using examples from small molecule new chemical entities in the pharmaceutical industry.

EXPERIMENTAL SECTION

Selection of Zeneth Parameters. Five pharmaceutical companies have participated in the benchmark study, each with 4−10 small molecule, organic, active pharmaceutical ingredients, for a total of 27 compounds for which forced degradation, accelerated, and long-term studies had been carried out. Table 3 gives an overview of the various functional groups and ring systems that are represented by the compounds in the study. All compounds selected for this study underwent appropriated drug substance forced degradation,[8,12] drug substance solid-state stability,[17] and drug product stability[33] studies to support development up to phase 2 or later.

Table 3. Counts of Functional Groups Present in the Benchmark Compounds[a]

Functional group	Count	Functional group	Count
1,3-Diene	1	Ether—aliphatic−aromatic	10
Acetal	1	Ether—aromatic−aromatic	4
Alcohol—aliphatic, primary	2	Fluoride—aromatic	5
Alcohol—aliphatic, secondary	2	Hydrazone	3
Alcohol—aliphatic, tertiary	1	Imide	1
Alcohol—phenol	2	Indole	5
Alkene—1,2-disubstituted	2	Ketone—aliphatic−aliphatic	1
Alkene—tetrasubstituted	1	Ketone—aromatic−aromatic	1
α-Aminoether	1	Naphthalene	2
Amide—secondary	5	Nitrile	3
Amide—tertiary	4	Other fused aromatic rings	4
Amidine	2	Other heterocycles	5
Amine—aliphatic, secondary	5	Oxazole	1
Amine—aliphatic, tertiary	7	Piperazine	5
Amine—aromatic, primary	4	Piperidine	5
Amine—aromatic, secondary	7	Purine	2
Amine—aromatic, tertiary	1	Pyrazole	2
Benzene	27	Pyridine	8
Benzothiophene	1	Pyrimidine	1
Carboxylic acid	10	Quinoline	1
Chloride—aromatic	7	Sulfonamide	2
Cyclopentane/hexane	3	Thiazole	1
Cyclopropane	3	Thioether	1
Enamide	3	Thiophene	3
Enol ether	1	Triazoles	2
Ester	3	Trifluoromethyl	5

[a]Specific aromatic ring types have been counted as isolated rings (*e.g.*, benzene, pyrrole) and separate from fused aromatic ring systems (*e.g.*, indole, naphthalene, "other"). The category "other heterocycles" contains all nonaromatic heterocycles (fused or not) not listed elsewhere.

The main input for Zeneth is the structure of the query compound. In addition, the prediction is controlled by parameters called processing constraints. These constraints include the minimum level of likelihood, the number of reaction steps (length of pathway), the reaction conditions, inclusion or exclusion of dimerization reactions, and whether tautomerization of reacting compounds is considered. Within the constraints, the program finds all applicable transformations in the knowledge base and generates a first level (*i.e.*, step) of degradation

Table 4. Forced Degradation Parameters and Processing Constraints Selected in Zeneth[a]

	Parameter	Values
Forced degradation conditions	Temperature	80 °C
	pH	1 and 13
	Hydrolysis	Water
	Oxidation	Oxygen
		Metal initiated
		Radical initiated
		Peroxide
	Photolysis	Light
Software parameters	Maximum number of degradants	400
	Steps	Initially 2 (4 for additional confirmation)
	Tautomers	Consider
	Reaction constraint (dimerization options)	Q (Versions 1−3); Q, Q+Q, Q+D (Versions 4−5)
	Absolute reasoning	Equivocal (and above)
	Relative reasoning	Off

[a]Q is the query compound (*e.g.*, chemical structure of the active pharmaceutical ingredient), and D_n are degradation products.

products. If more than one step is allowed, each individual predicted degradant will be processed in the same way as the drug substance, and this will be repeated until the number of steps defined by the user has been reached.[27] The choice of processing constraints was kept constant across all Zeneth predictions to ensure consistent analysis across data sets as well as for optimizing sensitivity. The forced degradation conditions and processing constraints in Zeneth that were modeled in this study are listed in Table 4. The parameters were chosen to give the program a reasonable ability to predict the largest number of observed major and minor degradants under standard forced degradation conditions, while still not constraining the program to look for specific degradation products.

For the Zeneth analysis of each compound, all the forced degradation conditions were selected in the software simultaneously. As a result, the total predictions of a given analysis result from all conditions together, even though an individual pathway step may be the result of only a single condition. Note that each transformation in the knowledge base has its own condition requirement. Thus, when in step 1 a degradation product is formed from a single transformation (as a result of a particular condition), that same degradation product may react in step 2 as if directly exposed to any of the conditions. The prediction was run twice, once in acidic (pH 1) and once in a basic (pH 13) environment. Currently, there are no phase-specific transformations, so the solid/solution state setting was not set. The API structure (query compound, Q) is allowed to react with its environment and with itself (Q+Q) in the environment and with its degradation products (Q+D). Reactions with counterions, excipients, and/or their impurities or reaction between degradants were not considered in the study, but are now possible in the more recent versions of Zeneth. Degradation pathways were limited to 400 structures and 2 steps in the first pass. If more than 400 degradants were predicted in the first pass, the application stopped further prediction (truncation of the degradation tree). This is a practical limitation based on computing/processing time and the user's ability to interrogate the output. In this study, there were 2 query structures where the degradation predictions were limited to the user-defined constraint of 400 degradants. Additionally, to further assess the sensitivity and to explore situations where some degradation products were observed experimentally but not predicted, Zeneth was allowed to predict the next 2 steps (steps 3−4) in the reaction paths (although still limited to 400 degradants). Knowledge base development was carried out independently from the benchmark study; improvements in the knowledge base and additional transformations were not aimed at improving the predictions of the benchmark study. Instead, separate knowledge base improvements and prioritization brainstorming sessions were undertaken. However, the results of this benchmarking effort are now being used to guide ongoing knowledge base development.

RESULTS

Table 5 shows the summarized results for the 27 compounds of the benchmark study.

From Table 5, it was noted that the predictive performance of the software showed steady progress with development of the knowledge base. Through the most recent version of the knowledge base included here, several APIs (4) had all of the observed degradation products predicted by Zeneth, whereas five compounds showed poor predictive performance (fewer than 30% of their degradants were predicted) that did not improve over time. For instance, compound 14 contained an uncommon/unusual heterocycle that was not within the focus of initial knowledge base development. There were another three compounds that scored below 40% as well. Interestingly, for five of the compounds, the predictive performance went *down* between the two earliest knowledge base versions. This observation can be explained through changes in the knowledge base through development and refinement. In the earlier knowledge base, *N*-hydroxylation transformations were deemed likely. Their products would give (in two further steps) N-dealkylated products, which were observed in the experimental *in vitro* stability studies. Feedback from the Zeneth users at that time suggested that the likelihood of these reactions was too high relative to practical observations. Consequently, the likelihood was reduced in version 2009.2.0, which resulted in the degradation products no longer being predicted. In more recent knowledge base versions, N-dealkylated products are predicted again, this time via a newly added transformation (C-oxidation alpha to a nitrogen atom; *e.g.*, formation of iminium ion and/or hemiaminal). For two compounds (5 and 8) the predictive performance decreased with the latest knowledge base version. This was caused by the fact that predictions with more recent knowledge base versions are more likely to generate more than the predefined user-set limit of 400 degradants, causing the prediction to be truncated. Degradants that were found

dx.doi.org/10.1021/mp5003976 | *Mol. Pharmaceutics* 2014, 11, 4179−4188

Table 5. Summary of Benchmark Study Results[a]

		No. Obsd Degradants Correctly Predicted by Zeneth, Using Each Knowledge Base					
Compd ID	No. Obsd Degradants	2009.1.0	2009.2.0	2011.1.0	2011.2.0	2012.1.1	2012.2.0
1	11*	3	3	3	3	3	4
2	7	3	2	4	4	5	5
3	9	4	5	5	5	5	6
4	4	1	1	1	1	1	1
5	10	2	1	7	5	6	5 (6)
6	6	1	3	2	3	2	3
7	2	0	1	1	1	1	1
8	18	4	7	11	13	14	11 (14)
9	8	0	0	0	1	2	2
10	7*	1	1	1	1	1	1
11	11*	2	2	2	4	5	5
12	6	2	2	2	4	4	4
13	8	3	0	3	6	6	6
14	9*	1	0	0	0	1	1
15	7*	2	2	4	4	5	5
16	6	4	4	4	4	5	4
17	7	3	3	6	6	7	7
18	15*	6	6	6	5	5	5
19	7	2	2	2	2	3	3
20	5	3	4	4	4	4	4
21	3	2	2	2	2	2	2
22	3	3	3	3	3	3	3
23	5	0	1	1	1	1	1
24	6	3	2	3	3	5	5
25	3	1	1	1	1	1	1
26	3	1	1	1	2	2	3
27	5	2	2	5	5	5	5
Total	191	59	61	84	93	104	104 (108)
Percentage of all obsd degradants combined that were predicted by Zeneth		31%	32%	44%	49%	54%	54% (57%)

[a]Each row depicts the number of observed degradation products that were predicted by Zeneth for each compound for each knowledge base. For compounds #5 and #8 and for the total count, the number in parentheses represents the number of degradants that would be correctly predicted if the degradation tree was not truncated at 400. Observed degradants that arise from interactions with excipients or counterions are included in totals noted with an asterisk (*).

previously ended up in the discarded section of the tree and were effectively "lost". In particular, degradation products formed via longer predictive paths (*e.g.*, 3 steps) were not predicted for compound 6.

The maximum number of unique degradants was chosen to be 400 in this study, but it can be set as high as 1000. During the benchmarking study one trimerization and four dimerization products were observed experimentally, of which Zeneth predicted three dimerization products, all within two steps (the knowledge base did not contain information affording the formation of the remaining two products). Therefore, in this study no predictions would have been lost by omitting dimerizations when processing with longer pathways.

Figure 2 shows the percentage of correctly predicted degradants for each of the 27 compounds listed in Table 5. For clarity, results are shown for three of six knowledge bases that were used in the benchmarking study and are overlaid with a plotted line showing the progressive increase in the number of transformations contained in those knowledge bases. Using the first knowledge base, on average, Zeneth predicted only 31% of observed degradants for any given compound, and that average increased to 54% for the latest knowledge base assessed in the study. Figure 2 also shows that as the size of the knowledge base increased, on average the software was able to correctly predict a greater percentage of empirically observed degradants, as represented by the black bars in the figure. It is worth noting that the three knowledge bases that are not presented had similar increases in the number of transformations (29−34, cf. Table 2) from the immediate precursor knowledge base.

Table 6 summarizes the prediction results using the most recent knowledge base available in this study (version 2012.2.0). Of the 191 empirically observed degradation products, 104 (54%) were predicted by Zeneth within the parameters set for the benchmark study. It is interesting to note that most of these (97, or 51%) were predicted within two steps. Only four compounds were actually processed with more than two steps, yielding an additional seven predicted degradation products. The remaining 3-step and 4-step degradants were found by processing the compounds with the dimerization options switched off and the maximum number of degradants set to 1000.

The largest number of experimentally observed degradants that were not predicted by Zeneth (57 or 30%) was caused by the knowledge base not containing the relevant transformation(s). A smaller number (13 or 7%) was caused by the scope of the relevant transformation(s) being too narrow. These deficiencies provide good indicators where the knowledge base needs expansion and improvement.[27] Some of these areas are epimerization, N-dealkylation of *N*-alkylheteroaromatic

dx.doi.org/10.1021/mp5003976 | *Mol. Pharmaceutics* 2014, 11, 4179−4188

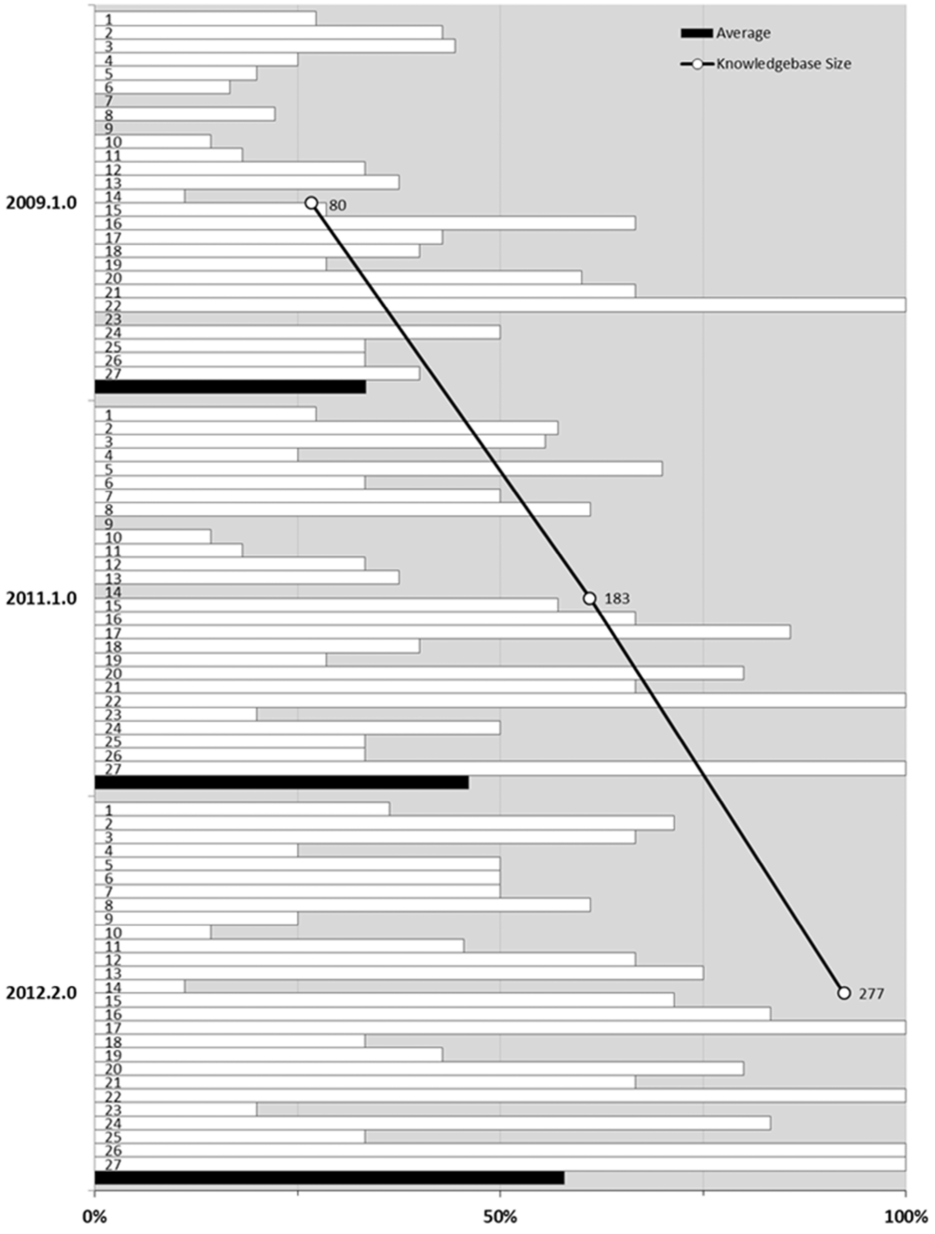


Figure 2. Efficacy of Zeneth predictions. Each white bar represents the percentage of correctly predicted degradants for each of the 27 compounds listed in Table 5, and the black bar represents the average value for all 27 compounds calculated using the equation:

$$\text{Average Compound Prediction} = \frac{1}{27}\sum_{i=1}^{27}\frac{(\text{Predicted Degradants})_i}{(\text{Obsd Degradants})_i} \times 100\%$$

Results are shown for three of six knowledge bases that were used in the benchmarking study and are overlaid with a line showing the progressive increase in the number of transformations contained in those knowledge bases.

compounds, photochemical decarboxylation, and (photochemical) electrocyclic reactions.

The functionality to deal with excipients and their impurities was added relatively recently to the Zeneth program[27] and was therefore not within the focus of this study, which mainly considers knowledge base development. However, only a limited number (6) of such interactions were observed in the current study. Three compounds had a degradation product that originated from a condensation with formaldehyde or formic acid. These products are predicted by Zeneth 5 (knowledge base 2012.2.0), where the capability to include interactions with excipients and their associated impurities has been built into the Zeneth software. Three other compounds had a degradation product that originated from an interaction with chloride ion. The knowledge base did not contain the transformations to make these predictions, and so these observations point toward an additional area that needs to be added to the knowledge base. It is worth noting that Zeneth was able to predict 50% of the observed interactions between the compounds and excipients, counterions, or impurities. As noted in the results (Table 6) these products are included as "not predicted" and result in an underestimation (by 3 additional correct predictions) of the actual sensitivity.

dx.doi.org/10.1021/mp5003976 | *Mol. Pharmaceutics* 2014, 11, 4179−4188

Table 6. Summary of Degradation Prediction Results with Knowledge Base Version 2012.2.0

Degradant category	Number of Degradants
Observed	191
Predicted in step 1	66
Predicted in step 2	31
Predicted in step 3	7
Not predicted—absent from knowledge base	57
Not predicted—outside scope of transformation	13
Not predicted—pathway too long/truncated from degradation tree (outside of benchmark parameters)	11
Not predicted—interaction with excipient impurity or counterion (outside of benchmark parameters)	6
Predicted within benchmark parameters	104

Areas for Further Development. A large number of potential degradation products are generated by Zeneth. In some cases, it makes the likely major degradants difficult to discern from those less likely to form. This is not a problem when the system is used to find pathways to specific structures or structures with a specific mass (use of Zeneth to help identify peaks observed by LC-MS); however, whenever the number of degradants exceeds the user preset maximum, such degradants may end up beyond the cutoff and disappear from the prediction results. This problem can be addressed in various ways. One method that may be utilized is to increase the specificity of the system through reduction of false negative predictions. This is a challenging approach, requiring a thorough analysis of why some predicted degradants should be considered less likely. Such reasoned likelihoods may then be implemented in additional (reasoning) rules. A second method that may help identify the major degradants that Zeneth currently predicts to form in multiple steps is to reduce the individual transformation to a single step in the knowledge base (with the intermediate structures embedded within the transformation). Zeneth would then be able to predict such products at the first level of degradants, thus reducing the need for processing using longer pathways with their correspondingly (very) large number of degradants. These developments have already been initiated by Lhasa Limited.

One other rationale for the lack of predictivity by Zeneth is that the scope of transformations is initially based on a particular structural class of compounds, but the query compound structure is from a different structural class. In these cases, it may be feasible to expand the scope of the transformation in the knowledge base. As an example, the knowledge base contains the transformation "Hydrolysis of *N*-sulfonylurea". The original scope was restricted to carbon- or hydrogen-substitution on the nitrogen atoms. However, closer examination of this chemistry suggested that the scope should be modified to include a larger class of *N*-sulfonylurea-type compounds (Figure 3). Moreover, it was noted that N-sulfonylamides (a related compound class) would more readily undergo the same hydrolysis. Therefore, a new transformation was written for their hydrolysis (Figure 3).

In order to improve upon the 54% prediction success rate, additional transformations would need to be added along with the respective reasoning rules. From a user perspective, prediction of the *relative* amount of each degradation product in a forced degradation study may be considered to be a key objective. The prediction of relative abundance is complex, since the rate of formation of a degradant does not simply depend on whether a particular transformation (*i.e.*, degradation pathway) is

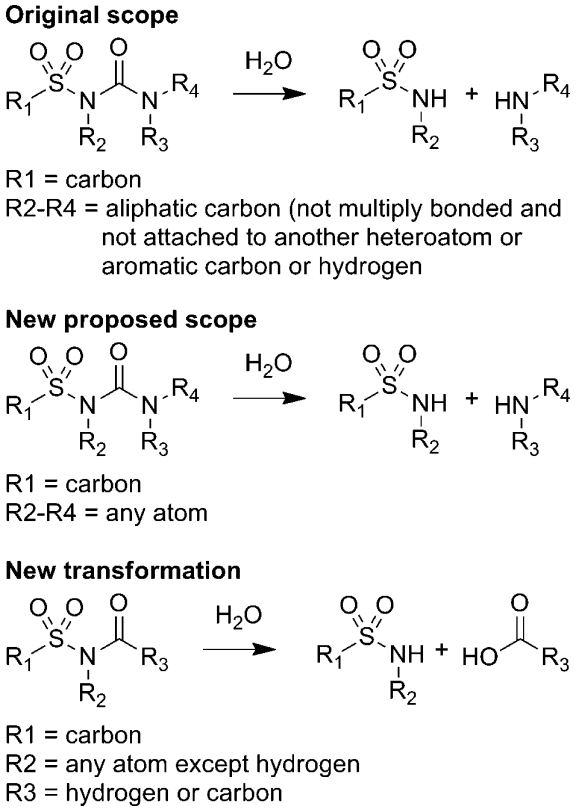


Figure 3. Scope extension of the transformation "Hydrolysis of *N*-sulfonylurea" to yield a new transformation "Hydrolysis of *N*-sulfonylamide or related compound". R3 is limited to groups that are poorer leaving groups than sulfonylamide.

possible based on reaction conditions and absolute reasoning rules. Within Zeneth, the absolute reasoning rules assign the likelihood of a transformation in five levels, from very likely to very unlikely depending on the reaction type and conditions. In the case of competing pathways, faster transformations (kinetic control) or more thermodynamically favored reactions may deplete a shared reagent (*e.g.*, oxygen) and prevent less favored reactions from taking place. Consequently, the prediction of relative abundance requires additional information that is not currently input into Zeneth.

Relative reasoning rules are also present in the knowledge base and describe the relative reactivity of functional groups. At this time these rules are based on commonly known reactivity principles of the functional groups present and the experience of current Zeneth users and forced degradation practitioners. While the authors have discussed the possibility of including calculations of physical chemical properties (*i.e.*, bond dissociation energy, polarizability, strain, pK_a, UV−vis absorption spectrum, and enthalpy of formation) that may support improved prediction accuracy and prediction of relative abundances, there are no immediate plans for their inclusion in the next version of Zeneth.

Currently, the authors use Zeneth in conjunction with their experimental forced degradation studies. Zeneth is used to augment the knowledge gained in the *in vitro* studies, as it aids the identification of degradant structures and elucidation of pathways. It accomplishes this by affording access to pre-existing knowledge/rules and also helps to bridge knowledge to new compounds in a series (for example, where degradation is understood for one compound but not for a related compound),

as well as to assess formulation development and excipient selection. Zeneth also affords the mass spectrometrist the ability to scan through likely degradation products to help match up the empirically observed peaks with specific m/z values. It is envisioned that Zeneth can continue to be used in this role and that it is not expected to supplant experimental forced degradation studies in the near future.

CONCLUSION

Zeneth is a new software application capable of predicting potential degradation products derived from small molecule active pharmaceutical ingredients. This study was aimed at understanding the current status of Zeneth's capabilities and assessing gaps in predictivity. Steady progress in predictivity was observed as the knowledge bases grew and were developed over several revisions. An average of 54% of the observed degradants from forced degradation or pharmaceutical stability studies (ICH long-term or accelerated stability studies) was predicted by Zeneth, up from 31% in the first knowledge base. Zeneth is currently utilized in a number of settings to help elucidate the identity of experimentally observed degradation products and their formation pathways. It affords access to pre-existing knowledge and reactivity rules to help assess formulation development and excipient compatibility. Zeneth also provides the capability to pair empirically observed peaks by LC-MS (m/z) with those predicted by the software. This simplifies the tentative assignment of structures to degradation products. Additionally, gaps (incomplete predictivity) were noted in epimerizations, N-dealkylation of *N*-alkylheteroaromatic compounds, photochemical decarboxylations, and electrocyclic reactions. The knowledge base can be enhanced by closing these identified gaps, expanding the scope of transformations where it is too narrow, and combining multiple steps within a single transformation to arrive at primary degradation products more effectively. Some of these developments are already underway.[35]

AUTHOR INFORMATION

Corresponding Author

*Telephone: 610-787-5703.

Present Address

●(C.F.) Gilead Sciences, 333 Lakeside Drive, Foster City CA, 94404.

Notes

The authors declare the following competing financial interest(s): Martin Ott is a full-time employee of Lhasa Limited, which develops the Zeneth software.

ACKNOWLEDGMENTS

R.S. and D.L.R. would like to thank Drs. Rick Chiu, Janan Jona, Minhui Ma, Jessica Tan, Charles Yang, and Peter Zhou for their assistance in compiling experimental statistics for Amgen's contribution to this publication. M.H.K. and D.W.R. thank Sonya Kennedy-Gabb and Simon Walsh for conducting forced degradation studies. M.D.M. and M.N. acknowledge Li Li for her assistance with the degradation analysis.

REFERENCES

(1) Waterman, K. C.; Adami, R. C.; Alsante, K. M.; Hong, J.; Landis, M. S.; Lombardo, F.; Roberts, C. J. Stabilization of pharmaceuticals to oxidative degradation. *Pharm. Dev. Technol.* **2002**, *7* (1), 1−32.

(2) Foti, C.; Alsante, K.; Cheng, G.; Zelesky, T.; Zell, M. Tools and workflow for structure elucidation of drug degradation products. *Trends Anal. Chem.* **2013**, *49*, 89−99.

(3) Singh, S.; Junwal, M.; Modhe, G.; Tiwari, H.; Kurmi, M.; Parashar, N.; Sidduri, P. Forced degradation studies to assess the stability of drugs and products. *Trends Anal. Chem.* **2013**, *49*, 71−88.

(4) Hovorka, S.; Schoneich, C. Oxidative degradation of pharmaceuticals: theory, mechanisms and inhibition. *J. Pharm. Sci.* **2001**, *90* (53), 253−269.

(5) Waterman, K. C.; Adami, R. C.; Alsante, K. M.; Antipas, A. S.; Arenson, D. R.; Carrier, R.; Hong, J.; Landis, M. S.; Lombardo, F.; Shah, J. C.; Shalaev, E.; Smith, S. W.; Wang, H. Hydrolysis in pharmaceutical formulations. *Pharm. Dev. Technol.* **2002**, *7* (2), 113−146.

(6) Byrn, S. R.; Xu, W.; Newman, A. W. Chemical reactivity in solid-state pharmaceuticals: formulation implications. *Adv. Drug Delivery Rev.* **2001**, *48* (1), 115−136.

(7) Alsante, K. M.; Ando, A.; Brown, R.; Ensing, J.; Hatajik, T. D.; Kong, W.; Tsuda, Y. The role of degradant profiling in active pharmaceutical ingredients and drug products. *Adv. Drug Delivery Rev.* **2007**, *59* (1), 29−37.

(8) Alsante, K. M.; Martin, L.; Baertschi, S. W. A Stress Testing Benchmarking Study. *Pharm. Technol.* **2003**, *27* (2), 60−72.

(9) Baertschi, S. W. Analytical methodologies for discovering and profiling degradation-related impurities. *Trends Anal. Chem.* **2006**, *25* (8), 758−767.

(10) Jackson, R. A. A Forced Degradation Strategy: How Johnson & Johnson PRD Approaches Forced Degradation. *Am. Pharm. Rev.* **2007**, *10*, 59−63.

(11) Joshi, B. K.; Reynolds, D. W. Regulatory and Experimental Aspects of Oxidative Stress Testing of Pharmaceuticals. *Am. Pharm. Rev.* **2005**, *8*, 105−107.

(12) Reynolds, D. W.; Facchine, K. L.; Mullaney, J. F.; Alsante, K. M.; Hatajik, T. D.; Motto, M. G. Available guidance and best practices for conducting forced degradation studies. *Pharm. Technol.* **2002**, *26* (2), 48−56.

(13) Reynolds, D. W. Forced Degradation of Pharmaceuticals. *Am. Pharm. Rev.* **2004**, *7*, 56−61.

(14) Min, L. *Organic Chemistry of Drug Degradation*; Royal Society of Chemistry: Cambridge, 2012.

(15) Carstensen, J. T.; Rhodes, C. T. *Drug Stability, 3rd ed., Revised, and Expanded: Principles and Practices*; Taylor & Francis: New York, 2000.

(16) Baertschi, S. W.; Alsante, K. M.; Reed, R. A. *Pharmaceutical Stress Testing: Predicting Drug Degradation*; Informa Healthcare: London, 2011.

(17) ICH. *Q1A(R2): Stability Testing of New Drug Substances and Products* (Second Revision). February 6, 2003. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q1A_R2/Step4/Q1A_R2__Guideline.pdf.

(18) ICH. *Q1B: Stability Testing: Photostability Testing of New Drug Substances and Products*. November 6, 1996. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q1B/Step4/Q1B_Guideline.pdf.

(19) Smith, M. B.; March, J. *March's Advanced Organic Chemistry: Reactions, Mechanisms, and Structure*; Wiley: 2007.

(20) *Drug Degradation Database (D3)*, d3.arxspan.com, Accessed March 30, 2014.

(21) Boyd, D. B.; Sharp, T. R. The power of computational chemistry to leverage stress testing of pharmaceuticals. In *Pharmaceutical Stress Testing: Predicting Drug Degradation*, 2nd ed.; Baertschi, S. W., Alsante, K. M., Reed, R. A., Eds.; Informa Healthcare: New York, 2011; pp 499−539.

(22) Reid, D. L.; Calvitt, C. J.; Zell, M. T.; Miller, K. G.; Kingsmill, C. A. Early prediction of pharmaceutical oxidation pathways by computational chemistry and forced degradation. *Pharm. Res.* **2004**, *21* (9), 1708−1717.

(23) Sharp, T. R. Calculated carbon-hydrogen bond dissociation enthalpies for predicting oxidative susceptibility of drug substance molecules. *Int. J. Pharm.* **2011**, *418* (2), 304−317.

(24) Chen, J. H.; Baldi, P. No Electron Left Behind: A Rule-Based Expert System To Predict Chemical Reactions and Reaction Mechanisms. *J. Chem. Inf. Model.* **2009**, *49* (9), 2034−2043.
(25) Jorgensen, W. L.; Laird, E. R.; Gushurst, A. J.; Fleischer, J. M.; Gothe, S. A.; Helson, H. E.; Paderes, G. D.; Sinclair, S. CAMEO: a program from the logical prediction of the products of organic reactions. *Pure Appl. Chem.* **1990**, *62* (10), 1921−1932.
(26) Pole, D. L.; Ando, H. Y.; Murphy, S. T. Prediction of drug degradants using DELPHI: an expert system for focusing knowledge. *Mol. Pharmaceutics* **2007**, *4* (4), 539−549.
(27) Parenty, A. D. C.; Button, W. G.; Ott, M. A. An Expert System To Predict the Forced Degradation of Organic Molecules. *Mol. Pharmaceutics* **2013**, *10* (8), 2962−2974.
(28) Marchant, C. A.; Briggs, K. A.; Long, A. In silico tools for sharing data and knowledge on toxicity and metabolism: DEREK for windows, METEOR, and VITIC. *Toxicol. Mech. Methods* **2008**, *18* (2−3), 177−187.
(29) Albini, A.; Fasani, E. *Drugs, Photochemistry and Photostability*; Royal Society of Chemistry: London, 1998.
(30) Connors, K. A.; Amidon, G. L.; Stella, V. J. *Chemical Stability of Pharmaceuticals: A Handbook for Pharmacists*, 2nd ed.; Wiley: 1986; pp 847.
(31) Tonnesen, H. H. *Photostability of Drugs and Drug Formulations*, 2nd ed.; CRC Press: Boca Raton, 2010.
(32) Yoshioka, S.; Stella, V. J. *Stability of Drugs and Dosage Forms*; Springer: New York, 2000.
(33) ICH. *Q3A(R2): Impurities in New Drug Substances*. October 25, 2006. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q3A_R2/Step4/Q3A_R2__Guideline.pdf.
(34) ICH. *Q3B(R2): Impurities in New Drug Products*. June 2, 2006. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q3B_R2/Step4/Q3B_R2__Guideline.pdf.
(35) The current version of the Zeneth knowledge base is 2014.1.0 and contains 327 transformations. Released June 13, 2014.

Ogenstad Declaration Exhibit F

Journal of Pharmaceutical and Biomedical Analysis 115 (2015) 487–501

Contents lists available at ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

journal homepage: www.elsevier.com/locate/jpba

The application of electrochemistry to pharmaceutical stability testing — Comparison with *in silico* prediction and chemical forced degradation approaches

Susana Torres[a,*], Roland Brown[a], Roman Szucs[a], Joel M. Hawkins[b], Todd Zelesky[b], Garry Scrivens[a], Alan Pettman[a], Mark R. Taylor[a]

[a] Pfizer Worldwide R+D, Ramsgate Road, Sandwich, Kent, CT139NJ, United Kingdom
[b] Pfizer Worldwide R+D, Eastern Point Road, Groton, CT, United States

ARTICLE INFO

Article history:
Received 19 May 2015
Received in revised form 3 August 2015
Accepted 9 August 2015
Available online 13 August 2015

Keywords:
Stability
Oxidation
Electrochemistry
Forced degradation
In silico methods

ABSTRACT

The aim of this study was to evaluate the use of electrochemistry to generate oxidative degradation products of a model pharmaceutical compound. The compound was oxidized at different potentials using an electrochemical flow-cell fitted with a glassy carbon working electrode, a Pd/H_2 reference electrode and a titanium auxiliary electrode. The oxidative products formed were identified and structurally characterized by LC-ESI–MS/MS using a high resolution Q-TOF mass spectrometer. Results from electrochemical oxidation using electrolytes of different pH were compared to those from chemical oxidation and from accelerated stability studies. Additionally, oxidative degradation products predicted using an *in silico* commercially available software were compared to those obtained from the various experimental methods. The electrochemical approach proved to be useful as an oxidative stress test as all of the final oxidation products observed under accelerated stability studies could be generated; previously reported reactive intermediate species were not observed most likely because the electrochemical mechanism differs from the oxidative pathway followed under accelerated stability conditions. In comparison to chemical degradation tests electrochemical degradation has the advantage of being much faster and does not require the use of strong oxidizing agents. Moreover, it enables the study of different operating parameters in short periods of time and optimisation of the reaction conditions (pH and applied potential) to achieve different oxidative products mixtures. This technique may prove useful as a stress test condition for the generation of oxidative degradation products and may help accelerate structure elucidation and development of stability indicating analytical methods.

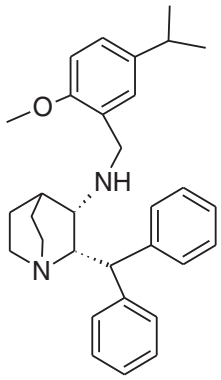
1. Introduction

Understanding the stability of a drug product is of great importance in protecting patient safety and development of robust products [1]. The launch of a drug product onto the market requires that extensive stability studies are performed in order to assure that quality, safety and efficacy are maintained and that listed shelf life is accurate and appropriate [2]. It is of great importance to be able to predict as soon as possible what may happen during the shelf life of a given pharmaceutical product as this information can significantly affect product development strategy [3]. Pharmaceutical "stress testing", often referred to as "forced degradation" studies are commonly applied in the early stages of product development in order to gain a comprehensive understanding of the possible oxidation products profiles prior to running accelerated stability studies [4] and formal ICH stability studies [5] on drug substance and drug product. Pharmaceutical stress testing studies typically include acid/base hydrolysis and various oxidation conditions for solution state stress testing and thermal, thermal/humidity and photostability for solid state stress testing [6]. Stress testing studies are important not only because they help obtain information about potential degradation products and their degradation pathways, but also help generate reference materials used in development of stability-indicating methods.

Oxidative reactions are amongst the most commonly observed chemical degradation pathways for pharmaceuticals [7,8]. Stress testing methods for generating oxidative forced degradation products include the use of hydrogen peroxide, metal ions, oxygen and

* Corresponding author.
E-mail address: susana.dasilvatorres@pfizer.com (S. Torres).

http://dx.doi.org/10.1016/j.jpba.2015.08.010
0731-7085/



1

Fig. 1. Chemical structure of the model compound ((2S,3S)-2-(diphenylmethyl)-*N*-[2-methoxy-5-(propan-2-yl) benzyl]-1-azabicyclo[2.2.2]octan-3-amine) used in this study.

radical initiators (such as 4,4-Azobis(4-cyanovaleric acid), ACVA; 2,2′-Azobis(2-methylpropionitrile), AIBN) [9,10]. These methods can result in a distribution of oxidative degradation products that are dependent on the intrinsic stability of the drug substance and the conditions used.

The use of electrochemistry (EC) to study the oxidative stability of pharmaceuticals is described in the literature by the work of Lombardo and Campos [11] and Gamache et al. [12]. By using a coulometric array detector they were able to examine the oxidation sensitivity of different drugs and rank them as stable or unstable based on the oxidation potential without making any structural characterization of the generated oxidative products. In a recently published article [13] we showed the possibility of using electrochemistry as an oxidative stress test condition and also as a means of synthetizing milligram quantities of oxidative pharmaceutical products for subsequent NMR characterization, for use as markers to aid method development and to assist in stability and bioactivity/toxicity studies.

Electrochemistry proved to be useful in the imitation of drug metabolism by cytochrome P450 (CYP) enzymes with a series of reactions being simulated electrochemically [14–16]. Benzylic hydroxylation [17], hydroxylation of aromatic rings [17,18], allylic and aliphatic hydroxylation [19], N-dealkylation [20,21], N, S-oxidation [17,22], dehydrogenation [17,23] and O-dealkylation [17,24] are examples of metabolic reactions that were mimicked using the electrochemical approach.

As a complement to forced degradation studies, *in silico* computational predictions are used to predict likely degradation products of drugs and/or drug products and their respective degradation pathways [25,26]. In this study we assess the effectiveness of electrochemistry as a method to generate the oxidation products of a model drug substance ((2S,3S)-2-(diphenylmethyl)-*N*-[2-methoxy-5-(propan-2-yl) benzyl]-1-azabicyclo[2.2.2]octan-3-amine, Fig. 1). Furthermore, we compare the results obtained by electrochemical oxidation to those obtained from forced degradation studies and from *in silico* oxidation predictions. The oxidative degradation products formed in the experimental chemical and electrochemical approaches were separated by liquid chromatography and characterised using on-line high resolution MS/MS. The experimentally recorded oxidation product profiles have been compared with previously published accelerated stability study data [27].

Table 1
Gradient profile used for the LC–UV–MS analysis. Gradient profile outside brackets was used for forced degradation studies. Gradient profile inside brackets was used for electrochemical oxidation studies to achieve better resolution of observed oxidation products.

Time/min	0.00	12.00	12.10
Acetonitrile (%)	12.0 (15.0)	90.0 (85.0)	12.0 (15.0)

2. Experimental

2.1. Chemicals and reagents

The model compound used was obtained from Pfizer Worldwide R+D, Groton, USA in the form of hydrochloride salt. Ammonium acetate (Optima® LC/MS grade), formic acid (Optima® LC/MS grade), acetonitrile (HPLC grade), methanol (HPLC grade) and trifluoroacetic acid (Optima® LC/MS grade) were purchased from Fisher Scientific (Loughborough, UK). Ammonium hydroxide solution (≥25% in H_2O), hydrogen peroxide (3 wt.% solution in H_2O), 2,2′–Azobis(2- methyl–propionitrile), and acetic acid (≥99.99%) were obtained from Sigma–Aldrich (Poole, UK). Water used for sample preparation and HPLC mobile phases was purified with a Millipore Milli-Q water purification system.

2.2. Instrumentation

2.2.1. Electrochemical oxidation

The electrochemical oxidation was performed using a 3-electrode cell (μ-PrepCell™), fitted with a glassy carbon (GC) working electrode, a Pd/H_2 (HyREF™) reference electrode and a titanium auxiliary electrode connected to a ROXY potentiostat controlled by Dialogue software (Antec, Zoeterwoude, The Netherlands).

2.2.2. LC-UV-MS analysis

A HPLC–MS system consisting of a degasser, a quaternary pump, a thermostatted column compartment, a UV diode array detector and an autosampler (all 1100 Series) with 6120 single quadrupole mass spectrometer fitted with a multimode source controlled by Agilent OpenLAB software (Waldbronn, Germany) was used.

A superficially porous reversed-phase column (Phenomenex, Kinetex 2.6 μm XB-C18, 100 Å, 150 × 4.60 mm) was used. Aqueous trifluoroacetic acid (0.05%) and acetonitrile were used as eluting solvents and separation was achieved by means of a gradient profile described in Table 1. The flow rate was set to 1.0 mL/min, the injection volume was either 5.0 μL (forced degradation studies) or 50 μL (electrochemical oxidation studies). The temperature of the column was maintained either at 40 °C (forced degradation studies) or 50 °C (electrochemical oxidation studies). Compounds were detected by UV at a wavelength of 225 nm or by mass spectrometry as described below. The UV detector and MS were connected in parallel using a flow splitter.

The mass spectrometer settings for LC–MS analysis are given in Table 2.

2.2.3. LC–MS/MS experiments

A 1290 Infinity HPLC system consisting of a binary pump, thermostat, auto-sampler, thermostatted column compartment, variable wavelength UV detector and 6550-Q-TOF ion funnel mass spectrometer fitted with an electrospray source controlled by Masshunter software (Agilent, Waldbronn, Germany) was used for high resolution mass and tandem mass spectrometry experiments. The oxidation products were detected in the positive ion mode. A reversed-phase column (ACQUITY UPLC® HSS T3 1.8 μm, 2.1 × 100 mm) was used. Aqueous formic acid (0.1%) and acetonitrile were used as eluting solvents and separation was achieved

Table 2
Single quadrupole parameters used for the detection of the compound under study and its oxidation products.

Parameter	
Scan range (*m*/*z*)	100.00-1000.00
Ionization mode	Positive
Fragmentor	50
Gain	1.00
Threshold	150
Step Size	0.20
Scanning Speed (μ/sec)	945
Drying gas flow (l/min)	5.0
Nebulizer Pressure (psig)	60
Drying gas temperature (°C)	300
Vaporizer Temperature (°C)	250
Capillary voltage (V)	2100
Corona Current (μA)	4.0
Charging voltage (V)	0

Table 3
HPLC gradient profile used for the analysis performed on the Agilent Technologies 1290 Infinity HPLC system.

Time/min	0.00	1.00	8.40	9.50	9.60	12.00
Acetonitrile (%)	5.00	5.00	95.00	95.00	5.00	5.00

by means of a gradient profile described in Table 3. The flow rate was set to 0.4 mL/min, the injection volume was 0.4 μL and the temperature of the column was maintained at 45 °C. Compounds were detected by UV at a wavelength of 225 nm and by high resolution mass spectrometry using conditions as described in Table 4. An internal reference mass solution was used to ensure consistent mass accuracy and reliable elemental composition determination.

For tandem mass spectrometry (MS/MS experiments), the precursor ion of interest was selected using the quadrupole analyzer and the product ions were analyzed using a TOF analyzer. Ultra high pure nitrogen was used as collision gas. All the spectra were recorded under identical experimental conditions using collision energies of 20 eV or 40 eV.

2.3. Electrochemical oxidation

0.086 mg/mL solutions of Compound **1** (Fig. 1) dissolved in 10 mM aqueous ammonium acetate solution at three different pH values (3.9, 7.1, and 8.8) were pumped through the thin-layer cell, fitted with a glassy carbon electrode at a flow rate of 40 μL/min using a glass syringe (1.0 mL, HAMILTON CO., Nevada, USA) and syringe pump (HARVARD APPARATUS, Holliston, USA) connected with 1 meter of PEEK tubing (0.005'' ID, Fisher Scientific, Loughborough, UK). The pH of the 10 mM ammonium acetate solution was adjusted either by adding ammonium hydroxide or acetic acid. Samples analyzed at basic pH were dissolved in a mixture 10 mM ammonium acetate (pH = 8.8)/acetonitrile (1:1) to ensure solubility of the substrate. The applied potential was increased from 0 to +2 V in steps of 100 mV, and a sample of the cell effluent was collected into separate HPLC vials at each potential. The collected samples were analyzed offline by HPLC-UV-MS.

Table 4
iFunnel Q-TOF parameters used for the detection of the compound under study and its oxidation products and for the determination of exact masses in EC-LC-HRMS measurements.

Parameter	
Scan range (*m*/*z*)	100-1700.00
Ionization mode	Positive
Gas temperature (°C)	200
Drying gas flow (l/min)	11
Nebulizer pressure (psig)	35
Sheath gas temperature (°C)	200
Sheath gas flow (l/min)	12
Capillary voltage (V)	4000
Skimmer Voltage (V)	65
Nozzle voltage (V)	1000
Fragmentor (V)	400
OCT 1 RF Vpp (V)	750
Acquisition rate (spectra/s)	1

2.4. Oxidative forced degradation

Two oxidative forced degradation studies, using peroxide and a radical-initiated oxidation, were performed on Compound **1**. The oxidative forced degradation protocol is outlined in Table 5. The nominal concentration of each sample was kept at 0.86 mg/mL and all samples were protected from light using aluminum foil. For each reaction 3 different samples were prepared (drug substance exposed to the challenge condition, H_2O_2 for peroxide oxidation or AIBN for radical-initiated oxidation; challenge condition control and drug substance control). The samples prepared for peroxide oxidation were diluted using 50:50 (v/v) acetonitrile: water and 50:50 (v/v) methanol: water was used as a diluent for the radical-initiated oxidation. For the peroxide oxidation, samples were kept at room temperature and for the radical-initiated oxidation the samples were kept at 40 °C. The samples were analyzed at given time points (Table 5) by HPLC–UV–MS.

2.5. In silico predictions

In silico predictions were performed using the Zeneth 6.0 application with Knowledge Base Z2014.2.0.mdb (Lhasa Limited, Leeds, UK). This application is an expert decision support system which predicts the forced degradation pathways of organic compounds. A Zeneth prediction analysis has been conducted for the model compound for oxidative susceptibility, to assist in anticipating and/or identifying potential degradation products which may arise during stability studies. These predictions were performed in AutoZeneth mode with processing constraints set for pH 1 and pH 13 at 80 °C with equivocal setting and maximum degradation products and steps set to 400 and 1, respectively.

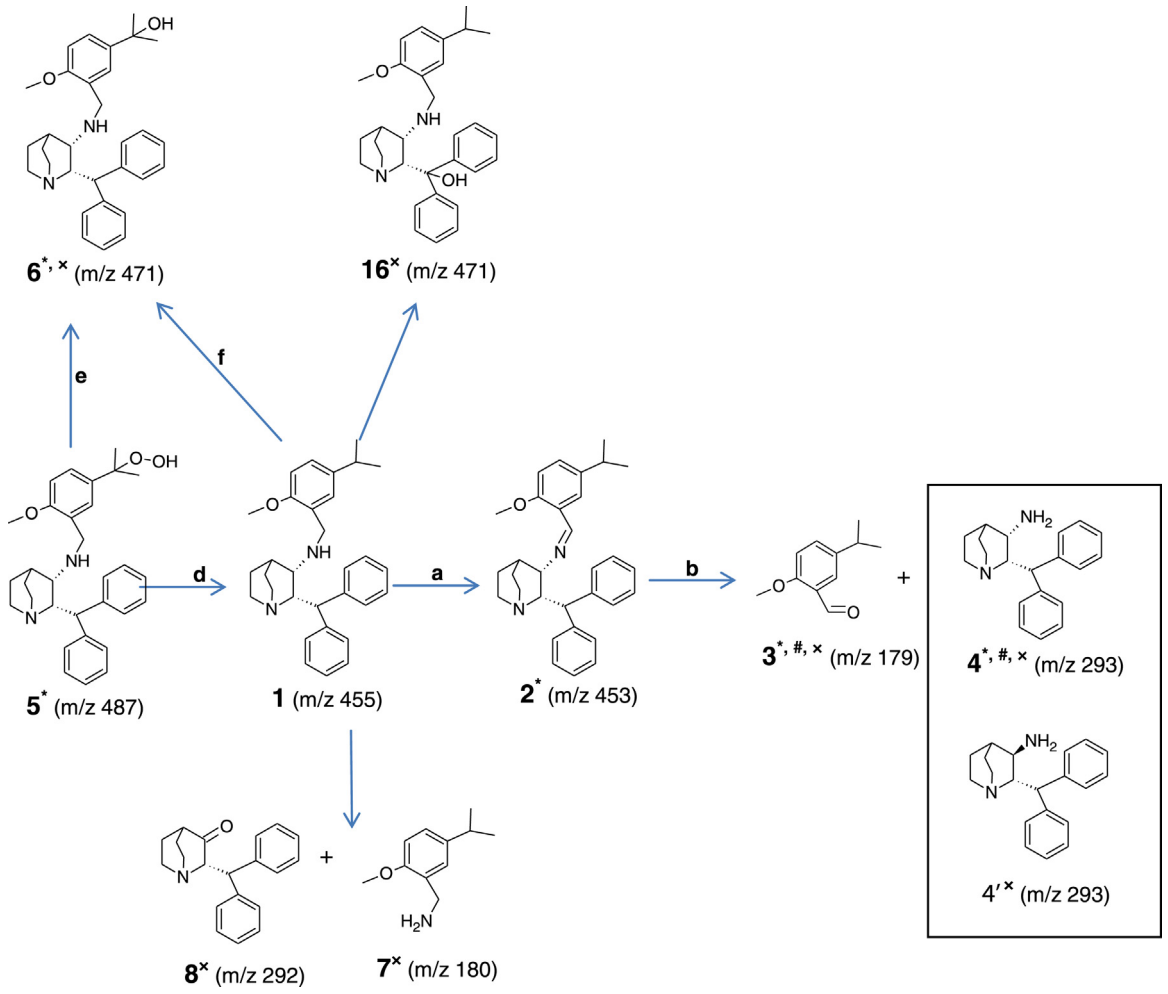
3. Results and discussion

3.1. Electrochemical oxidation

With this work we were interested in understanding the extent to which electrochemistry can be used to generate oxidative degradation products of drug substances as well as comparing it with others approaches used by the pharmaceutical industry, namely forced degradation studies and *in silico* predictions. For that purpose Compound **1** (Fig. 1) was used as model compound because

Table 5
Conditions used for the oxidative forced degradation studies.

Type of Study	Challenge condition	Temperature	Time point (hours)
Peroxide Oxidation	0.3% H_2O_2	Room Temperature	0, 4, 8, 16, 24, 48, 72
Radical Oxidation	5 mM AIBN	40 °C	0, 12, 24, 48, 72



Scheme 1. Overview of oxidation products of model compound **1** for those which was possible to elucidate their chemical structure. For a complete list of all the generated oxidation products see Table 1. Compounds **2–6** and respective degradation pathways (**a**–**e**) were obtained from accelerated stability studies (*) [27]. Compounds **3** and **4** could be generated both chemically (#) (stress testing studies) and electrochemically (×). Compound **6** and **16** were only electrochemically generated. Chemical structures for compounds **3, 4, 4′, 6, 7, 8** and **16** were derived on the basis of exact masses, and MS/MS fragmentation patterns. The *m/z* values shown correspond to the $[M+H]^+$ ions.

it has a known oxidative degradation pathway from accelerated stability studies [27]. The primary oxidative product of **1** is the imine **2** which readily hydrolyses to give products **3** and **4.** Another oxidative product is the alcohol **6** which reportedly forms via the peroxide intermediate **5** which subsequently decomposes to give alcohol **6** or can be formed directly from compound **1** (Scheme 1) [27].

Compound **1** was oxidized in solution using a commercially available flow-through electrochemical cell (μ-PrepCell™) using glassy carbon as the working electrode. With 10 mM ammonium acetate (pH 7.1) as the supporting electrolyte the electrochemical oxidation of the compound was initially studied using different cell potentials increasing from 0 to +2 V in steps of 100 mV. Samples were collected into vials using each cell potential and were analysed offline by LC–UV–MS. The electrochemical oxidation of the test compound was observed at potentials at 400 mV and above. At 400 mV dealkylation of the secondary amine was observed with consequent generation of compounds **4** and **7** (Scheme 1, oxidation products bearing a primary amine chemical moiety) and compounds **3** and **8** (Scheme 1, oxidation products bearing a carbonyl chemical moiety). Increasing the potential led to an increase on the amount of the products resulting from the dealkylation of the secondary amine. It was observed that under these experimental conditions an applied voltage of 700 mV was the optimal potential for the generation of products **7** and **8** and that for the generation of compounds **3** and **4** the optimal potential was 1200 mV. For that reason 700 mV and 1200 mV were the potential values chosen to study the effect of pH on the degradation products profile. A comparison of the products generated by electrochemistry and forced degradation studies methods is presented in Table 6. The proposed molecular formulas from LC–MS analysis are shown and the deviations between measured and calculated masses are presented in ppm. A comparison to an *in silico* approach and published accelerated stability study data [27] is also presented. For some of the electrochemically-generated products it was not possible to obtain a high resolution mass and for that reason neither a proposed molecular formula is presented nor their chemical structures are elucidated. For those cases a low resolution mass is presented. No further work was conducted to identify the formulae of these species as they were considered unlikely to be relevant to real life degradation of the compound. The peak-assigned HPLC-UV chromatograms (225 nm) from samples oxidized at 700 mv and 1200 mV cell potentials with electrolytes of three different pH values (pH = 3.9, 7.1 and 8.8) are shown in Figs. 2 and 3. A threshold of 1% of total peak area was set for selection of products to undergo further characterization by LC–MS/MS. Several small peaks arising from electrochemical degradation were observed below this threshold but were not identified. With an applied potential of 700 mV and approximately neutral conditions (pH = 7.1) (Fig. 2A) the substrate conversion was approximately 55% (based on ini-

Table 6
Overview of the oxidation products generated via pharmaceutically relevant stress testing studies (ST) and/or electrochemically (EC). A comparison with oxidation products predicted by computational approaches (*in silico*) and generated under accelerated stability studies is also presented. A proposed molecular formula and deviation between measured and calculated masses is shown for some of the oxidation products formed. Oxidation products for which a chemical structure was assigned are marked with *. (Chemical structures can be seen in Scheme 1). Compounds can be identified in the chromatograms by their numbers. (√ - generated and × - not generated)

Oxidation products	Molecular formula	Observed *m/z* ([M+H]$^+$)	Error (ppm)	EC	ST	*in silico* [25,26]	Accelerated Stability [27]
1* (Drug substance)	$C_{31}H_{38}N_2O$	455.3073	0.29	–	–	–	–
2*	$C_{31}H_{36}N_2O$	–	–	×	×	√	√
3*	$C_{11}H_{14}O_2$	179.1065	1.85	√	√	√	√
4*	$C_{20}H_{24}N_2$	293.2011	0.12	√	√	√	√
4′*	$C_{20}H_{24}N_2$	293.2013	-0.22	√	×	×	×
5*	$C_{31}H_{38}N_2O_3$	–	–	×	×	√	√
6*	$C_{31}H_{38}N_2O_2$	471.2998	0.26	√	×	√	√
7*	$C_{11}H_{17}NO$	180.1378	2.28	√	×	×	×
8*	$C_{20}H_{21}NO$	292.1693	0.67	√	×	×	×
9	–	183.2	–	√	×	×	×
10	–	208.2	–	√	×	×	×
11	–	167.2	–	√	×	×	×
12	$C_{20}H_{18}N_2O$	303.1490	0.50	√	×	×	×
13	$C_{31}H_{32}N_2O$	449.2586	0.45	√	×	×	×
14	$C_{18}H_{24}N_2O$	285.1962	0.08	√	×	×	×
15	–	167.2	–	√	×	×	×
16*	$C_{31}H_{38}N_2O_2$	471.3004	0.77	√	×	√	×

tial peak area) and four oxidation products were recorded with peak areas higher than 1% of the total peak area (**3**, **4**, **7** and **9**). The elemental formula for the oxidation product **4** (Scheme 1) (RT = 2.8 min) was determined by accurate mass and isotope ratio measurements. This product is the result of the substrate secondary amine dealkylation and this type of oxidative reaction is commonly observed for electrochemically generated metabolites and oxidative degradation products. [13,16]. According to the *N*-dealkylation mechanism, the reaction leading to formation of compound **4** (primary amine) should lead to the formation of the corresponding aldehyde. Electrochemically, oxidation occurs by the removal of an electron from the secondary amine, followed by deprotonation of an adjacent carbon giving neutral radicals (species **II** and **III**). (Scheme 2) Removal of a second electron gives iminium ion species (**IV** and **V**), which hydrolyse to the observed carbonyl and amine compounds. Deprotonation of the secondary carbon leads to iminium ion **V** which hydrolyses to aldehyde **3** and amine **4**, and deprotonation of the tertiary carbon leads to the iminium ion **IV** which hydrolyses to ketone **8** and amine **7**.

Analysis of the chromatogram shows that the corresponding aldehyde (**3**) is also formed (Scheme 1) with a measured retention time of 10.3 min. The elemental formulae was determined based on accurate mass and isotope ratio measurements assisted by software algorithms (Table 6). Both products were previously reported as degradation products generated under accelerated stability conditions [27]. The previous study reported the primary oxidative product of the compound to be the imine **2** (*m/z* 453) which then hydrolyses to the primary amine (**4**) and aldehyde (**3**). In the case of electrochemical oxidation the imine intermediate **2** was not observed likely because products **3** and **4** are formed via a different mechanism as earlier discussed. In contrast to electrochemical oxidation, autoxidation of compound **1** proceeds via neutral species. (Scheme 3) Hydrogen atom abstraction from the secondary aminal carbon of **1**, radical trapping with O_2, and elimination of H_2O_2 yields imine **2**, which hydrolyses to aldehyde **3** and amine **4**. According to the autooxidation mechanism in Scheme 3, ketone **8** and amine **7** are similarly formed by hydrogen atom abstraction from the tertiary aminal carbon of **1**. The detection of a product with *m/z* = 453 in electrochemical infusion experiments (data not shown) may correspond to the iminium ions likely formed during electrochemical dealkylation and for which subsequent hydrolysis leads to the formation of a primary amine and a carbonyl product [16,28]. According to the electrochemical mechanism, the oxidation should originate two different primary amines and their corresponding carbonyl compounds. (Scheme 2) Therefore in addition to primary amine **4** the formation of another amine with *m/z* 180 should also be observed. This was confirmed with the peak eluting at 5.2 min (compound **7**). The formation of compound **7** (primary amine) via the *N*-dealkylation reaction leads to the formation of the corresponding ketone moiety (**8**). This was detected at low levels via an extracted ion chromatogram at *m/z* 292.2 ([M+H]$^+$) with retention time of at 5.5 min. Elemental formulae for compounds **7** and **8** were determined from accurate mass measurements (Table 6) and their proposed molecular structures are shown in Scheme 1. Compound **9** (*m/z* = 183.2) detected by MS with the multimode (ESI/APCI) ionisation source was also formed under these experimental conditions but did not respond by electrospray MS making it impossible to obtain high resolution mass data and propose its molecular formula using available equipment. This oxidation product was not previously published nor observed under forced degradation conditions or predicted by the *in silico* approach reported in this study. We hypothesize that this oxidation product may result from the O-dealkylation of oxidation product **3** with addition of water to the aldehyde moiety.

The oxidation of the sample at 700 mV using acidic electrolyte (pH 4.9) showed little conversion of the substrate. (Fig. 2 B). This result is similar to those previously published on the electrochemical synthesis of two N-dealkylated products of Fesoterodine where it was shown that the electrochemical *N*-dealkylation reaction is not favoured under acidic conditions [13]. This may be related to the fact that at acidic conditions the lone pair of electrons on the amine nitrogen is less readily accessible for abstraction than it is at basic conditions. Using the 700 mV potential with basic (pH 8.8) and near neutral (pH 7.1) electrolytes the substrate conversion was 65% and 55%, respectively. Under these conditions there was an increase in the peak area of compounds **7** and **9** and a slight increase in the peak area of compound **3**. Interestingly no increase was observed for the peak area of compound **4** which may be explained by subsequent degradation in solution after its formation or by attachment of the amino group to the surface of the carbon electrode [29,30]. Under these experimental conditions a new oxidation product **10** (*m/z* 208.2) was observed. A high resolution mass was not obtained for oxidation product **10** and consequently its chemical structure could not be elucidated. Its formation was not reported for accelerated stability studies, it was not generated under the conditions used for the stress testing studies and it was not predicted by the *in silico* methods.

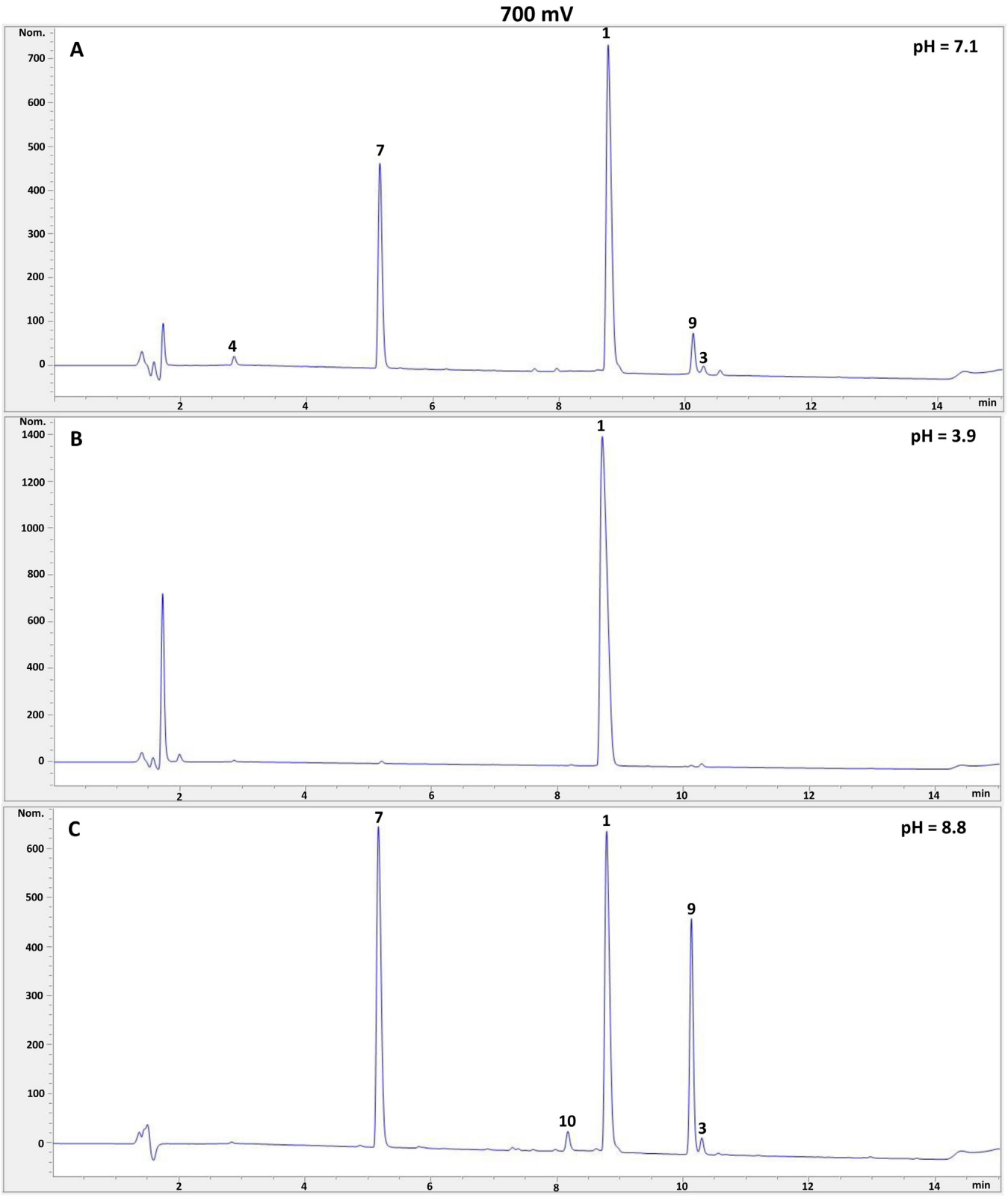


Fig. 2. UV chromatograms at 225 nm for electrochemically oxidized samples at 700 mV using 10 mM ammonium acetate at three different pH values (pH = 3.9, 7.1, and 8.8). Peaks that display an area percent higher than 1% of the total peak area are identified in the chromatograms. Peak **1** corresponds to drug substance under study. For proposed chemical structures see Scheme 1. (No chemical structures were derived for oxidation products **9** and **10**).

Accelerated stability studies previously published describe the generation of five oxidative degradation products (products **2**, **3**, **4**, **5**, and **6**) [27]. The intermediate oxidation products **2** and **5** were not observed for the samples produced for LC–MS by electrochemical oxidation. This may be because the electrochemical and accelerated oxidative degradation pathways are different and do not proceed via the same intermediate species as earlier discussed or may be due to instability of the analytes in the diluents prior to analysis.

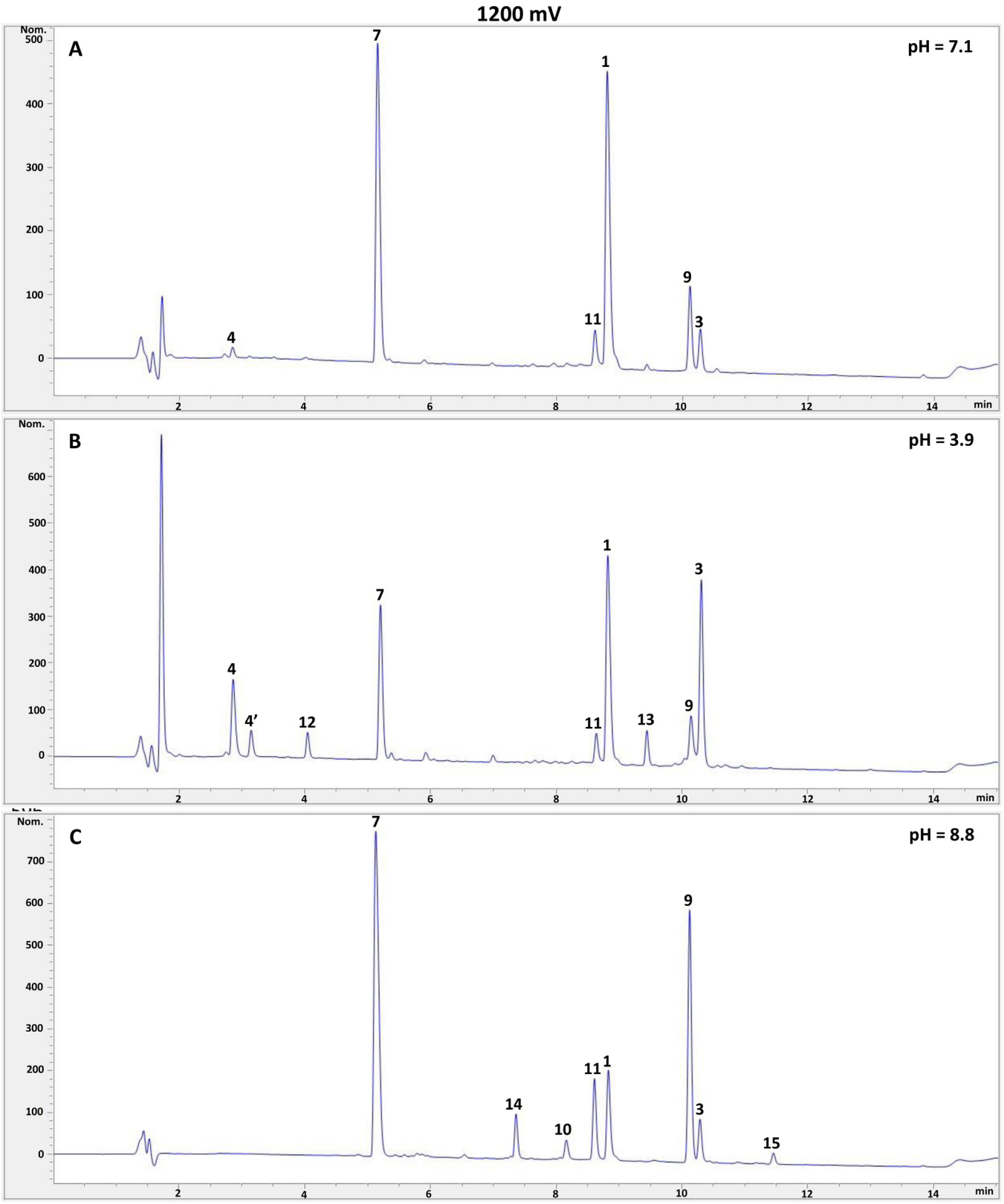
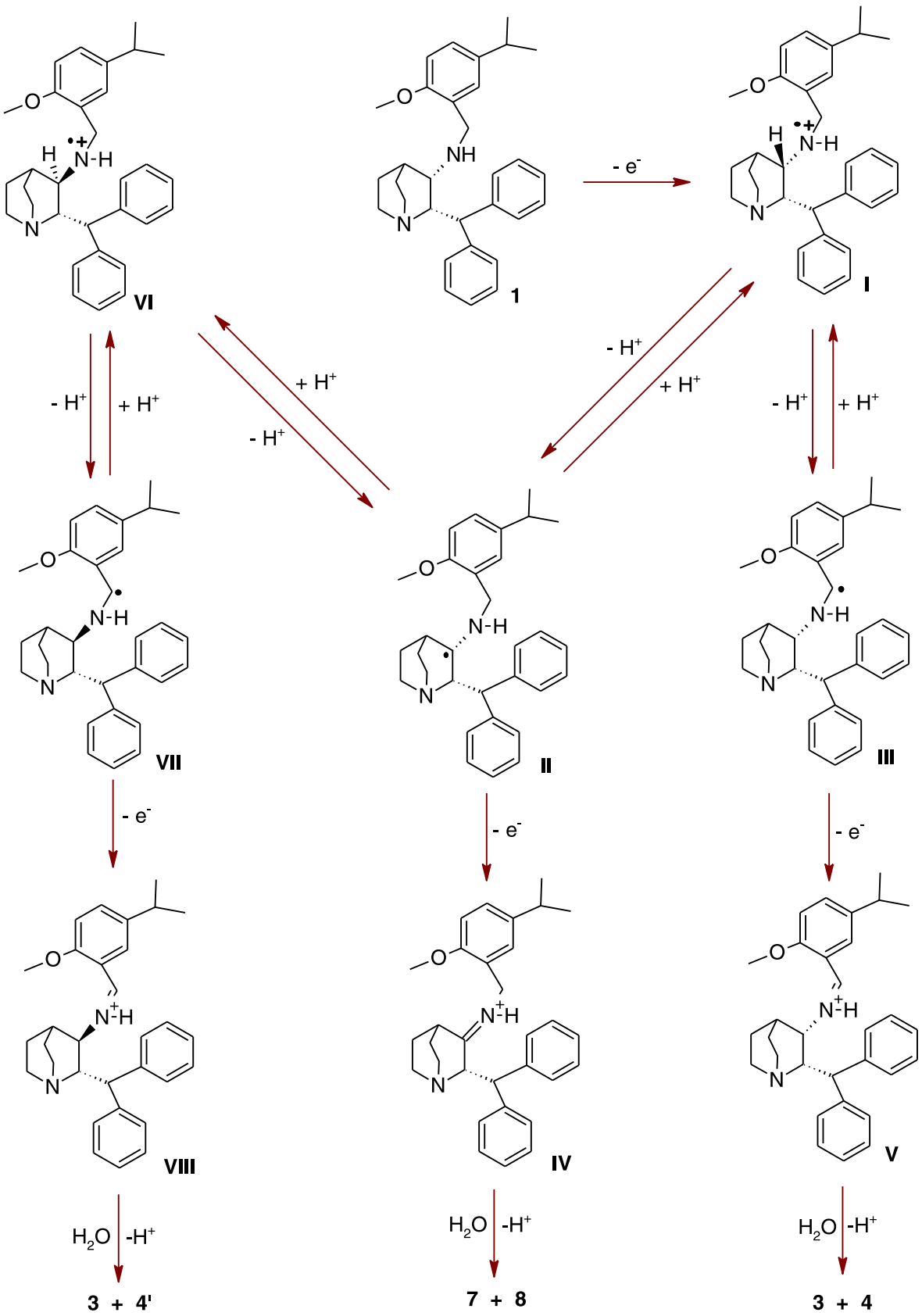
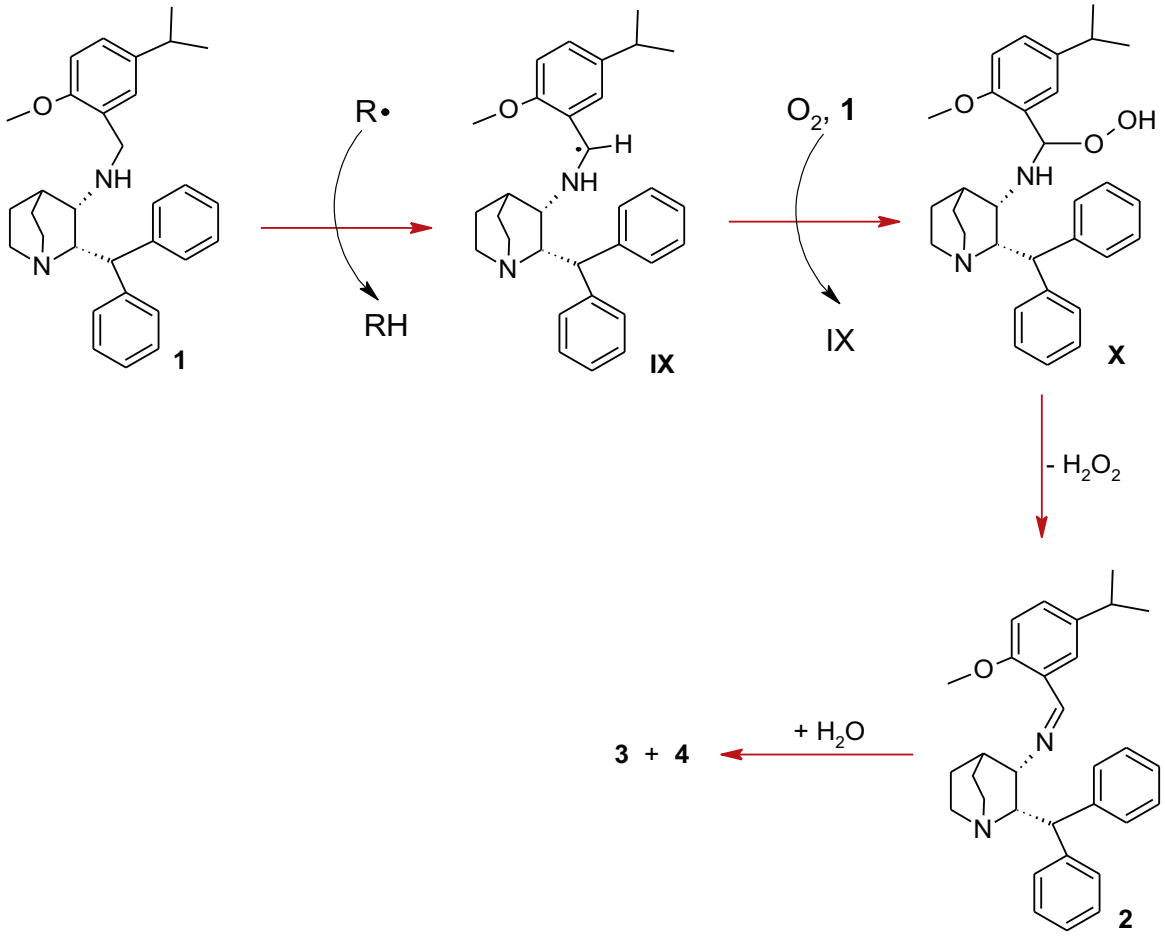


Fig. 3. UV chromatograms at 225 nm for electrochemically oxidized samples at 1200 mV using 10 mM ammonium acetate at three different pH values (pH = 3.9, 7.1, and 8.8). Peaks that display an area percent higher than 1% of the total peak area are identified in the chromatograms. Peak **1** corresponds to the drug substance under study. For proposed chemical structures see Scheme 1. (No chemical structures were obtained for oxidation products **9–15**).



Scheme 2. Electrochemical mechanism for the secondary amine cleavage and corresponding formation of oxidation products **3**, **4**. **4′**, **7** and **8**.



Scheme 3. Autoxidation mechanism of model compound **1,** showing formation of oxidation products **3** and **4** via hydrolysis of the imine product **2**.

An extracted ion chromatogram at *m*/*z* 471 revealed the presence in solution of four mono-oxygenated species. In order to assign a chemical structure to each of the observed species LC–ESI–MS/MS experiments were performed in combination with elemental compositions derived from accurate mass measurements. For two of those four species it was not possible to elucidate their chemicals structure because signal intensities were too low for MS/MS experiments. For the other two species the ESI–MS/MS spectrum of $[M+H]^+$ ion (*m*/*z* 471) was obtained (Figs. 4 and 5) showing structure indicative fragment ions which are explained in the fragmentation scheme below each spectrum. The analysis of the ESI–MS/MS spectrum in Figure 4 shows the presence of a fragment ion at *m*/*z* 179 which suggests that the oxygenation occurs either at the methoxybenzene ring, at the isopropyl moiety or at the benzylic position. However, the fact that ions at *m*/*z* 161 and *m*/*z* 121 are found suggests that the oxygenation of the drug substance likely takes place at the isopropyl moiety. These results show that alcohol **6** which is reported as being one of the oxidation products found under accelerated stability studies can also be generated by electrochemistry. However, it is unlikely that product **6** forms through decomposition of peroxide intermediate **5** since this was not detected in samples oxidized by electrochemistry. In Scheme 4 is depicted the benzylic oxidation of **1** through electrochemical and autoxidation mechanisms. Electrochemically, removal of an electron from the methoxyphenyl ring gives a cation radical (**XI**) which can be deprotonated at the benzylic position giving benzyl radical **XII**. Removal of a second electron gives the benzyl cation which can add water giving alcohol **6**. In contrast, under accelerated and chemical degradation conditions (discussed below), autoxidation is initiated by trace radical sources or radical initiators, producing radical **XII**. Benzyl radical **XII** is trapped by oxygen giving hydroperoxide **5**, which is observed under accelerated degradation conditions. Electrochemical oxidation under air could also initiate autoxidation by producing radical **XII**. Hydroperoxide **5** is not observed by electrochemical oxidation even though there is air exposure, suggesting that the electrochemical process is faster than the electrochemically initiated autoxidation, or that **5** converts to alcohol **6** under these conditions, e.g., via solvolysis by water.

The analysis of the ESI–MS/MS spectrum in Fig. 5 (oxidation product **16**) shows a fragment ion at *m*/*z* 163 which indicates that the oxygenation of the drug substance molecule does not occur at the methoxybenzene ring nor at the $NHCH_2$ carbon; since if that was the case we would likely observe a fragment ion at *m*/*z* 179 as previously observed for compound **6**. The observation of a fragment ion at *m*/*z* 453 (loss of H_2O) indicates that oxygenation likely occurs at the benzylic position and not at the one of the aromatic rings. The proposed chemical structure of mono-oxygenated species **16** is shown in Scheme 1. The species was also predicted by the *in silico* approach.

The UV chromatogram for the sample oxidized at 1200 mV and under neutral conditions is shown in Fig. 3 **A**. 76% of the substrate was converted and the products that were generated with 700 mV cell potential were also generated with a small increase in the peak areas of products **3**, **7**, and **9** and a small decrease in the level of compound **4**. This observation may be explained by subsequent degradation of this product in solution whilst awaiting analysis or by the binding of the amino group to the surface of the carbon elec-

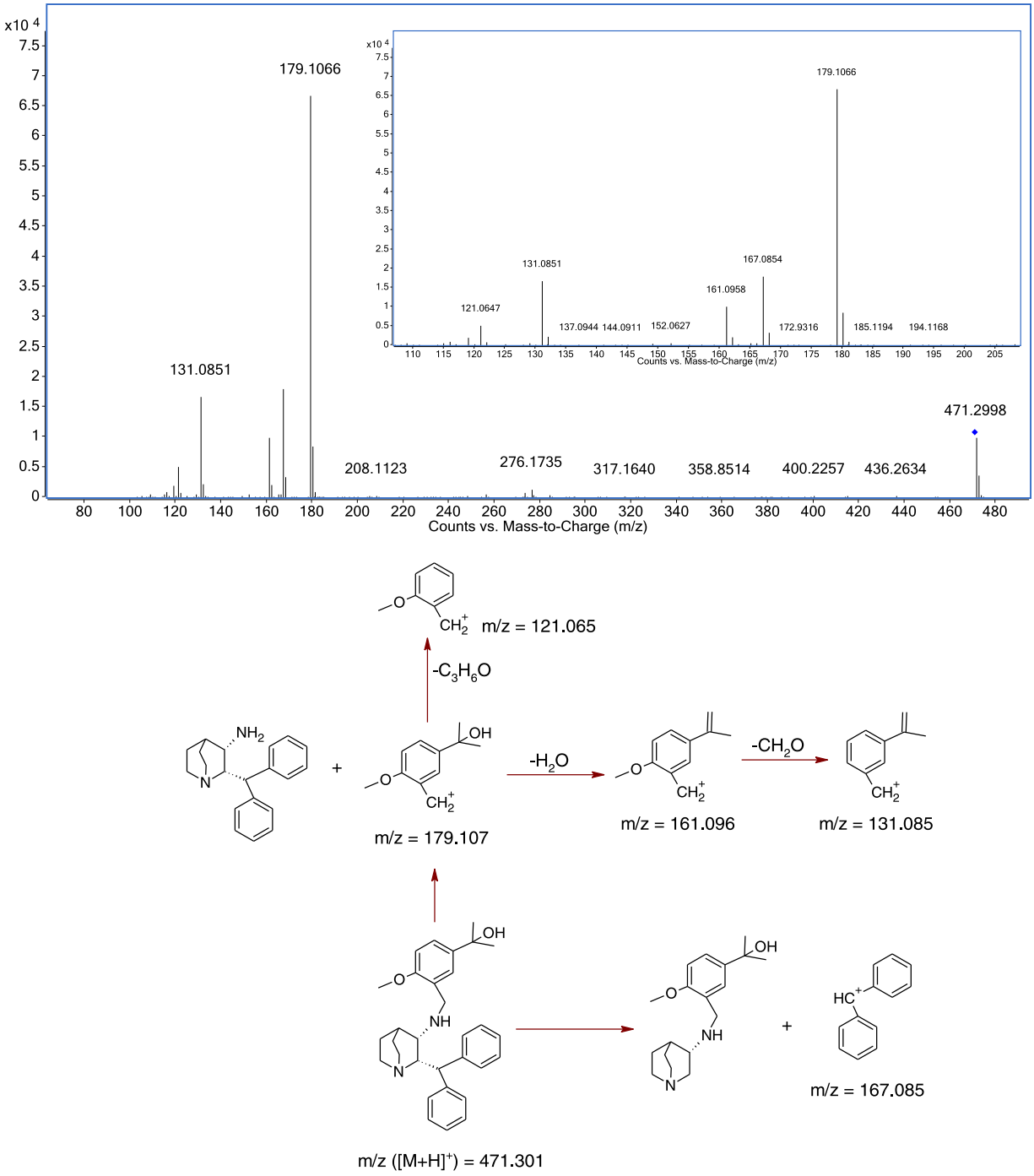


Fig. 4. ESI–MS/MS spectrum of $[M+H]^+$ ion (m/z 471) of the mono-oxygenated species identified as oxidation product 6 generated at 700 mV (pH = 7.1) at 40.0 eV (top) and its proposed fragmentation pathway (bottom).

trode. Another electrochemical oxidation product (**11**, m/z 167.2) was also generated.

With the higher (1.2 V) voltage and basic pH (8.8) 90% of the substrate was oxidised (Fig. 3C) with the formation of two new products; **14** (m/z 285.5) and **15** (m/z 167.0). For compound **14** a high resolution mass was obtained whereas for compound **15** only low resolution mass was obtained. With acidic pH (4.4) 79% of the substrate was oxidized (similar to the 76% conversion obtained with the same voltage at near-neutral pH). The oxidation products profile can be seen in Fig. 3 B. This oxidation profile was very different to the one observed for the same pH lower voltage where no substrate oxidation was observed. The analysis of the chromatogram in Fig. 3 B shows an increase in the amount of products **3** and **4** formed and a decrease of products **7** and **9** when compared to the substrate oxidized at the same voltage but at neutral pH. The amount of compound **11** was unchanged. Use of the acidic electrolyte (pH 4.4) increased the relative amount of products **4′**, **12** and **13**; which were observed with near-neutral electrolyte (pH 7.1) at levels less than 1% total peak area. Oxidation product **4′** was isobaric with compound **4** (m/z 293.2) both species providing similar MS/MS fragmentation patterns. Product **4′** was tentatively assigned as the diastereoisomer of product **4** (Scheme 1)

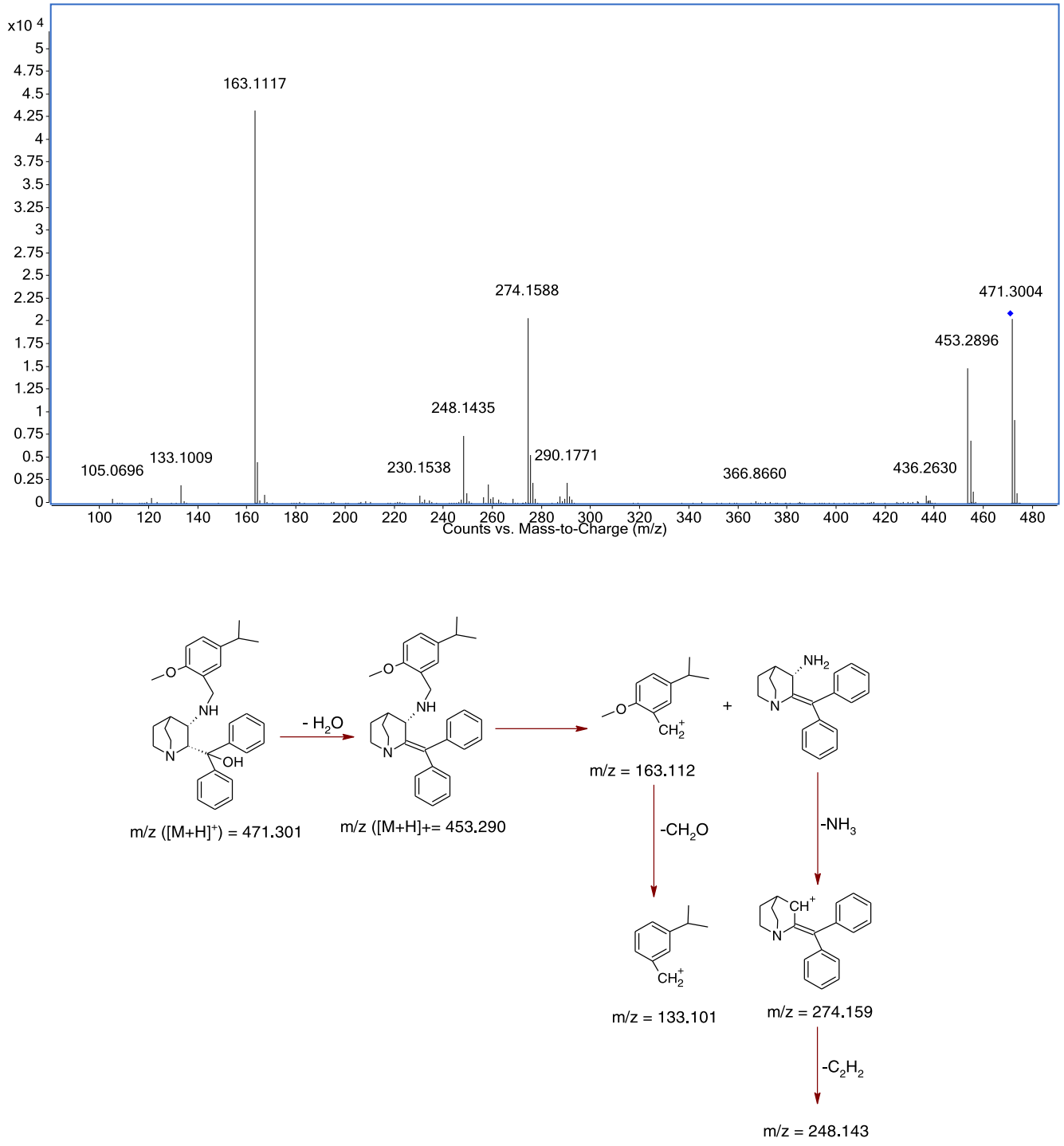
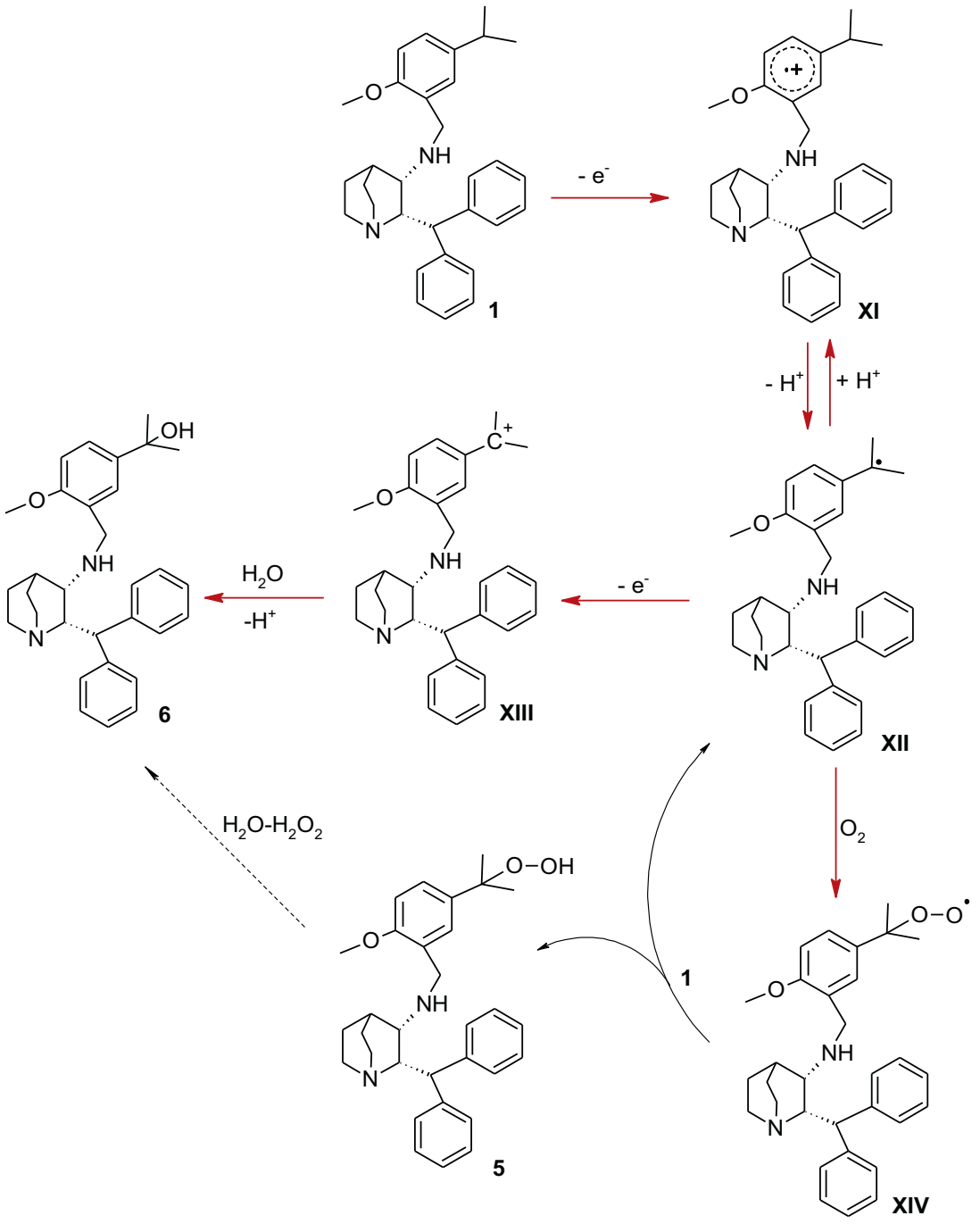


Fig. 5. ESI–MS/MS spectrum of [M + H] ion (*m/z* 471) of the mono-oxygenated species identified as oxidation product 16 generated at 700 mV (pH = 7.1) at 20.0 eV (top) and its proposed fragmentation pathway (bottom).

since during electrochemical *N*-dealkylation if the deprotonation step of the electrochemically produced cation radical is reversible radical **II** could be protonated giving the isomeric cation radical **VI**, which would then yield aldehyde **3** and amine isomer **4′**. Oxidation product **4′**, is not generated by autoxidation which suggests that oxidation products **3**, **4**, **7** and **8** observed in the electrochemically-generated samples are not formed via the autoxidation mechanism (Scheme 2) [31].

The elemental formula for product **13** was determined from accurate mass and isotope ratio measurements (Table 6). This compound can be the result of a triple dehydrogenation of the substrate molecule. However, it was not possible to determine where the dehydrogenation takes place. It was possible to derive an elemental formula for oxidation product **12** as shown in Table 6. This product seems to be related to oxidation product **13** after N-dealkylation and further oxygenation. However its chemical structure was not assigned. The results show that optimization of pH and the applied potential is important in order to selectively produce target oxidation product(s).



Scheme 4. Formation of oxidative product **6** via electrochemical and autoxidation pathways.

3.2. Oxidative forced degradation studies

For stress testing conditions the drug substance under study was exposed either to 0.3% H_2O_2 or 5 mM AIBN solution as described previously in the experimental section. After 3 days of exposure to the peroxide challenge condition no oxidation products were observed. After three days of exposure to the AIBN challenge condition two oxidation products were observed corresponding to a 3.9% total degradation. (Fig. 6B) An oxidation product with retention time of 3.70 min (m/z $[M+H]^+$ = 293.2, compound **4,** Scheme 1) and another oxidation product with a retention time of 10.36 min (m/z $([M+H]^+$ = 179.2, compound **3**). The two products generated under these conditions were observed in accelerated stability studies [27] and were also electrochemical degradation products.

3.3. In silico predictions

A Zeneth prediction analysis was conducted for the compound under study in order to understand how *in silico* methods can help in anticipating and/or identifying potential oxidative degradation products. Zeneth is an expert, knowledge-based software that gives forced degradation predictions quickly, helps to understand the forced degradation pathways of organic compounds and gives a prediction of the degradation products forming via various known mechanisms [25,26]. The chemical structures of the predicted oxidative products for the substrate under study are provided in Table 7. The predicted oxidation products can be used as an initial guide to potential degradation products that may form at a given condition and it is not anticipated that Zeneth will predict

Table 7
Chemical structures of oxidative degradation products of the compound under study predicted by Zeneth. The structures presented are the result of autoxidation (O_2/Initiator) and peroxide and are predicted as likely unless otherwise stated.

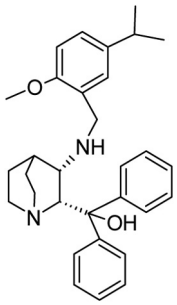
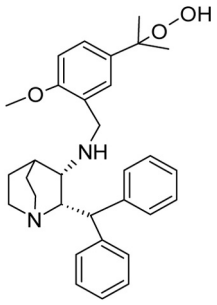
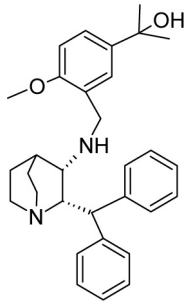
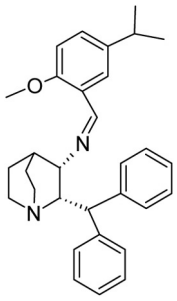
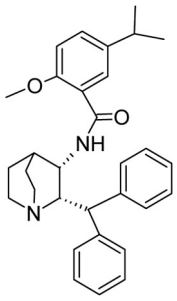
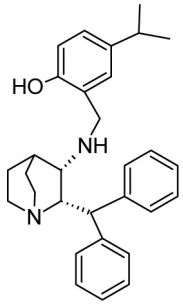
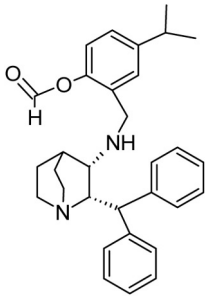
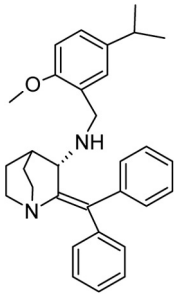
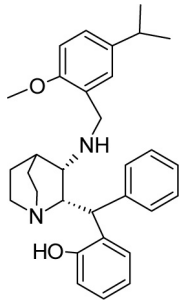
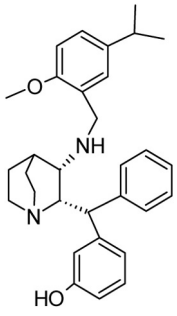
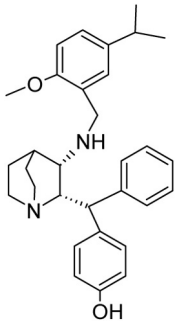
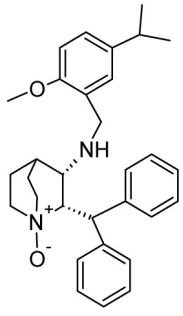
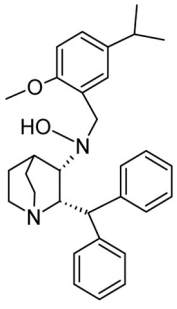
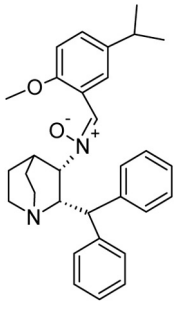
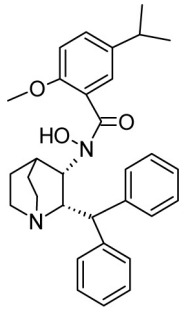
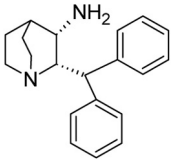
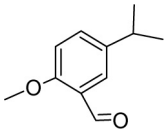
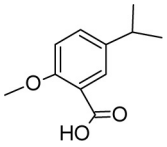
Entry		Entry		Entry	
I		II		III	
IV		V		VI[a]	
VII[a]		VIII[a,b]		IX[a,e]	
X[a,e]		XI[a,e]		XII[a,c]	
XIII[c]		XIV[c]		XV[c]	

Table 7(*Continued*).

Entry		Entry		Entry	
XVI[d,f]		XVII[d]		XVIII[f]	

[a] Entries VI–XI were predicted as equivocal and entry XII was predicted as very likely.
[b] Condition: O_2/Initiator (autoxidation) and metals.
[c] Condition: Peroxide/Base.
[d] Condition: Hydrolysis of IV, Acid/Base
[e] Condition: Fenton Chemistry.
[f] Hydrolysis of V, Acid/Base.

all degradations products that are actually observed during stability studies. In fact, additional primary and secondary degradation pathways may occur.

The computational method used predicted eighteen oxidative degradation products for the substrate under study including the three products generated by electrochemistry **3**, **4**, and **6** and the peroxide and imine intermediates reported in the accelerated stability studies (**2** and **5**). Using the stress testing approaches no reaction was observed using hydrogen peroxide and only products **3** and **4** were generated by the reaction with AIBN radical initiator after three days of incubation. It was not possible to generate product **6** or the intermediates using these stress test conditions.

The electrochemical approach generated a higher number of oxidized products which were not previously reported as degradation products nor predicted by Zeneth software. This is not surprising as the knowledge base used by Zeneth for the prediction of oxidative degradation does not include electrochemical oxidative pathways. There is potential to build

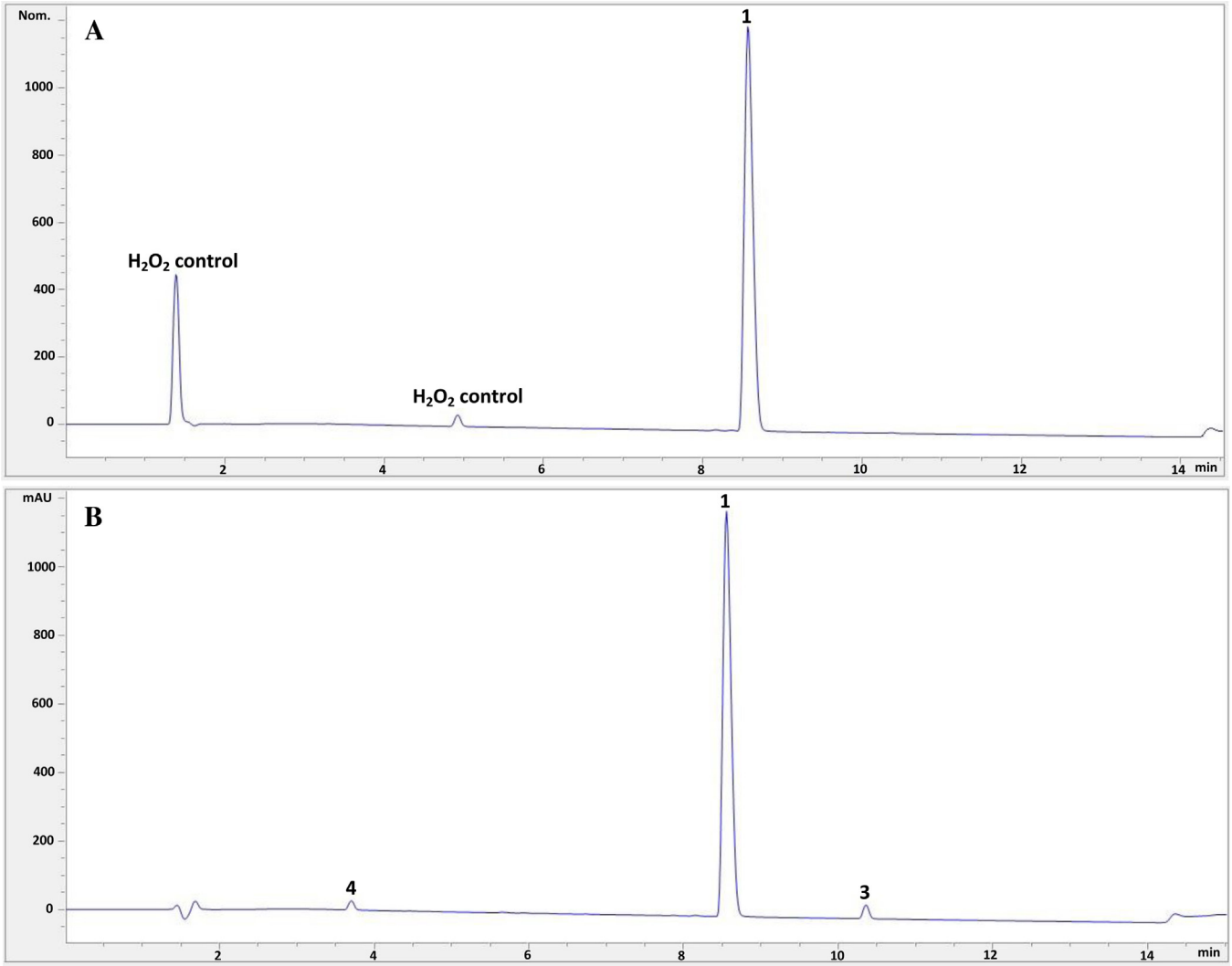


Fig. 6. LC–UV chromatogram at 225 nm of oxidative degradation samples after 3 days for challenge condition (A) 0.3% H_2O_2 and (B) 5 mM AIBN.

knowledge-based predictions into *in silico* databases to improve the accuracy of prediction with respect to electrochemical degradation. These results show that the electrochemical oxidation approach can help produce oxidative products-enriched solutions for use in structure elucidation and methods development studies in an accelerated way compared to traditional chemical approaches.

4. Conclusions

The model pharmaceutical compound was electrochemically oxidized using an electrochemical flow-through cell and the oxidation products generated were compared to those obtained from chemical oxidation (stress testing studies) and from accelerated stability studies and those predicted by a computational approach (*in silico* method). The five reported oxidation products previously found in accelerated stability studies [27] were among the eighteen oxidation products predicted by the *in silico* approach. From those five products electrochemistry was able to generate the three reported stable degradation products. The previously proposed intermediate species were not observed either because of a different oxidation mechanism or chemical instability under the applied experimental conditions. From the three stable products generated by electrochemistry only two were generated using the chemical approach (stress testing studies) using the radical initiator AIBN. Electrochemical stress testing was much faster and provided more comprehensive synthesis of the targeted degradants than the chemical stress test methods evaluated for this compound. These results demonstrate the potential of electrochemistry as an oxidative stress testing method alongside other chemical and physical approaches. Electrochemistry has the advantage of not requiring the use and storage / disposal of strong oxidizing agents and it enables study of reactions in real-time as well as the ability to select optimum experimental conditions for generation of the desired oxidation products profile(s). Additionally, it also enables the synthesis of oxidative products that are difficult to obtain by traditional synthetic methods and that can be further used as analytical markers or in toxicity or bioactivity tests. The electrochemical oxidation technique is able to generate a high number of structurally-related oxidation products and whilst in this case some were not recorded in traditional stability studies it demonstrates a way of producing mixtures of similar compounds which could find utility in many pharmaceutical R + D studies.

References

[1] S. Yoshioka, J.V. Stella, Stability of Drugs and Dosage Forms, Springer, 2000.
[2] S. Bajaj, D. Singla, N. Sakhuja, Stability testing of pharmaceutical products, J. Appl. Pharm. Sci. 02 (2012) 129–138.
[3] M. Blessy, R.D. Patel, P.N. Prajapati, Y.K. Agrawal, Development of forced degradation and stability indicating studies of drugs—a review, J. Pharm. Anal. 4 (2014) 159–165.
[4] S.W. Baertschi, Pharmaceutical Stress Testing: Predicting Drug Degradation, Taylor & Francis, 2005.
[5] ICH Q1A(R2) Stability Testing of New Drugs Substances and Products, ICH Steering Committee, (2003).
[6] S.W. Baertschi, P.J. Jansen, K.M. Alsante, Stress testing: a predictive tool, in: S.W. Baertschi (Ed.), Pharmaceutical Stress Testing: Predicting Drug Degradation, 2005, pp. 10–48.
[7] S.W. Baertschi, K.M. Alsante, D. Santafianos, Stress testing: the chemistry of drug degradation, in: S.W. Baertschi (Ed.), Pharmaceutical Stress Testing: Predicting Drug Degradation, 2005, pp. 49–141.
[8] S.W. Hovorka, C. Schöneich, Oxidative degradation of pharmaceuticals: theory, mechanisms and inhibition, J. Pharm. Sci. 90 (2001) 253–259.
[9] P. Harmon, G. Boccardi, Oxidative susceptibility testing, in: S.W. Baertschi (Ed.), Pharmaceutical Stress Testing: Predicting Drug Degradation, 2005, pp. 168–191.
[10] K.M. Alsante, A. Ando, R. Brown, J. Ensing, T.D. Hatajik, W. Kong, Y. Tsuda, The role of degradant profiling in active pharmaceutical ingredients and drug products, Adv. Drug Deliv. Rev. 59 (2007) 29–37.
[11] F. Lombardo, G. Campos, How Do We Study Oxidative Chemical Stability in Discovery? Some Ideas Trials and Outcomes, AAPS Press, Virginia, 2004.
[12] P. Gamache, R. McCarthy, J. Waraska, I. Acworth, Pharmaceutical oxidative stability profiling with high-throughput voltammetry, Am. Lab. 35 (2003) 21–25.
[13] S. Torres, R. Brown, R. Szucs, J.M. Hawkins, G. Scrivens, A. Pettman, D. Kraus, M.R. Taylor, Rapid synthesis of pharmaceutical oxidation products using electrochemistry: a systematic study of N-Dealkylation reactions of fesoterodine using a commercially available synthesis cell, Org. Process Res. Dev. (2014), http://dx.doi.org/10.1021/op500312e.
[14] A. Baumann, U. Karst, Online electrochemistry/mass spectrometry in drug metabolism studies: principles and applications, Expert Opin. Drug Metab. Toxicol. 6 (2010) 715–731.
[15] S. Jahn, U. Karst, Electrochemistry coupled to (liquid chromatography/) mass spectrometry–current state and future perspectives, J. Chromatogr. A 1259 (2012) 16–49.
[16] U. Jurva, H.V. Wikstrom, L. Weidolf, A.P. Bruins, Comparison between electrochemistry/mass spectrometry and cytochrome P450 catalyzed oxidation reactions, Rapid Commun. Mass Spectrom. 17 (2003) 800–810.
[17] T. Johansson, L. Weidolf, U. Jurva, Mimicry of phase I drug metabolism–novel methods for metabolite characterization and synthesis, Rapid Commun. Mass Spectrom. 21 (2007) 2323–2331.
[18] S.M. van Leeuwen, H. Hayen, U. Karst, Liquid chromatography-electrochemistry-mass spectrometry of polycyclic aromatic hydrocarbons, Anal. Bioanal. Chem. 378 (2004) 917–925.
[19] A. Baumann, W. Lohmann, B. Schubert, H. Oberacher, U. Kart, Metabolic studies of tetrazepam based on electrochemical simulation in comparison to in vivo and in vitro methods, J. Chromatogr. A 1216 (2009) 3192–3198.
[20] W. Lohmann, U. Karst, Generation and identification of reactive metabolites by electrochemistry and immobilized enzymes coupled on-line to liquid chromatography/mass spectrometry, Anal. Chem. 79 (2007) 6831–6839.
[21] S.M. van Leeuwen, B. Blankert, J.M. Kauffmann, U. Karst, Prediction of clozapine metabolism by on-line electrochemistry/liquid chromatography/mass spectrometry, Anal. Bioanal. Chem. 382 (2005) 742–750.
[22] B. Blankert, H. Hayen, S.M. van Leeuwen, U. Karst, E. Bodoki, S. Lotrean, R. Sandulescu, N.M. Diez, O. Dominguez, J. Arcos, J.M. Kauffmann, Electrochemical, chemical and enzymatic oxidations of phenothiazines, Electroanalysis 17 (2005) 1501–1510.
[23] S. Jahn, A. Baumann, J. Roscher, K. Hense, R. Zazzeroni, U. Kart, Investigation of the biotransformation pathway of verapamil using electrochemistry/liquid chromatography/mass spectrometry - a comparative study with liver cell microsomesInvestigation of the biotransformation pathway of verapamil using electrochemistry/liquid chromatography/mass spectrometry - a comparative study with liver cell microsomes, J. Chromatogr. A 1218 (2011) 9210–9220.
[24] W. Lohmann, U. Karst, Electrochemistry meets enzymes: instrumental on-line simulation of oxidative and conjugative metabolism reactions of toremifene, Anal. Bioanal. Chem. 394 (2009) 1341–1348.
[25] M.H. Kleinman, S.W. Baertschi, K.M. Alsante, D.L. Reid, M.D. Mowery, R. Shimanovich, C. Foti, W.K. Smith, D.W. Reynolds, M. Nefliu, M.A. Ott, In silico prediction of pharmaceutical degradation pathways: a benchmarking study, Mol. Pharm. 11 (2014) 4179–4188.
[26] A.D. Parenty, W.G. Button, M.A. Ott, An expert system to predict the forced degradation of organic molecules, Mol. Pharm. 10 (2013) 2962–2974.
[27] K.C. Waterman, M.C. Roy, Use of Oxygen Scavengers to Stabilize Solid Pharma Dosage Forms: A Case Study, Pharmaceutical Development and Technology 7 (2002) 227–234.
[28] M. Masui, H. Sayo, Anodic oxidation of amines. Part II. Electrochemical dealkylation of aliphatic tertiary amines, J. Chem. Soc. (B) (1971) 1593–1596.
[29] I. Gallardo, J. Pinson, N. Vilà, Spontaneous attachment of amines to carbon and metallic surfaces, J. Phys. Chem. B 110 (2006) 19521–19529.
[30] M.M. Alain Adenier, I. Chechimi, J. Gallardo, N. Pinson V, ila, Electrochemical oxidation of aliphatic amines and their attachment to carbon and metal surfaces, Langmuir 20 (2004) 8243–8253.
[31] M. Masui, H. Sayo, Y. Tsuda, Anodic oxidation of amines. Part I. Cyclic voltammetry of aliphatic amines at a stationary glassy-carbon Electrode, J. Chem. Soc. (B) (1968) 973–976.

Ogenstad Declaration Exhibit G

Simulation for Designing Clinical Trials

A Pharmacokinetic-Pharmacodynamic Modeling Perspective

Hui C. Kimko
Johnson & Johnson Pharmaceutical Research & Development
Raritan, New Jersey, U.S.A.

Stephen B. Duffull
University of Queensland
Brisbane, Queensland, Australia

MARCEL DEKKER, INC. NEW YORK • BASEL

Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the Library of Congress.

ISBN: 0-8247-0862-8

This book is printed on acid-free paper.

Headquarters
Marcel Dekker, Inc.
270 Madison Avenue, New York, NY 10016
tel: 212-696-9000; fax: 212-685-4540

Eastern Hemisphere Distribution
Marcel Dekker AG
Hutgasse 4, Postfach 812, CH-4001 Basel, Switzerland
tel: 41-61-260-6300; fax: 41-61-260-6333

World Wide Web
http://www.dekker.com

The publisher offers discounts on this book when ordered in bulk quantities. For more information, write to Special Sales/Professional Marketing at the headquarters address above.

Current printing (last digit):
10 9 8 7 6 5 4 3 2 1

PRINTED IN THE UNITED STATES OF AMERICA

7

Analysis of Simulated Clinical Trials

Ene I. Ette, Christopher J. Godfrey, and Stephan Ogenstad
Vertex Pharmaceuticals, Inc. Cambridge, Massachusetts, U.S.A.

Paul J. Williams
Trials by Design, LLC, La Jolla and Stockton, and University of the Pacific, Stockton, California, U.S.A.

7.1 INTRODUCTION

Maximization of knowledge gained during the drug development process with the intent of increasing the probability of "success" in a clinical trial has received considerable attention recently (1). Clinical trial simulation has been recognized as being potentially useful for streamlining clinical drug development and evaluation with a view to maximizing information content of clinical trials (2–5).

Simulation of a clinical trial based on pharmacokinetic, pharmacodynamic (PK/PD), and clinical outcome link models is a technique that is being newly developed. When there is a considerable amount of information about the drug, simulation provides the means of synthesizing the information into a coherent package that indicates the drug developer (sponsor) has good control over the pharmacology and, eventually, the therapeutics of the drug.

The objective of the clinical trial simulation paradigm in drug development is to increase the efficiency of drug development, i.e., minimizing the cost and

time of the development of a drug while maximizing the informativeness of data generated from a trial or trials. Simulation of clinical trials offers the means of generating complex data sets, which may include the influence of prognostic factors, sample size, and dropouts, for testing new analysis methods (6, 7). It offers the possibility of improving the chances of an efficacy trial succeeding by allowing the user to ask questions which would involve perturbing different aspects of the input model (5). For instance, what would be the impact of a 20% noncompliance rate on the study outcome? What would be the minimal size of a study that would be commensurate with a reliable detection of the treatment effect? Thus, clinical trial simulation as an abstraction of the clinical trial process is used to investigate assumptions built into a proposed study, study designs, perform sensitivity, and power analysis in order to maximize the information content that can be gained from a study.

Hale et al. (8) used simulation to choose experimental design, sample size, and study power for a pivotal clinical trial of mycophenolate mofetil, a prodrug of mycophenolic acid—an immunosuppressant, in combination with cyclosporin and corticosteroids in transplantation. This was a randomized concentration-controlled trial (RCCT) in which subjects were randomized to one of three target drug exposure [area under the plasma concentration-time curve (AUC)] levels instead of three dose levels. An RCCT was chosen because modeling and simulations based on phase II data had indicated a study using AUC rather than dose would have much greater power to detect treatment effect.

Bias and precision of the estimates of quantitative descriptors that reflect treatment effect size, time to peak effect, drug disposition, and effect with associated variability can be evaluated using simulation. In addition, simulation has been used as a tool for investigating the performance of various sampling designs and design factors employed in population pharmacokinetics (PPK) and population PK/PD studies (9–16). It has also been used for dose selection for clinical trial structure and design (17, 18).

Clinical trial simulation can be likened to meta-analyses of clinical trials. Meta-analysis has been defined as statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings (19). Without meta-analysis useful data can be left fallow or are at least not utilized to their maximum extent (20). The results of meta-analysis of clinical trials can point to specific areas that need to be addressed, either because there are few available data or because the available data suggest a particular hypothesis that requires additional attention (e.g., an age group of patients that may be at an increased risk, or an unexpected exposure from a low dose may appear harmful). These issues can be addressed by simulating the clinical trials, but unlike meta-analysis that is retrospective, clinical trial simulation is prospective. However, a simulated trial is similar to meta-analysis in that it consists of many replications of the trial. It differs, however, in that the replications of the trial

by simulation are usually of greater homogeneity than the trials included in meta-analyses. Hence, statistical analysis methods used in analyses of simulated clinical trials usually incorporate a random effects component which is generally unnecessary for the meta-analysis of actual clinical trials. Nevertheless, they are analogous systems.

One of the challenging tasks for a pharmacometrician/statistician is to convey findings from the analyses of a simulated trial to clinicians and other members of the drug development team. Communication of the results can be effected through words, tables, or graphics. The use of high-quality graphics can effectively enhance the communication of the outcome of a simulation project to a drug development team. Graphics, in particular, are essential for conveying relations and trends in an informal and simplified visual form (21).

The remainder of the chapter deals with analysis of simulated data from single and replicated trials with some examples from PPK and efficacy trials, multiplicity in simulated efficacy trials, and the communication of the results of a clinical trial simulation (CTS) project.

7.2 ANALYSIS OF SIMULATED DATA

The procedure for analyzing a simulated trial should be specified in a simulation plan. A simulated trial should be analyzed using the same data analysis method as the actual trial. Insights into a single replication of the trial or integrative and comparative insights can be gained by analyzing a single-trial replication or the group of trial replications performed. Thus, we consider these two levels of analysis next.

7.2.1 Single-Trial Replication Analysis

The experimental unit of a replication of a single clinical trial is each individual in the trial. Thus, analyzing one replication of a simulated trial allows the responses across subjects within that replication to be summarized. In this case, the analyst would be interested in individual subject outcome measures, for example, the duration of drug effect, before proceeding to perform the complete simulation experiment.

7.2.2 Simulated Experiment Analysis

The replicated in silico clinical trial is akin to the meta-analysis of clinical trials as discussed previously. In this case, a response across trials is defined. A summary of the analysis of the outcome measures obtained from each clinical trial replication yields the outcome of each simulated experiment. For a specific clinical trial design the power of the study is determined by the number of trials that reject the null hypothesis. This is often the case in a confirmatory trial. It is good

practice in performing CTS to assess study power that the type I error rate (α) be estimated to verify the size of the test (e.g., $\alpha = 0.05$). An error rate is a probabilistic measure of erroneous inference in a given set of hypotheses from which some common conclusions are drawn. Bias and precision of parameter estimates such as drug clearance in a population pharmacokinetic study or maximum drug effect in a PK/PD study are typically computed.

7.2.2.1 A Population Pharmacokinetic Study Example: Determination of Power

Kowalski and Hutmacher (22, chapter by Hutmacher and Kowalski) performed clinical trial simulations to assess the power to detect subpopulation differences in apparent drug clearance (CL/F) and examine sample size requirements for a population pharmacokinetic substudy (3) of a phase III clinical trial. The simulations were based on a population PK model developed from a phase I healthy volunteer study. Taking into account the practical constraints of the proposed study, a sparse sampling design was developed for the study. Due to the narrow range of the sampling times, it was expected that a sparse sampling design would not support a two-compartment model that was used to describe the phase I data. Thus, a minimal model, a one-compartment model, was used to fit the data from the simulated experiment. The key parameter of interest in the simulated study was CL/F, and a 40% reduction was considered to be of clinical significance. That is, this degree of reduction in CL/F would result in a need for dosage adjustment.

Three hundred hypothetical clinical trials were simulated to determine the sample size necessary to detect 40% reduction in CL/F in a subpopulation of proportion $p = 0.05$ or $p = 0.10$ with at least 90% power. Sample sizes of 150 and 225 were investigated. The power of the study was estimated as the percentage of trials out of 300 in which statistically significant ($\alpha = 0.05$) difference in CL/F was observed using the likelihood ratio test.

To obtain the empirical estimates of α, Kowalski and Hutmacher (22) simulated 300 clinical trials for each combination of sample size and p, where the proportional reduction in CL/F (ϕ) was constrained to zero. Covariate and base models were fitted to each of the trials and the likelihood χ^2 ratio tests were performed at the 5% level of significance. The percentage of trials where a statistically significant difference in CL/F was observed provided an empirical estimate of α (i.e., H_0: $= 0$ is rejected when H_0 is true). The data were analyzed with the NONMEM (23) software. The results suggested that approximately a 9 point change in the objective function should be used to assess statistical significance at the 5% level rather than the commonly used χ^2 critical value of 3.84 for 1 degree of freedom. Also, the proposed sampling design with the minimal model yielded accurate and precise estimates of apparent drug clearance (CL/F) and apparent volume of distribution at steady state. This study was executed in accor-

dance with the spirit and methods of the consensus document on clinical trial simulation (24).

a. Use of the Likelihood Ratio Test. The work of Kowalski and Hutmacher (22) also provides a poignant example of the risk inherent in the use of the likelihood ratio test (LRT) in the analysis of simulated trials, particularly in the context of mixed effects modeling.

If minus twice the log likelihood associated with the fit of a saturated model A with $p + q$ parameters is designated ℓ_A, and a reduced version of this model (model B) with p parameters has minus twice the log likelihood ℓ_B, the difference in minus twice the log likelihoods $(\ell_A - \ell_B)$ is asymptotically χ^2 distributed with q degrees of freedom. This formulation is widely used to assess the statistical significance level of the parameters associated with the q degrees of freedom.

For the determination of the significance level of fixed effects, the LRT is known to be anticonservative, i.e., the empirical p value will be greater than the nominal p value (25). Generally, as the number of parameters (degrees of freedom) being tested increases, the more liberal the test.

Conversely, Stram and Lee (26) noted that the LRT tends to be asymptotically conservative for the assessment of random effects significance level. In this context, the conservative nature is attributable to the null hypothesis consisting of setting the variance term at a boundary condition, i.e., zero. While the inaccuracy in p value is modest when the number of random effects being tested is small, the conservativeness increases with an increase in the number of random effects being tested.

Wählby et al. (27) explored via simulation a number of factors influencing the disparity between nominal and actual significance level of tests for covariate (fixed) effects in nonlinear mixed effects models using the software, NONMEM. Approximation method [first order (FO) versus first-order conditional estimation (FOCE) with interaction between interindividual variability η and residual variability ε], sampling frequency, and magnitude and nature of residual error were determined to be influential on the bias associated with the p value. An important finding was that the use of the FOCE method with η-ε interaction resulted in reasonably close agreement of actual and nominal significance levels, whereas the application of the LRT after estimation using the FO approximation generally resulted in marked bias in p values.

The implications of the disparity between actual and nominal significance levels of the likelihood ratio test in mixed effects modeling and simulation are clear; however, simple algorithms for p value correction are not readily available. The bias in likelihood ratio test–determined p value for fixed effects could be very influential on trial simulation findings. Ultimately, simulation exercises should provide for determination of empirical p values to avoid faulty conclusions about

power and sample size, for example. This is what Kowalski and Hutmacher did in the example discussed in the previous section.

b. Reliability and Robustness of Parameter Estimates. Determining the reliability of parameter estimates from the pharmacokinetic/pharmacodynamic (input-output) models is necessary as this may affect study outcome. Not only should bias and precision associated with parameter estimation be determined, but also the confidence with which these parameters are estimated should be examined. Confidence interval estimates are a function of bias, standard error of parameter estimates, and the distribution of parameter estimates. Paying attention to these measures of parameter estimation efficiency is critical to a simulation study outcome (11, 12).

Simulation is useful for evaluating the merits of competing study designs (3, 4). Competing study design should be evaluated for power, efficiency, robustness, and informativeness. In evaluating the power of a study with a particular design, the ability to reject a null hypothesis or to estimate a parameter such as drug clearance with the design is examined. Efficiency examines the ratio of effort/cost to expected result. Efficiency is sacrificed but power is improved when more subjects are enrolled, sampled more often (when an efficacy variable is involved) and for longer periods of time.

It is also important to evaluate the quality of the results of a simulated population pharmacokinetic, pharmacokinetic/pharmacodynamic, or confirmatory efficacy study for robustness. Robustness addresses the question, "If my assumptions underlying the study design are wrong, am I still able to meet the objectives of the research project?" Evaluation for robustness may be approached by sensitivity analysis. Evidence of robustness renders the results acceptable and independent of the analyst. Informativeness addresses the question of how much can be learned from the study. Very often informativeness can both be increased in a study simply by collecting extra data on each subject without enrolling additional subjects (10).

The intended objective of a study will dictate which of the criteria are most important for any given design. For studies where it is important to continue to learn about the effects of factors such as renal function, size, gender, or race on drug pharmacokinetics, dynamics, or other effects, one would be most interested in an informative study design. When one is interested in showing a drug effect is superior to a placebo, power becomes a major issue and informativeness may be sacrificed. A study design should be selected that balances the four criteria for evaluation.

7.2.2.2 An Efficacy Study Example: Dosage Optimization

Another well-executed example of CTS is a report on dosage optimization for docetaxel in a proposed oncology trial (17). This application is also consistent

with the spirit and methods of the Simulation in Drug Development: Good Practices (24) consensus document. The investigation was performed to determine whether non–small-cell cancer patients might benefit from dose intensification. Prior PK/PD analysis performed during the development of the drug showed that those non–small-cell cancer patients who had high baseline α1-acid glycoprotein (AAG) levels had shorter time to disease progression and death. Thus, PK/PD models were developed for time to disease progression, death, and dropout and validated with data from 151 non-small-cell lung cancer patients from phase II studies. Separate hazards were estimated for each event type. Different types of hazard models (exponential, Gompertz, and Weibull) were tested, and the Weibull model was found to provide the best explanation of the data. Prognostic factors (dose at first cycle, cumulative dose, and cumulative AUC) were also included in the models for death and disease progression. Posterior predictive check (3, 28) was used to evaluate whether the models provided adequate description of the data.

The simulation process was evaluated by simulating 100 complete trials of previous phase II trials from which the PK/PD models were derived. Kaplan-Meier analyses were performed on each simulated trial. Several statistics, such as median time to progression, 1 year survival, number of deaths, and disease progression, etc. were computed and compared with those obtained from the real trials. The statistics from real and simulated trials were in good agreement. For instance, from the Kaplan-Meier analysis a similar time course of death was observed in the simulations and phase II data. Thus, the simulation process was evaluated as the basis for simulating a phase III trial of docetaxel in non-small-cell lung cancer.

The primary objective of the simulated phase III trial was the comparison of overall survival between two doses (100 and 125 mg/m^2) of docetaxel. A secondary objective was the comparison of the time to progression and the safety of the two dosage regimens. This was a proposed randomized phase III trial in non-small-cell lung cancer patients with high AAG levels (i.e., AAG levels higher than 1.92 g/L). It was thought that dose intensification would result in a clinical benefit to this group of patients. Patients were randomized to one of the two doses, and the drug was administered as a 1 h intravenous infusion every 3 weeks. Two hundred patients were randomized to each treatment arm, and 100 trials were simulated. This sample size was assumed to be adequate and would permit the detection of a survival advantage of 8 weeks in the 125 mg/m^2 treatment arm with an α of 5% and 80% power. A comparison of the overall survival and time to progression for the two treatments was done using the log-rank test at the 5% level of significance.

The difference in the median time to progression between the two treatments (i.e., 9.18 and 9.48 weeks for the 100 and 125 mg/m^2, respectively) was only significant in 11 of 100 trials. A slightly longer median survival time of

5.49 months was observed for patients on the 125 mg/m^2 regimen as opposed to a 5.31 month median survival time for patients on the 100 mg/m^2 regimen. However, the difference in the median survival times was only significant in 6 of 100 trials. The 1 year survival rate (14% for patients on the 100 mg/m^2 regimen and 15% for those on the 125 mg/m^2 regimen) was similar in both treatment arms. From this finding it was concluded that dose intensification would not result in any significant clinical benefit to patients with high AAG. Thus, a real trial was not performed because of the outcome of the simulated study.

These examples have been chosen to highlight the importance of appropriate analysis of simulated trials. It is worth noting that the number of replications in a simulated trial should be justified by the objectives of the study and the precision required (3). The presence of more than one endpoint raises the issue of multiplicity.

7.2.3 Multiplicity

Response variables in clinical trials are usually classified as primary or secondary endpoints. The multiplicity problem arises when performing many hypothesis tests on the same data set or having more than one endpoint. The presence of two or more primary endpoints in a clinical trial usually means that some adjustments of the observed p values for the multiplicity of tests may be required for the control of the type I error rate. To promote good clinical trial simulation practice, multiplicity should be addressed in simulated trials that lend themselves to it.

Consistency in the design and analysis of simulated and actual clinical trials is essential for CTS to be useful. The conclusions drawn from the interpretation of the results of a clinical trial are dependent on factors such as the disease under study, patient population, endpoints, study design, study conduct, appropriateness of statistical analysis for the given design, and the sensitivity of the chosen statistical test(s) to the scientific question(s) the study is designed to answer. The presence of multiplicity in a clinical trial that is unaccounted for in the design and ensuing analysis is one of the factors that may make interpretation of results difficult if not impossible.

7.2.3.1 Sources of Multiplicity

Multiplicity can arise from the presence of multiple active treatment arms in a study as in dose ranging studies, multiple endpoints because of the nature of the disease, multiple analyses and/or tests. Multiple analyses of clinical trials are usually the rule and not the exception and are executed in an effort to understand the trial outcome. There are usually per protocol analysis, evaluable subset analysis, protocol-defined subset analysis, and the all randomized (intent-to-treat) or all patients treated analysis in the case of confirmatory trials. Also, the need to

minimize the cost of obtaining data (interim analyses), the exploration of alternative statistical methods, and the desire to discover new aspects to the data (subset analyses) are necessary ingredients in the conduct of a clinical trial that will also introduce multiplicity. Thus, there is an inherent multiplicity component in most clinical trials, and it is important to consider and address multiplicity in the analysis of simulated clinical trials.

7.2.3.2 Error Rate and Control

Performing multiple tests is often reasonable in a clinical trial because of the cost of obtaining data. However, a negative feature of multiple testing is the greatly increased probability of declaring false significance (type I error). To prevent this, some adjustments for the observed p values for the multiple tests is required for the control of the type I error rate. By an error rate, one means the probability of false rejection of either the null hypothesis (referred to as the type I error rate), or the alternative hypothesis (referred to as the type II error rate). A type I error is committed when an ineffective drug gains entrance into the market, and a type II error is committed when an effective drug fails to reach the market. Both of these errors are equally serious.

The probability of false rejection of at least one of K null hypotheses in a multiple hypothesis testing problem can be controlled via one of two approaches: comparisonwise error rate or experimentwise error rate. If one opts to control the individual type I error rates at a given nominal α level for each of K hypotheses, then one is controlling what is called the comparisonwise error rate. On the other hand, controlling overall type I error rate at a given nominal α level for all possible K hypotheses is controlling the experimentwise error rate. If, for instance, $K = 4$ with a prespecified nominal α level of 0.05, a comparisonwise error rate method of control would imply testing each of the four corresponding null hypothesis at the 0.05 nominal level to ensure a 0.05 significance level for each test. On the other hand, an experimentwise type I error rate method of control would imply testing each of the four null hypotheses at some α level so as to ensure an overall significance level of 0.05 or less for all four tests (i.e., testing each of the null hypothesis at $\alpha = 0.013$ using the Bonferroni procedure, for instance).

Increased chance of a higher type I error rate is a well-known disadvantage associated with the comparisonwise error rate control approach. The probability of committing at least one type I error increases as the number of comparisons increases. If n independent tests are separately performed at $\alpha = 0.05$ level of significance, the probability that at least one will be significant is $1 - (1 - 0.05)^n$. As the type I error rate increases, the type II error rate decreases giving rise to more powerful tests (an experimenter's delight?). On the contrary, a type I error rate less than or equal to the designated α level would result from controlling the experimentwise error rate. From an experimenter's perspective, the downside of this approach is that the resulting type I error rate is usually smaller than

the designated α level which may lead to less sensitive tests (an experimenter's dismay?). Controlling experimentwise error rate instead of comparisonwise error rate does not depend on which null hypotheses are true. Controlling the experimentwise error rate seems to be a more appropriate option since it is not practical to require that an improvement be realized for every component of a global test in a given clinical trial. In addition, a multiple test procedure that controls the experimentwise error rate also controls the global error rate. The reverse is not true (29). A global test (which is outside the scope of this chapter) focuses on providing an overall probability statement about the efficacy of treatments by considering simultaneously all endpoints (30–40). It does not focus on any individual endpoint effect, which may sometimes be an important question of research.

7.2.3.3 Multiplicity Adjustment Approaches

The Bonferroni procedure is the most commonly used approach for multiplicity adjustment. There are improvements to this approach and they are referred to as stepwise procedures (41–44).

a. The Bonferroni Procedure. The simplest multiplicity adjustment procedure that is applicable to most multiplicity adjustment problems is the Bonferroni test. No assumptions are made about the correlation structure among the endpoints and/or data distributions. Thus, it is too conservative when there are many endpoints and/or the endpoints are highly correlated. If there are K endpoints, one accepts as statistically significant all those p values $\leq \alpha/K$, where α is the overall error type I error rate. $K \times p_k$ are the adjusted p values, where p_k are the observed p values and $k = 1, 2, \ldots, K$. This methodology would be applied to each replicate of the simulated trial.

There are other "less conservative" and more powerful improvements to the Bonferroni procedure that are also used to control experimentwise error rate. Among these are the Holm's procedure (39) and modifications and improvements on the Holm's procedure [i.e., the Hochberg (41) and Hommel (42) procedures] that are less conservative than the Holm's procedure. However, Holm's procedure requires fewer assumptions. The basis of these approaches is the realization that of the K null hypotheses to be tested, the only ones to protect against rejection at any given step are those not yet tested.

b. Bonferroni Modified (Stepwise) Procedures

HOLM'S PROCEDURE. A step-down adjustment procedure introduced by Holm (41) was an improvement on the Bonferroni method by providing additional power while maintaining the experimentwise error rate. The testing is done in a decreasing order of significance of (ordered) hypotheses. Testing for significance is continued until a null hypothesis is accepted. Thereafter all remaining (untested) null hypotheses are accepted without further testing.

The following is the algorithm for Holm's procedure:

1. Let $p_1 \leq p_2 \leq \ldots \leq p_K$ be the ordered p values and $H_{01}, H_{02}, \ldots, H_{0K}$ be the corresponding ordered null hypothesis.
2. Reject H_{01} if $p_1 < \alpha/K$ and go to the next step; otherwise stop and accept all null hypotheses.
3. Reject H_{02} if $p_{(2)} < \alpha/(K-1)$ and go to the next step; otherwise stop and accept all remaining $K-1$ null hypotheses.
4. In general reject H_{0k} if $p_k < \alpha/(K-k+1)$, k = 1, 2, . . . , K, and go to the next step; otherwise stop and accept the remaining null hypotheses. The adjusted p values are $p_{\text{adj},k} = \max\{(K-j+1) \times p_k\}$, $j = 1, 2, \ldots, k$ and $k = 1, 2, \ldots, K$.

Consider, for example, $K = 4$ endpoints and the following (ordered) p values were observed: $p_1 = 0.005 < p_2 = 0.020 < p_3 = 0.024 < p_4 = 0.081$. With the Holm procedure:

1. Reject H_{01} since $p_1 = 0.005 < 0.0125 = 0.05/4$ and go to the next step.
2. Given that $p_2 = 0.020 > 0.017 = 0.05/3$, stop and accept H_{02} and all remaining untested null hypotheses, H_{0k} ($k = 2, 3, 4$) since $p_2 > \alpha/3$. The resulting adjusted p values are $p_{\text{adj},1} = \max\{4 \times p_1\} = 0.020$, $p_{\text{adj},2} = \max\{4 \times p_1, 3 \times p_2\} = 0.060$, $p_{\text{adj},3} = \max\{4 \times p_1, 3 \times p_2, 2 \times p_3\} = 0.060$ and $p_{\text{adj},4} = \max\{4 \times p_1, 3 \times p_2, 2 \times p_3, 1 \times p_4\} = 0.081$.

HOCHBERG'S AND HOMMEL'S PROCEDURES. With the Hochberg procedure (41) testing is done in an increasing order of significance of the (ordered) hypotheses. Once a null hypothesis is rejected, significance testing is discontinued. Thereafter all remaining untested null hypothesis are rejected.

The following is the algorithm for the Hochberg procedure:

Let $p_1 \geq p_2 \geq \ldots \geq p_K$ be the ordered p values and $H_{01}, H_{02}, \ldots, H_{0K}$ be the corresponding null hypotheses. Reject H_{0k} and H_{0j} for $j > k$ and $k = 1, 2, \ldots, K$ if $p_k \alpha/k$. The adjusted p values are $p_{\text{adj},k} = \min\{j \times p_j\}$ for $j = 1, 2, \ldots, k$ and $k = 1, 2, \ldots, K$.

Using the same example as with the Holm procedure, the Hochberg procedure would proceed as follows:

1. Accept H_{01} since $p_1 = 0.081 > 0.05$ and go to the next step.
2. Given that $p_{02} = 0.024 < 0.025 = 0.05/2$, stop and reject H_{02} and all remaining null hypotheses $H_{0k}(k = 3, 4)$ that are untested. The resulting adjusted p values are $p_{\text{adj},1} = \min\{p_1\} = 0.081$, $p_{\text{adj},2} = \min\{p_1, 2 \times p_2\} = 0.048$, $p_{\text{adj},3} = \min\{p_1, 2 \times p_2, 3 \times p_3\} = 0.048$, and $p_{\text{adj},4} = \min\{p_1, 2 \times p_2, 3 \times p_3, 4 \times p_4\} = 0.020$.

The algorithm for the Hommel procedure (43) is as follows: Let $p_1 \leq p_2 \leq \ldots \leq p_K$ be the ordered p values, and $H_{01}, H_{02}, \ldots, H_{0K}$ the corresponding ordered null hypothesis. Assume an a priori p value of 0.05. The procedure is performed by starting in succession with $m = 1, 2, \ldots, K$ until the maximum m that has $p_{K-m+j} \geq j\ \alpha/m$ for $j = 1, 2, \ldots, m$ is found. Suppose the maximum m is equal to t. Then we reject all H_{0k}, $k = 1, \ldots, K$, for which $p_k \leq \alpha/t$.

Continuing with the example above, and starting with $m = 1$ and $j = 1$, $p_{K-m+j} = p_4 = 0.081$ which is greater than 0.05 ($j\times\ \alpha/m$). Proceed to $m = 2$, $j = 1$. $p_3 = 0.024 < 0.025$ ($j\times\ \alpha/m$). Therefore, the maximum m is 1. Reject all null hypotheses for which $p_k < 0.05$; i.e., reject H_{01}, H_{02}, H_{03} as with the Hochberg procedure. The resulting adjusted p values are $p_{\text{adj},k} = m \times p_k = p_k$, $k = 1, 2, 3, 4$; i.e., $p_{\text{adj},1} = 0.081$, $p_{\text{adj},2} = 0.024$, $p_{\text{adj},3} = 0.020$, and $p_{\text{adj},4} = 0.005$.

It is worth noting from the example that the Hochberg and Hommel procedures lead to two more null hypothesis rejections (H_{02} and H_{03}) than the Holm procedure. The Holm procedure is uniformly more powerful than the Bonferroni procedure, the Hochberg procedure is uniformly more powerful than the Holm, and the Hommel procedure has been shown to be only slightly more powerful than the Hochberg (42–43).

Some Ad Hoc Procedures. Some ad hoc procedures that make use of correlation information among endpoints without any distributional assumptions have been developed. Tukey et al. (45) have suggested that for strongly correlated endpoints and for a given nominal α level the adjustments $p_{\text{adj},k} = 1 - (1 - p_k)^{\sqrt{K}}$ and $\alpha_k = 1 - (1 - \alpha)^{1/\sqrt{K}}$ be used, where p_k and $p_{\text{adj},k}$ are the observed and adjusted p values and α_k is the adjusted critical level for the kth hypothesis for $k = 1, \ldots, K$. Dubey (46) and Armitage and Parmar (47) have suggested the use of $p_{\text{adj},k} = 1 - (1 - p_k)^{m_k}$ and $\alpha_k = 1 - (1 - \alpha)^{1/m_k}$, where

$$m_k = K^{1-r_k}$$

and

$$r_k = \frac{1}{K-1} \sum_{j \neq k}^{K} r_{jk}$$

replaces $\sqrt{K}$ in the Tukey et al. (45) formula for $p_{\text{adj},k}$; r_{jk} is the correlation coefficient between the jth and the kth endpoints. A nice feature of the Dubey (46) and Armitage and Parmar (47) procedure is that when the average of the correlation coefficients is 1, the adjusted and the unadjusted p values are the same; and when it is zero, the adjustment is according to the Bonferroni test. For equicorrelated endpoints with 0.5 correlation coefficient, the Dubey (46) and Armitage and Parmar (47) procedure is equivalent to the Tukey et al. (45) procedure.

Some of the approaches for addressing multiplicity in clinical trials have been discussed with their applicability. It is important to use the appropriate mul-

tiplicity adjustment procedure in the analysis of simulated clinical trials. This would lead to a meaningful interpretation of the outcome of the simulation experiment.

7.3 COMMUNICATION OF ANALYSIS RESULTS

The results of the analysis of a simulated clinical trial should be presented in a manner that can be readily understood by the intended audience. Communication of the results of a simulation project can be effected through words, tables, or graphics. The efficient conveyance of the message from a simulation project should involve the use of all three communication media, which has been described respectively as infantry, artillery, and cavalry of the pharmacokinetic/pharmacodynamic defense force (21). These methods should be used to supplement one another, although the effectiveness of each depends on the contents of the message. Tables should be used to communicate information that can best be conveyed using this means of presentation. Graphics, in particular, are essential for conveying relations and trends in an informal and simplified visual form.

Graphs are analogous to written language; they communicate quantitative and categorical information. Written language communicates thoughts, ideas, observations, emotions, theories, hypotheses, numbers, etc. Graphical language is used extensively to convey information because it does so, effectively. Quantitative patterns and relationships in data are readily revealed by graphs because of the enormous power of the eye-brain system to perceive geometrical patterns. The eye-brain system can quickly summarize vast amounts of quantitative information on a graph, perceiving salient features, or focusing on specific detail. The power of a graph is its ability to enable one to take in the quantitative information, organize it, and see patterns and structure not readily revealed by other means of studying and presenting data (21, 48).

Numerical data can be displayed in different formats, but only some are well suited to the information processing capacity of the human vision. The phrase *graphical perception* has been coined by Cleveland and McGill (49) to refer to the role of visual perception in analyzing graphs. These authors have studied several elementary visual tasks (such as discrimination of slopes or lengths of lines) relevant to graphical perception. They attribute the great advantage of graphical displays (e.g., scatterplots) over numerical tables to the capacity of the human vision to process pattern information globally at a glance. In comparing the effectiveness of data display using graphs and numerical tables, Legge et al. (50) found perceptual efficiencies to be very high for scatterplots, $\geq$60%. Efficiencies were much lower for numerical tables, 10%. Efficiency in the study referred to the performance of a real observer relative to the ideal observer. The ideal observer makes an optimal use of all available information (50). Performance with scatterplots was reported to have the hallmark of a parallel process:

weak dependence on viewing time. Processing of tables of numbers was found to be performed in a much more serial fashion. Efficiencies dropped roughly with increasing information content in the tables and increased in rough proportion to viewing time. They concluded that entries in tables are processed sequentially at a fixed rate. Given enough viewing time, efficiency of information processing from tables could approach that of graphics (50). Thus, information content of tables should be kept to a minimum to allow efficient extraction of such information by the reader, or the audience in the case of oral presentation of the outcome of a simulation project.

However, it is worth noting that the merits of a graphical display depend on the information chosen for display and the amount of effort that will be expended by the reader in deciphering what is encoded in the graph (51). In summarizing how sample size affects the power of a study, an integrated display such as a line graph would be a superior method of display. This is because in decoding the graph the reader would have to compare the power of the study for the different sample sizes and integrate that information to form his opinion. Judging change requires comparing quantities and integrating that information (51). A line graph is more effective in conveying change than other types of display (51). This is because the eye is focused on the physical slope of the line. Bar plots are also effective in conveying change (trends) in that the eye, in decoding change (or a trend) with bar plots, is tracing a perceived slope (51). The effectiveness of a graph, therefore, depends on the amount of work that is to be performed by the reader in decoding the information contained in the graphical display.

In summarizing the results of a population pharmacokinetic study in which the effect of sample size on the bias and precision with which population pharmacokinetic parameters were estimated, Ette et al. (52) used line plots. A similar line plot display (Figure 1) of the effect of sample size and intersubject variability on the estimation of population pharmacokinetic parameters was created from an aspect of data generated in a simulation study performed to determine the performance of mixed designs in population pharmacokinetic studies (53). The plot shows the influence of intersubject variability on parameter estimation as sample size was varied. In the study, the effect of three different levels of intersubject variability, ranging from 30 to 60% coefficient of variation, and different sampling designs on the sample size required for efficient population pharmacokinetic parameter estimation were investigated. However, in Figure 1 data for only one of the designs are plotted to illustrate the effectiveness of a line plot.

The *coplot* (21, 54) is a powerful tool for studying how a response depends on two or more factors. It presents conditional dependence in a visually powerful manner. Two variables are plotted against each other in a series of overlapping ranges. This enables one to see how a relationship between two variables (y and x) changes as a third variable z changes, i.e., $y \sim x \mid z$. Thus y is plotted against

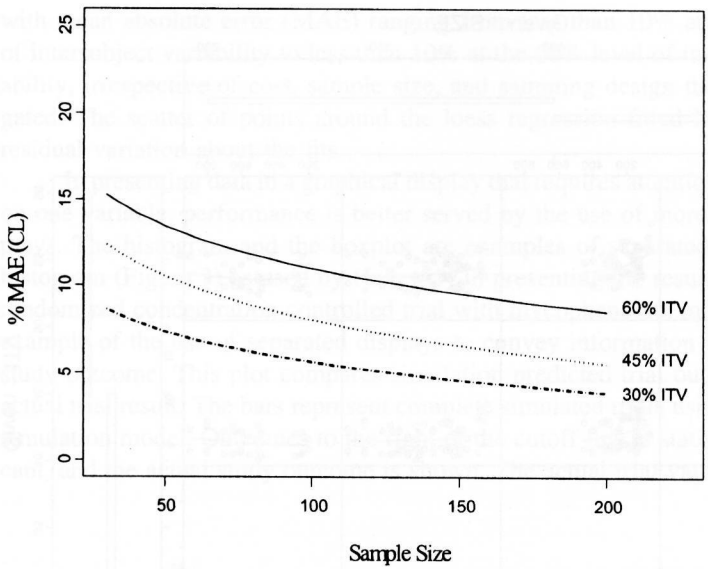


FIGURE 1 A line plot of the effect of sample size and intersubject variability on the precision (mean absolute error: %MAE) with which clearance (CL) was estimated in a simulated population pharmacokinetic study. [Created from data reported in Fadiran et al. (53).]

x for a series of conditioning intervals determined by z. Coplots may have two simultaneous conditioning variables, i.e., $y \sim x \mid z_1 * z_2$.

In studying the performance of mixed design in population pharmacokinetic studies Fadiran et al. (53) used conditioning plots (coplots) to summarize their findings (Figure 2). They investigated the influence of total cost of a population design (i.e., the combined cost of individual designs expressed as the total number of samples for a sample size given the sampling design considered), sample size, and intersubject variability on bias and precision associated with population pharmacokinetic parameter estimation. The presentation of the results from that study using coplots is a good example of using multipanel display to summarize the results of a simulated study; and this is recommended in the consensus document on clinical simulation (24). In this example, the use of the coplot allowed information from four variables to be presented using a two-dimensional graphical display.

Figure 2, extracted from the work of Fadiran et al. (53), is a coplot illustrating the effect of cost on the precision associated with the estimation of drug

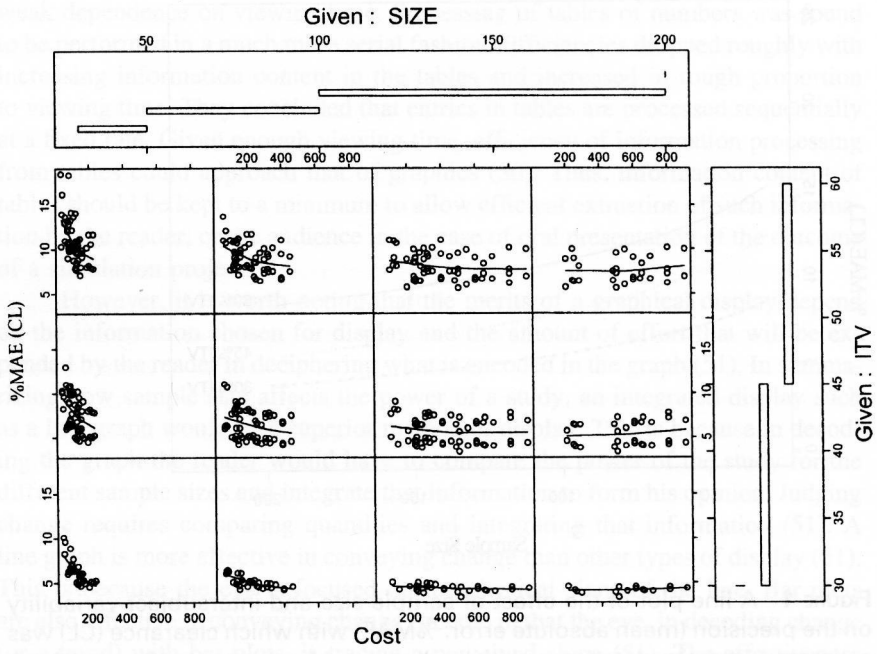


FIGURE 2 A conditioning plot of a simulated population pharmacokinetic study showing the dependence of precision (expressed as percent mean absolute error (%MAE) in clearance (CL) on cost (i.e., total number of samples) conditioned on sample size (SIZE, top panel) and levels of intersubject variability (ITV, right panel). The study evaluated the performance of mixed designs in population pharmacokinetic study. [Excerpted from Fadiran et al. (53).]

clearance (CL) as sample size and intersubject variability are altered (53). The panel at the top is the *given panels* (*or conditioning intervals*), the panels below is the *dependence panel* (cost). The panel at the right is also the *given panels* (*or conditioning intervals*). Locally weighted regression (loess) lines are added to the plots to aid perception. The rectangles of the top conditioning panel specify the sample size, while those on the right specify the intersubject variability. On the corresponding dependence panel precision in clearance estimation (expressed as percent mean absolute error, %MAE) is graphed against the cost of a population design for those estimates whose value lie in the interval. The graph is read from left to right of the bottom row, then from left to right of the next row, and so forth. It can be seen from the figure that the CL was reasonably well estimated

with mean absolute error (MAE) ranging from less than 10% at the 20% level of intersubject variability to less than 10% at the 60% level of intersubject variability, irrespective of cost, sample size, and sampling design that was investigated. The scatter of points around the loess regression fitted lines reflect the residual variation about the fits.

In presenting data in a graphical display that requires attention to be focused on one variable, performance is better served by the use of more separated displays. The histogram and the boxplot are examples of separated displays. The histogram (Figure 3) as used by Hale (55) in presenting the results of simulated randomized concentration controlled trial with mycophenolate mofetil is a good example of the use of separated displays to convey information on a simulated study outcome. This plot compares simulation predicted trial outcomes and the actual trial result. The bars represent complete simulated trials using a developed simulation model. Outcomes to the right of the cutoff line is statistically significant, and the actual study outcome is shown. The actual trial value fell between

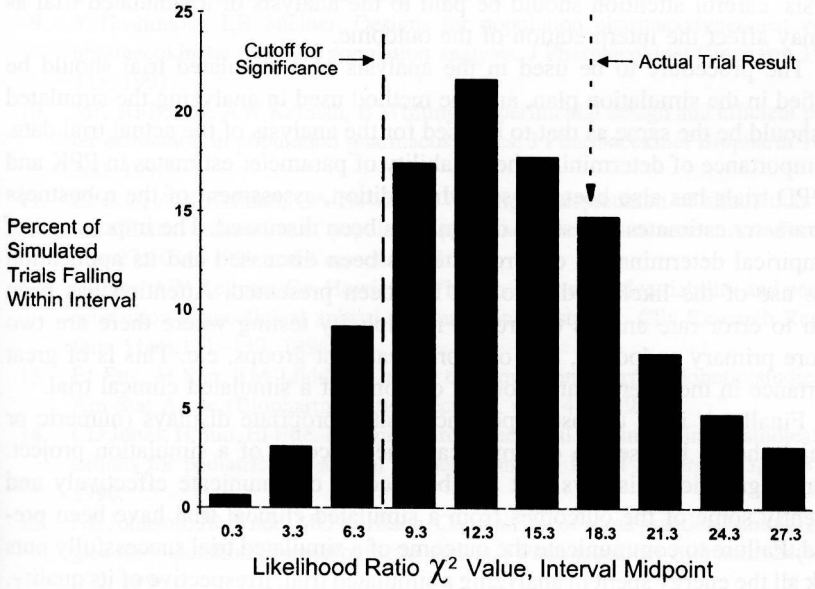


FIGURE 3 Frequency distribution of the test statistic for the primary analysis resulting from 500 completed simulated RCCTs with mycophenolate mofetil under "worst case" trial conditions, completed before real study initiation. The actual study outcome is shown, falling in the central portion of the distribution in the interval centered at 18.3 [Excerpted and adapted from Hale [55].]

the 80th and 90th percentile of the simulated results, which means that the actual trial outcome is not unusual based on the simulation model, which reflects that the simulation model was a reasonable description of the trial process.

The boxplot has proven to be a popular graphical method for displaying and summarizing univariate data, to compare parallel batches of data and to supplement more complex displays with univariate information. Its appeal is due to the simplicity of the graphical construction (based on quartiles) and the many features that it displays (location, spread, skewness, and potential outliers). Boxplots are useful for summarizing distributions of treatment outcomes. A good example would be the comparison of the distribution of response to treatment at different dose levels.

7.5 SUMMARY

A simulated clinical trial has been likened to a meta-analysis of clinical trials. While meta-analysis is retrospective, a simulated trial is prospective. Like meta-analysis, careful attention should be paid to the analysis of a simulated trial as this may affect the interpretation of the outcome.

The procedure to be used in the analysis of a simulated trial should be specified in the simulation plan, and the method used in analyzing the simulated trial should be the same as that to be used for the analysis of the actual trial data. The importance of determining the reliability of parameter estimates in PPK and PPK/PD trials has also been stressed. In addition, assessment of the robustness of parameter estimates and study designs has been discussed. The importance of an empirical determination of error rate has been discussed and its application in the use of the likelihood ratio test has been presented. Attention has been drawn to error rate and its control in multiplicity testing where there are two or more primary endpoints, two or more treatment groups, etc. This is of great importance in the interpretation of the outcome of a simulated clinical trial.

Finally, it is of utmost importance that appropriate displays (numeric or graphic) should be used to communicate the outcome of a simulation project. Different graphical displays that can be used to communicate effectively and efficiently some of the outcomes from a simulated clinical trial have been presented. Failure to communicate the outcome of a simulated trial successfully puts at risk all the energy spent in analyzing a simulated trial, irrespective of its quality.

REFERENCES

1. CC Peck. Drug development: improving the process. Food and Drug Law J 52(2): 163–167, 1997.

2. M Hale, WR Gillespie, SK Gupta, B Tuk, NH Holford. Clinical trial simulation: streamlining your drug development process. Appl Clinic Trials 5:35–40, 1996.
3. US Food and Drug Administration. Guidance for Industry: Population Pharmacokinetics. February 1999. http://www.fda.gov/cder/guidance/index.htm
4. H Sun, EO Fadiran, CD Jones, L Lesko, SM Huang, K Higgins, C Hu, S Machado, S Maldonado, R Williams, M Hossain, EI Ette. Population pharmacokinetics: a regulatory perspective. Clin Pharmacokinet 37:41–58, 1999.
5. NHG Holford, HC Kimko, JPR Monteleone, CC Peck. Simulation of clinical trials. Annu Rev Pharmacol Toxicol 40:209–234, 2000.
6. DR Jones. Computer simulation as a tool for clinical trial design. Int J Bio-Med Comp 10:145–150, 1979.
7. KL Lee, M Frederick, CF Starmer, PJ Harris, RA Rosati. Clinical judgment and statistics: lessons from a simulated randomized trial in coronary artery disease. Circulation 61(3):508–515, 1980.
8. MD Hale, AJ Nicholls, RE Bullingham, R Hene, A Hoitsma, JP Squifflet, W Weimar, Y Vanrenterghem, FJ Van de Woude, GA Verpooten. The pharmacokinetic-pharmacodynamic relationship for mycophenolate mofetil in renal transplantation. Clin Pharmacol Ther 64:672–683, 1998.
9. Y Hashimoto, LB Sheiner. Designs for population pharmacodynamics: value of pharmacokinetic data and population analysis. J Pharmacokinet Biopharm 19:333–353, 1991.
10. MK Al-Banna, AW Kelman, B Whiting. Experimental design and efficient parameter estimation in population pharmacokinetics. J Pharmacokinet Biopharm 18:347–360, 1990.
11. EI Ette, AW Kelman, CA Howie, B Whiting. Interpretation of simulation studies for efficient estimation of population pharmacokinetic parameters. Ann Pharmacother 27:1034–1039 and Correction 27:1548, 1993.
12. EI Ette, AW Kelman, CA Howie, B Whiting. Inter-animal variability and parameter estimation in preclinical animal pharmacokinetic studies. Clin Research Regul Affairs 11(2):121–139, 1994.
13. EI Ette, H Sun, TM Ludden. Design of population pharmacokinetic studies. Proc Am Stat Assoc (Biopharmaceutics Section), 487–489, 1994.
14. CD Jones, H Sun, EI Ette. Designing cross-sectional pharmacokinetic studies: implications for pediatric and animal studies. Clin Res Regul Affairs 13(3,4):133–165, 1996.
15. NE Johnson, JR Wade, MO Karlsson. Comparison of some practical sampling strategies for population pharmacokinetic studies. J Pharmacokinet Biopharm 24(6):245–272, 1996.
16. H Sun, EI Ette, TM Ludden. On error in the recording of sampling times and parameter estimation from repeated measures pharmacokinetic data. J Pharmacokinet Biopharm 24(6):635–648, 1996.
17. C Veyrat-Follet, R Bruno, R Olivares, GR Rhodes, P Chaikin. Clinical trial simulation of docetaxel in patients with cancer as a tool for dosage optimization. Clin Pharmacol Ther 68(6):677–687, 2000.

18. P Williams, JR Lane, CC Turkel, EV Capparelli, Z Dziewanowska, AW Fox. Dichloroacetate: population pharmacokinetics with a pharmacodynamic sequential link model. J Clin Pharmacol 41:259–267, 2001.
19. GV Glass. Primary, secondary and meta-analysis of research. Educ Res 5:3–8, 1976.
20. K Dickersin, JA Berlin. Meta-analysis: state-of-science. Epidemiol Rev, 14:154–176, 1992.
21. EI Ette. Statistical graphics in pharmacokinetics and pharmacodynamics: a tutorial. Ann Pharmacother 32:818–828, 1998.
22. KG Kowalski, MM Hutmacher. Design evaluation for a population pharmacokinetic study using clinical trial simulations: a case study. Stat Med 20:75–91, 2001.
23. SL Beal, LB Sheiner. NONMEM Users Guide, parts I–VI. San Francisco, CA: Division on Clinical Pharmacology, University of San Francisco, 1979–1992.
24. NHG Holford, M Hale, HC Ko, J-L Steimer, LB Sheiner, CC Peck. Simulation in Drug Development: Good Practices. Washington, DC: CDDS, Georgetown University, 1999. http://cdds.georgetown.edu
25. Pinheiro JC, Bates DM, eds. Mixed-Effects Models in S and S-PLUS. New York: Springer, 2000.
26. DO Stram, JW Lee. Variance components testing in the longitudinal mixed effects model. Biometrics 50:1171–1177, 1994.
27. U Wählby, EN Jonsson, MO Karlsson. Assessment of actual significance levels for covariate effects in NONMEM. J Pharmacokinet Pharmacodyn 28:231–252, 2001.
28. Gelman A, Carlin JB, Stern HS, Rubin DB, eds. Bayesian Data Analysis. London: Chapman and Hall, 1995.
29. P Bauer. Multiple testing in clinical trials. Stat Med 10:871–890, 1991.
30. G Hommel. Tests of overall hypothesis for arbitrary dependent structures. Biometrical J 25(5):423–430, 1983.
31. G Hommel. Multiple test procedures for arbitrary dependence structures. Metrika 33(6):321–336, 1986.
32. W Lehmacher, G Wassmer, P Reimer. Procedures for two-sample comparisons with multiple endpoints controlling the experimentwise error rate. Biometrika 67:655–660, 1991.
33. PC O'Brien. Procedures for comparing samples with multiple endpoints. Biometrics 40:1079–1089, 1984.
34. SJ Pocock, NL Geller, A Tsiatis. The analysis of multiple endpoints in clinical trials. Biometrics 43:487–498, 1987.
35. AJ Sankoh, MF Huque. A note on O'Briens OLS and GLS tests. Biometrics 51: 1580–1581, 1995.
36. DI Tang, C Gnecco, NL Geller. Design of group sequential trials with multiple endpoints. J Am Stat Assoc 84:776–779, 1989.
37. DI Tang, C Gnecco, NL Geller. An approximate likelihood ratio test for a normal mean vector with nonnegative components with application to clinical trials. Biometrika 76:751–754, 1989.
38. DI Tang, NL Geller, SJ Pocock. On the design and analysis of randomized clinical trials with multiple endpoints. Biometrics 49:23–30, 1993.
39. S Holm. A simple sequentially rejective Bonferroni Test Procedure. Scandinavian Journal of Statistics 6:65–70, 1979.

40. RJ Simes. An improved Bonferroni procedure for multiple tests of significance. Biometrika 67:655–660, 1976.
41. Y Hochberg. A sharper Bonferroni procedure for multiple significance testing. Biometrika 75:800–803, 1988.
42. G Hommel. A comparison of two modified Bonferroni procedures. Biometrika 76: 624–625, 1989.
43. GA Hommel. A stagewise rejective multiple test procedure based on a modified Bonferroni test. Biometrika 75:383–386, 1988.
44. CW Dunnet, AC Tamhanne. A step-up multiple test procedure. J Am Stat Assoc 87:162–170, 1993.
45. JW Tukey, JL Ciminera, JF Heyse. Testing the statistical certainty of a response to increasing doses of a drug. Biometrics 41:295–301, 1985.
46. SD Dubey. Adjustment of p values for multiplicities of intercorrelating symptoms. Proceedings of the VIth International Society of Clinical Biostatisticians, Germany, 1985.
47. P Armitage, M Parmar. Some approaches to the problem of multiplicity in clinical trials. Proceedings of the XIIth International Biometrics Conference, Seattle, Washington, 1986.
48. EI Ette, P Williams, H Sun, FO Ajayi, LC Onyiah. The process of knowledge discovery from large pharmacokinetic data sets. J Clin Pharmacol 41:25–54, 2001.
49. WS Cleveland, R McGill. Graphical perception and graphical methods of analyzing scientific data. Science 229:828–833, 1985.
50. GE Legge, Y Gu, A Luebker. Efficiency of graphical perception. Percep Psychophys 46:365–374, 1989.
51. JS Hollands, I Spence. Judgement of change and proportion in graphical perception. Human Factors 34:313–334, 1992.
52. EI Ette, H Sun, TM Ludden. Balanced designs in longitudinal population pharmacokinetic studies. J Clin Pharmacol 38:417–423, 1998.
53. EO Fadiran, CD Jones, EI Ette. Designing population pharmacokinetic studies: performance of mixed designs. Eur J Drug Metab Pharmacokinet 25:231–239, 2000.
54. WS Cleveland. Visualizing Data. Summit, NJ: Hobart Press, 1993.
55. MD Hale. Using population pharmacokinetics for planning a randomized concentration–controlled trial with a binary response. In: L Aarons, LP Balant, M Danhof, M Gex-Fabry, UA Gundert-Remy, MO Karlsson, F Mentre, PL Morselli, F Rombout, M Rowland, JL Steimer, S Vozeh, eds. The Population Approach: Measuring and Managing Variability in Response, Concentration and Dose. Commission of the European Communities, European Cooperation in the Field of Scientific and Technical Research; Brussels 1997, pp 228–235.

Ogenstad Declaration Exhibit H

Journal of Biopharmaceutical Statistics, 18: 611–630, 2008

ISSN: 1054-3406 print/1520-5711 online
DOI: 10.1080/10543400802071303

CLINICAL TRIALS SIMULATION: A STATISTICAL APPROACH

Peter H. Westfall[1], Kuenhi Tsai[4], Stephan Ogenstad[3], Alin Tomoiaga[1], Scott Moseley[2], and Yonggang Lu[5]

[1]*Texas Tech University, Lubbock, Texas, USA*
[2]*Vertex Pharmaceuticals Incorporated, Cambridge, Massachusetts, USA*
[3]*Statogen Consulting, LLC, Wake Forest, North Carolina, USA*
[4]*Merck & Co. Inc., North Wales, Pennsylvania, USA*
[5]*JP Morgan Chase Bank, N. A., Columbus, Ohio, USA*

A generic template for clinical trials simulations that are typically required by statisticians is developed. Realistic clinical trials data sets are created using a unifying model that allows general correlation structures for endpoint*timepoint data and nonnormal distributions (including time-to-event), and computationally efficient algorithms are presented. The model allows for patient dropout and noncompliance. A grid-enabled SAS-based system has been developed to implement this model; details are presented summarizing the system development. An example illustrating use of the system is given.

Key Words: Binary data; Correlation structure; Grid computing; Mixture data; Non-normality; Non-compliance; Ordinal data; Patient dropouts; Survival data.

1. INTRODUCTION

Drug development is not for the faint hearted. The clinical trial execution risk is high. In the process of bringing a new compound to the market, every delay can add millions of dollars in added expenses and lost revenues. Clinical trials simulation (CTS) helps the project team minimize risks and guide decision making by formalizing assumptions, quantifying and testing uncertainties. The simulations can be used for defining and testing interactive drug models, exploring and communicating study design attributes, and performing analyses around precision and accuracy of future endpoint estimates. CTS incorporates accessible scientific knowledge to help the entire project team communicate and test ideas, and to plan significant, effective trials for every phase of clinical development. The trial simulation helps the team anticipate risks and preview the range of expected results before huge investments are committed. An important use of CTS is the development of "mock up" trials: project team members from various disciplines

Received April 5, 2007; Accepted July 30, 2007
Address correspondence to Peter H. Westfall, Texas Tech University, Lubbock, Texas 79409-2101, USA; E-mail: peter.westfall@ ttu.edu

utilize the CTS to explore a series of scenarios, from different trial designs, to alternative ways of analyzing the generated data. As a result, the project team can receive prompt feedback on the impact on trial outcomes that alternative designs and analysis methods could have presented in the future.

Thus, optimization of clinical trial design is the main focus of this paper. Because trials and data resulting from it are often too complex to allow simple decision-theoretic solutions, interest in CTS has recently exploded in popularity among statisticians, clinicians, and pharmacokineticists. Current trends within the pharmaceutical industry and within the offices of some regulatory agencies suggest a promising future for modeling and simulation—incorporating pharmacokinetic modeling and clinical trial simulation (Holford et al., 2000; Gieschke and Steimer, 2000; Machado et al., 1999; Reigner et al., 1997; Sheiner and Steimer, 2000). The support from government authorities has boosted the interest in clinical trial simulations. Besides the more straightforward application of sample size allocation, design optimization may include protocols (for example, choice of ideal models and test statistics), and estimation of operating characteristics of nonstandard and computationally intensive procedures (including Bayesian and adaptive designs). Planning and conduct of late-phase clinical trials in drug development has the potential to be revolutionized by CTS. It has the ability to transform drug development by making better use of prior data and information and to explore important clinical trial designs. Peck et al. (2003) outline an ambitious but possible future that CTS might sometimes replace the second Phase III trial, and therefore needing only a single trial. Simulation gains credibility with the nonscientist because it is understandable without technical terms, and this gives transparency to several otherwise difficult principles affecting opinion and behavior. Therefore, CTS is useful in designing clinical trials in drug development (Hale et al., 1996; Holford et al., 2000; Peck and Desjardins, 1996). Since its beginning in a teaching environment (Maxwell et al., 1971), there have been several success stories recounting the value of simulation for design of clinical trials. These include reports where simulations proved helpful in explaining proposed trial designs not only to their own internal members but also to regulatory authorities, such as the Food and Drug administration (FDA) (Hale, 1997). The FDA has cosponsored several meetings to discuss the value of modeling and simulation and has advised the industry to perform modeling and simulation when deciding the best trial design. The first simulation of a proposed randomized concentration controlled trials of mycophenolate mofitile was the reported demonstration of the practical utility of clinical trial simulation (Hale et al., 1998). This simulation helped to settle trial feasibility, and influenced the design of a study.

Figure 1 displays the essential idea of what we mean when we refer to CTS; others may emphasize different aspects. Often, "variations in study design" refers simply to different sample sizes, but the idea is much broader, encompassing length of trial, measurement of endpoints (continuous, time-to-event, categorized, binary), and analysis methods (baseline covariate-adjusted vs. percentage change, use of compliance data as covariates, parametric vs. nonparametric, etc.). The horizontal axis of Figure 1 need not be ordinal; the graph suggests ordinality for illustrative purposes only.

Similarly, although "probability of successful outcome" often means "power," the possibilities are much broader, encompassing combination rules involving both

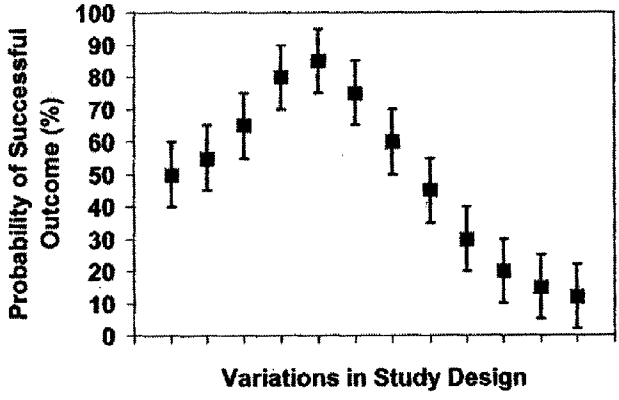


Figure 1 Simulation methodology to optimize trial design.

safety and efficacy, or complex rules like "3 out of 4 significant" for multiple co-primary endpoints (U.S. Department of Health and Human Services, Food and Drug Administration, 1999, p. 3), rules that include economic considerations (Poland and Wadda, 2001), and rules involving patient quality of life. In the more general case, the vertical axis of Figure 1 will be replaced by "Expected Benefit."

Some commercial software (e.g., Pharsight, http://www.pharsight.com/main.php) require pharmacokinetic/pharmacodynamic (PK/PD) inputs as drivers for the simulation output. Several references on clinical trials simulation using PK/PD exist, see Kimko and Duffull (2003) for an overview and further references. We have taken an alternative, more "statistical" approach for simplicity, to produce data with realistic characteristics that are useful for decisions that statisticians typically must make. An example of a similarly "statistical" approach is given by Anderson et al. (2003). Our model starts with a rich probabilistic structure to account for typical scenarios, using historical data where possible to validate the inputs and outputs, with specific emphasis on the parsimonious yet flexible input of correlation structures. The output data sets are massive, and the analyses are allowed to be computationally challenging, often requiring grid computing for feasibility.

Our framework for multivariate simulation is reasonably simple to code using a variety of software, yet flexible, retaining the realism the doubly-multivariate endpoint/timepoint correlation structures, informative dropout mechanisms, non-normal distributions, survival endpoints, and noncompliance effects. This research is based on the development of a real, currently existing CTS software system, and simplifying, albeit somewhat questionable assumptions are made at various places. Such assumptions reflect a necessary trade-off between ease of use of the system on one hand, and realism and flexibility of its outputs on the other hand. The simulation algorithms are not entirely new, but we hope that pharmaceutical statisticians will find it useful to have them all in one place, for convenient reference. We also hope that the models we present for compliance and dropout effects will stimulate research into this area, whether to instantiate the models we suggest using parameter estimates from designed experiments, or to propose alternative models that may be more appropriate for specific diseases. Section 2 provides an overview of the system, Sections 3–6 provide technical details

regarding data generation and inputs, Section 7 provides an application, and Section 8 concludes.

2. OVERVIEW OF A STATISTICAL CLINICAL TRIALS SIMULATION SYSTEM

An overview of the system is shown in Figure 2.

Patient responses are functions of underlying correlated $N(0,1)$ clinical quantities; all distributional forms and dropout effects are determined from these underlying values. Evaluation of trial success then follows from the analysis of the simulated data sets. Details follow.

3. THE FUNDAMENTAL CORRELATED QUANTITIES

Our goal is to generate realistic data sets having typical correlation structures for multiple endpoint/timepoint data with p endpoints (safety, efficacy or both) indexed by $j = 1, \ldots, p$, and $T + 1$ timepoints indexed by $t = 0, \ldots, T$ ($t = 0$ denotes baseline). To start, we generate for "patient i" a $p(T + 1)$-vector of

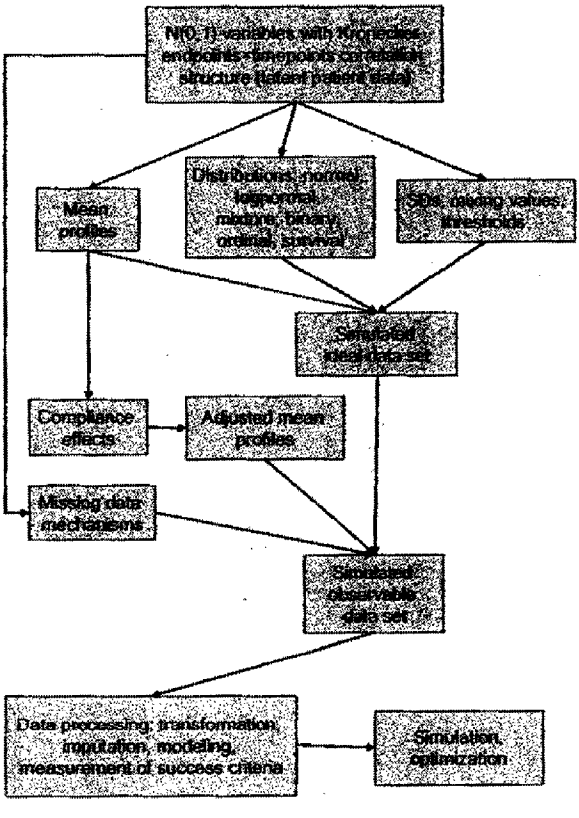


Figure 2 Overview of clinical trials simulation system.

correlated $N(0, 1)$ variates Z_{iji}, each of which may be thought of as a latent indicator of the patient's health relative to a population of similar patients, for endpoint j and timepoint t (see the upper left panel of Figure 2). Observations Z_{ijt} and $Z_{i'j't'}$ will always be considered independent when $i \neq i'$; random center effects violate the assumption of independence between patients; it is possible to include such effects, but this is not pursued here. To simplify notation, we frequently drop subscripts; meanings will be clear from context.

In some cases, code for simulating the data Z_{ijt} can utilize a high-level matrix language such as MAPLE or SAS/IML. Our intent is to avoid such software for two reasons: First, it is desirable that the simulation system be entirely self-contained within a single software system, from data generation through analysis, SAS in our case. However, in our experience, the SAS/IML component is frequently not licensed, thus there is a need to keep the code at as "native" a level as possible, using SAS/BASE and SAS/STAT only. Second, we exploit special structures to obtain more efficient algorithms by operating at a more native level. Nevertheless, we also provide instructions for use with matrix-intensive software.

3.1. Repeated Measures Data for Patient*Endpoint

For a given patient and endpoint, the timepoint data $Z_0, \ldots, Z_T$ are correlated because of subject and carryover effects. Frison and Pocock (1992) argue for generic use of the compound symmetry (CS) covariance structure, which accommodates subject effects only, and not time-series carryover effects, but also note "allowance in design for alternative non-equal correlation structures can and should be made when necessary." The CS model can be expanded easily to accommodate time-series carryover effects in addition to subject effects as

$$Z_t = \theta^{1/2} S + (1 - \theta)^{1/2} \epsilon_t, \tag{1}$$

where $S \sim N(0, 1)$ is the subject effect and $\epsilon_0, \ldots, \epsilon_T$ is a realization of a unit variance AR(1) process with parameter ρ. Since repeated-measures software such as PROC MIXED of SAS/STAT have become available, the routine use of correlation structures other than CS is now commonplace. The model can be estimated using existing data using *both* the 'RANDOM' and 'REPEATED' statements:

```
random subject_id; repeated /subject=subject_id type=ar(1);
```

For simulation purposes, the parameters θ and ρ must be specified (typically suggested by early Phase data and/or similar studies); it may be helpful to note that θ is the within-subject correlation for large time lags. The data may be generated easily using (1), where the $\{\epsilon_t\}$ are generated recursively as

$$\epsilon_t = \rho \epsilon_{t-1} + (1 - \rho^2)^{1/2} u_t, \tag{2}$$

with the variates $S, \epsilon_{-1}, u_0, \ldots, u_T$ generated as i.i.d. $N(0, 1)$. The resulting covariance structure within patient*endpoint combination is

$$\mathrm{Cov}(Z_0, \ldots, Z_T) = \boldsymbol{\Sigma} = \theta \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix} + (1-\theta) \begin{bmatrix} 1 & \rho & \cdots & \rho^T \\ \rho & 1 & \cdots & \rho^{T-1} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^T & \rho^{T-1} & \cdots & 1 \end{bmatrix}. \quad (3)$$

It is also possible to generate $\boldsymbol{Z}' = (Z_0, \ldots, Z_T)$ using matrix decomposition. Factoring $\boldsymbol{\Sigma}$ as $\boldsymbol{\Sigma} = \boldsymbol{U}'\boldsymbol{U}$ (e.g., using the Cholesky decomposition available in matrix languages such as SAS/IML, or using outputs from PROC PRINCOMP of SAS/STAT), and letting $\boldsymbol{W} = \{W_j\}$ denote a $(T+1)$-vector of independent $N(0, 1)$ variates, we may simply take $\boldsymbol{Z} = \boldsymbol{U}'\boldsymbol{W}$. However, it is worth noting that (1) and (2) are computationally more efficient and do not require specialized matrix functions.

3.2. Multiple Endpoint Data for Patient*Timepoint

For a given patient and timepoint t, the correlation between endpoints $Z_{1t}, \ldots, Z_{pt}$ is best left unstructured (UN), rather than assumed as CS, AR(1) or some other form. For estimation purposes it is often desirable to parameterize a UN correlation using a more parsimonious form such as factor-analytic approximation (e.g., FA(1) or FA(2) in PROC MIXED). However, the actual endpoint correlation matrix (e.g., from early phase or similar studies) is a more convenient input to the system. Denote the covariance matrix for the multiple endpoints as

$$\mathrm{Cov}(Z_{1t}, \ldots, Z_{pt}) = \boldsymbol{\Gamma} = \begin{bmatrix} 1 & \gamma_{12} & \cdots & \gamma_{1p} \\ \gamma_{12} & 1 & \cdots & \gamma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{1p} & \gamma_{2p} & \cdots & 1 \end{bmatrix};$$

like the parameters θ and ρ, the correlations $\{\gamma_{ij}\}$ are system inputs. For now, the basic quantities Z_{jt} have unit variance, so correlation equals covariance.

3.3. Correlation Structure for all Within-Patient Data

Observations between endpoints at different timepoints are correlated. There are a number of possibilities for defining this structure, the most convenient and common is the Kronecker product model used in multivariate longitudinal models (Gao et al., 2006). This model implies that the covariance matrix of the entire set of $p(T+1)$ endpoint and timepoint measurements for a given patient is

$$\mathrm{Cov}(\mathbf{Z}_1', \ldots, \mathbf{Z}_p') = \boldsymbol{\Gamma} \otimes \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma} & \gamma_{12}\boldsymbol{\Sigma} & \cdots & \gamma_{1p}\boldsymbol{\Sigma} \\ \gamma_{12}\boldsymbol{\Sigma} & \boldsymbol{\Sigma} & \cdots & \gamma_{2p}\boldsymbol{\Sigma} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{1p}\boldsymbol{\Sigma} & \gamma_{2p}\boldsymbol{\Sigma} & \cdots & \boldsymbol{\Sigma} \end{bmatrix}, \quad (4)$$

where $\mathbf{Z}_j' = (Z_{j0}, \ldots, Z_{jT})$, so that $(\mathbf{Z}_1', \ldots, \mathbf{Z}_p') = (Z_{10}, \ldots, Z_{1T}, \ldots, Z_{p0}, \ldots, Z_{pT})$. As noted above following (3), $(\mathbf{Z}_1', \ldots, \mathbf{Z}_p')$ may be generated using the Cholesky decomposition: generate a $p(T+1)$-vector $\boldsymbol{W}$ of independent $N(0, 1)$ variates, decompose $\boldsymbol{B} = \boldsymbol{\Gamma} \otimes \boldsymbol{\Sigma}$ as $\boldsymbol{B} = \boldsymbol{U}'\boldsymbol{U}$ and let $(\mathbf{Z}_1', \ldots, \mathbf{Z}_p') = \boldsymbol{W}'\boldsymbol{U}$.

However, the special structure of $\boldsymbol{B}$ can be exploited, giving a more efficient algorithm. First, generate p independent vectors $\mathbf{V}_i' = (V_{j0}, \ldots, V_{jT})$ having covariance matrix from (3) using (1) and (2). Then factor $\boldsymbol{\Gamma}$ as $\boldsymbol{\Gamma} = \boldsymbol{U}'\boldsymbol{U}$. Taking

$$\mathbf{Z}_1 = u_{11}\mathbf{V}_1 + \cdots + u_{p1}\mathbf{V}_p$$

$$\vdots$$

$$\mathbf{Z}_p = u_{1p}\mathbf{V}_1 + \cdots + u_{pp}\mathbf{V}_p$$

gives basic variables with the desired covariance structure.

An interesting fact about covariance structure (4) is that the correlation between variables at any time point is identical to the correlation between difference scores; specifically; $Corr(Z_{jt}, Z_{j't}) = Corr(Z_{jt} - Z_{j0}, Z_{j't} - Z_{j'0}) = \gamma_{j,j'}$. This fact can be helpful when deciding values of the $\gamma_{j,j'}$, because it does not matter whether the raw variables or difference scores are considered.

4. MEAN STRUCTURES

Our recommendation is to specify, as inputs, mean structures for the different endpoint*timepoint*treatment combinations. Such structures can be determined purely *a priori*, from earlier phase data, suggested by PK/PD models, or from studies on similar interventions. If little prior knowledge is available, there are a few ways to obtain helpful information. One of these is the allometric approach (Mahmood, 1999) that through interspecies scaling is used to predict pharmacokinetic parameters in humans from animal data. The method is based on a power function (not a statistical power function), where the body weight of the species is plotted against the pharmacokinetic parameters of interest. Another approach is the use of information obtained from toxicology studies. The information from these studies will help in determining, for example, the NOEL (no observable effects limit), ADI (acceptable daily intake), and MDD (maximum daily dose). Another useful tool is the pharmaceutical therapeutic index (TI), through which safety and effectiveness can be predicted. The therapeutic index is the ratio of "median lethal dose" over "median effective dose." A high TI value indicates an effective treatment at low doses and a lethal effect at higher doses. A low TI value indicates an ineffective treatment at low doses with lethality at the same low levels. The objective of using the therapeutic index is to have a pharmaceutical that is highly effective at low doses with minimal toxic effects.

By whatever method, we assume that users can assign hypothetical mean time-response profiles $\mu_{j1}^{(g)}, \ldots, \mu_{jT}^{(g)}$ for treatment g $(g = 0, \ldots, G)$, and endpoint j. To simplify the burden of assigning $(G+1) \times p \times (T+1)$ distinct values $\mu_{jt}^{(g)}$, one might assume linear, quadratic or other time-response functions; an alternative is to assign $\mu_{jt}^{(g)}$ for a few specific $t \in \{0, \ldots, T\}$, and linearly interpolate to specify the remaining $\mu_{jt}^{(g)}$. Because time-response profiles tend to be most curved at the

beginning of the study, it is best to assign more values $\mu_{jt}^{(g)}$ for small t; fewer values are needed for large t where the patients have (typically) reached a steadier state.

Bayesian predictive data generation can be accommodated by generating the $\mu_{jt}^{(g)}$ from probability distributions rather than assigning the values specifically. We do not pursue that angle here, but instead suggest sensitivity analysis using alternative configurations $\mu_{jt}^{(g)}$ that are considered probable, *a priori*.

5. DISTRIBUTION FILTERS

Clinical data can be non-normal: adverse event data are binary, self-reported and physician-reported patient conditions are often reported in five-point and seven-point Likert scales, laboratory data contain outliers, and primary endpoints are often censored time-to-event data (such as time to cure or time to death). In this section we show how to generate an ideal data set having desired distributional characteristics and correlation structures; the next section will consider the more realistic situation where there are patient dropouts and noncompliance effects.

5.1. Normal

Given that the basic random elements Z_{ijt} are normal, this is the simplest case. Standard deviations $\sigma_{jt}^{(g)}$ must be specified in addition to the means, and the data set contains elements

$$Y_{ijt}^{(g)} = \mu_{jt}^{(g)} + \sigma_{jt}^{(g)} Z_{ijt}. \tag{5}$$

5.1.1. Input Issues. Homoscedasticity assumptions simplify inputs; in some cases it may be reasonable to assume $\sigma_{jt}^{(g)} = \sigma_{jt}$; or even more simply that $\sigma_{jt}^{(g)} = \sigma_j$.

5.2. Lognormal

For positive, positively skewed clinical data, the lognormal distribution often is realistic. Here we assume simply

$$\ln\left(Y_{ijt}^{(g)}\right) = \mu_{jt}^{(g)} + \sigma_{jt}^{(g)} Z_{ijt}.$$

5.2.1. Input Issues. The $\mu_{jt}^{(g)}$ may be specified as means of the logged data. If it is more convenient to think in terms of the unlogged data, then the $\mu_{jt}^{(g)}$ may be specified as $\mu_{jt}^{(g)} = \ln\{\text{median}(Y_{ijt}^{(g)})\}$ (since the median is invariant to monotonic transformation).

Again homoscedasticity assumptions simplify inputs; in some cases it may be reasonable to assume $\sigma_{jt}^{(g)} = \sigma_{jt}$; or even more simply that $\sigma_{jt}^{(g)} = \sigma_j$, assuming that it is convenient to specify these quantities in terms of the logged data. However, even assuming $\sigma_{jt}^{(g)} = \sigma_j$ the raw data $Y_{ijt}^{(g)}$ are heteroscedastic across time and group, depending on the $\mu_{jt}^{(g)}$. If it is more convenient to think in terms of unlogged data,

then the $\sigma_{jt}^{(g)}$ may be specified in terms of $\mu_{jt}^{(g)}$ and the standard deviation $\sigma_{jt}^{\prime(g)}$ of the unlogged $Y_{ijt}^{(g)}$ by solving

$$\left\{\sigma_{jt}^{\prime(g)}\right\}^2 = \exp\left\{2\mu_{jt}^{(g)} + 2\sigma_{jt}^{(g)}\right\} - \exp\left\{2\mu_{jt}^{(g)} + \sigma_{jt}^{(g)}\right\}$$

for $\sigma_{jt}^{(g)}$.

The correlation inputs ρ and θ in the repeated measures covariance matrix Σ may be specified in terms of the logged data. If it is more convenient to consider unlogged data, one may use the identity

$$\mathrm{Corr}\left(Y_{ijt}^{(g)}, Y_{ijt'}^{(g)}\right) = \frac{\exp\left\{\sigma_{jt}^{(g)}\sigma_{jt'}^{(g)} \times \mathrm{Corr}(Z_{ijt}, Z_{ijt'})\right\} - 1}{\left\{\exp(\sigma_{jt}^{(g)}) - 1\right\}^{1/2}\left\{\exp(\sigma_{jt'}^{(g)}) - 1\right\}^{1/2}} \tag{6}$$

to help identify values of ρ and θ that are reasonably consistent with the values of $\mathrm{Corr}(Z_{ijt}, Z_{ijt'})$ that are implied by (6). However, when attempting to identify input values for the endpoint correlation matrix Γ, one cannot use (6) in general, as some endpoints might be normal, others lognormal, others ordinal, etc. Thus, it is best to identify inputs for lognormal case in terms of the logged data, rather than to use (6). On the other hand, (6) can be helpful as a diagnostic check.

5.3. Mixture

Although the lognormal distribution allows positive outliers, clinical data often contain outliers in both directions and can be negative; a common example of both is percentage change data. A general mixture random variable is easily generated as $X = I(U \geq c)X_1 + I(U < c)X_2$, where $X_i \sim F_i$, U is a $U(0, 1)$ random variable independent of X_i, and where c is the contamination fraction.

Suppose the values $E(Y_{ijt}^{(g)}) = \mu_{jt}^{(g)}$ and $Var(Y_{ijt}^{(g)}) = \{\sigma_{jt}^{(g)}\}^2$ are given as inputs. To simplify inputs and to facilitate consistency across the various distribution filters, one might specify normal mixing distributions with common means. This is accomplished easily: consider (5) and make the following substitution:

$$Z_{ijt} \leftarrow I(U \geq c_j)\frac{Z_{ijt}}{(1 - c_j + c_j r_j^2)^{1/2}} + I\left(U < c_j\right)\frac{r_j Z_{ijt}}{\left(1 - c_j + c_j r_j^2\right)^{1/2}}, \tag{7}$$

where c_j is the contamination fraction for endpoint j and r_j is the ratio of contaminated to normal standard deviation. The resulting Z_{ijt} remain unit variance, so the desired means and variances of $Y_{ijt}^{(g)}$ are achieved.

5.3.1. Input Issues. The mixing fraction c_j and the standard deviation ratio r_j might be estimated from historical data (e.g., using maximum likelihood, care here is needed, as noted by Day, 1969), or simply assigned. In the latter case, the relationship between excess kurtosis κ_j and (c_j, r_j) is helpful:

$$\kappa_j = \frac{3c_j(1 - c_j)(r_j - 1)^2}{(1 - c_j + c_j r_j^2)^2}.$$

For example, one might suspect a certain contamination fraction c_j (e.g., .05), and know the approximate kurtosis from historical data (e.g., 20), in which case the standard deviation ratio r_j can be set to 5.48.

Interestingly, the correlations between the repeated measures $Y_{ijt}^{(g)}$ obtained after applying (7), and then (5), within endpoint j are identical to the corresponding correlations between the basic quantities Z_{ijt}. Hence, for determining ρ and θ to generate the basic normal quantities Z_{ijt}, one may use raw data, despite its nonnormal characteristics. Correlations between mixture endpoints and other types of endpoints are attenuated, however. To use historical data to estimate or suggest correlations among the basic normal quantities Z_{ijt}, it is therefore advisable to transform such historical data to normal scores before estimation.

5.4. Binary

Binary data abound in clinical trials: adverse events are typically binary, and efficacy measures such as cure/no cure are binary as well. The binary outcomes may be generated from a multivariate probit model in terms of the basic variables Z_{ijt}:

$$Y_{ijt}^{(g)} = I(Z_{ijt} > d_{jt}^{(g)}),$$

where $d_{jt}^{(g)} = \Phi^{-1}(1 - \mu_{jt}^{(g)})$, and where $\mu_{jt}^{(g)}$ is the desired probability of success.

5.4.1. Input Issues. The $\mu_{jt}^{(g)}$ may be instantiated using PK/PD models, earlier phase studies, or studies on similar compounds. With binary data there is no need to specify standard deviations $\sigma_{jt}^{(g)}$.

When variables are binary, the correlations in (3) and (4) refer to tetrachoric correlations in the case of multiple binary variables, or to biserial correlations when some variables are binary and others are normal. Correlations involving the raw binary variables should not be used when specifying parameters in Γ and Σ; such correlations are generally too small.

5.5. Ordinal

Ordinal data are common in clinical data sets; such data arise from patient-reported pain scales and quality-of-life (QOL) surveys, as well as from Physician's Global Assessments. Such data are typically recorded as 5-point or 7-point Likert scales. To simulate data from such a process, the multivariate probit model is extended from 0, 1 to several categories, say $1, \ldots, k$.

Suppose the mean structure $E(Y_{ijt}^{(g)}) = m_{jt}^{(g)}$ is desired. Unlike the binary case, there are many probability distributions $P(Y_{ijt}^{(g)} = c)$, $(c = 1, \ldots, k)$, for which $E(Y_{ijt}^{(g)}) = \sum_c cP(Y_{ijt}^{(g)} = c) = m_{jt}^{(g)}$. It is unwieldy to specify separate probability distributions $P(Y_{ijt}^{(g)} = c)$ for all timepoints, treatment groups and endpoints, and the process can be simplified by specifying a single baseline distribution, then relating all distributions to the baseline distribution through an ordinal probit model.

To that end let the baseline distribution (assumed common for all treatment groups) be given by $P(Y_{ij0} = c) = p_{jc}$, with $\sum_c p_{jc} = 1$. In the baseline category, the data are generated in terms of the fundamental quantities as

$$Y_{ij0}^{(g)} = 1 + I(Z_{ij0} > d_{j1}) + \cdots + I(Z_{ij0} > d_{j,k-1}),$$

where $d_{jc} = \Phi^{-1}(p_{j1} + \cdots + p_{jc})$. For other timepoint*group combinations, set

$$Y_{ijt}^{(g)} = 1 + I(Z_{ijt} > d_{j1} - \mu_{jt}^{(g)}) + \cdots + I(Z_{ijt} > d_{j,k-1} - \mu_{jt}^{(g)}).$$

Here the $\mu_{jt}^{(g)}$ are determined by solving $E(Y_{ijt}^{(g)}) = m_{jt}^{(g)}$, which implies that they are the solution to

$$k - \sum_{c=1}^{k-1} \Phi(d_{jc} - \mu_{jt}^{(g)}) = m_{jt}^{(g)}. \tag{8}$$

Software to solve nonlinear equations is required for (8); nonlinear regression software such as PROC NLIN will work if more sophisticated tools are unavailable.

5.5.1. Input Issues. The $m_{jt}^{(g)}$ are suggested, as before, from early phase studies, PK/PD models, or studies on similar compounds. The baseline distributions $P(Y_{ij0} = c) = p_{jc}$ are specified similarly.

Correlations in (3) and (4) refer to polychoric correlations in the case of multiple ordinal variables, or to ordinal extensions of biserial correlation when some variables are ordinal and others are normal. Again, correlations involving the raw ordinal variables should not be used when specifying parameters in Γ and Σ; such correlations are generally too small.

5.6. Survival

As seen above, the simulation model we present revolves around the basic patient quantities Z_{ijt}, which may be thought of as "latent health indicators." This paradigm generalizes easily to a certain kind of survival model known as the "first hitting time model" (Lee and Whitmore, 2004). As in the case of binary variables, thresholds $d_{jt}^{(g)}$, $t = 1, \ldots, T$, must be assigned. Survival times $Y_{ij}^{(g)}$ and censoring indicators $C_{ij}^{(g)}$ are then given as

$$Y_{ij}^{(g)} = \min\{t : Z_{ijt} > d_{jt}^{(g)}\} \quad \text{and} \quad C_{ij}^{(g)} = 0, \quad \text{when such a } t \text{ exists} \tag{9}$$

$$Y_{ij}^{(g)} = t \quad \text{and} \quad C_{ij}^{(g)} = 1 \text{ otherwise} \tag{10}$$

Such a model is realistic when Z_{ijt} can be thought of as "progression of disease" or "progression of cure." Note that an equivalent form of the model has $Y_{ijt}^{(g)} = \min\{t : Z_{ijt} - d_{jt}^{(g)} > 0\}$ so that the quantities $Z_{ijt} - d_{jt}^{(g)}$ may be thought of as group-specific "progression," with 0 denoting the threshold that determines a survival "event."

5.6.1. Input Issues. In cases where "progression" $Z_{ijt} - d_{jt}^{(g)}$ is manifest, the inputs $d_{jt}^{(g)}$ can be determined from historical data, PK/PD models, or studies on similar compounds. In such a case the standard deviation of the progression measurement must be "absorbed" into the $d_{jt}^{(g)}$. When progression is not manifest, it may be convenient to specify survival probabilities $S_{jt}^{(g)} = P(Y_{ij}^{(g)} > t)$, and then determine the $d_{jt}^{(g)}$ as a function of the $S_{jt}^{(g)}$ using (9) by solving $P(\min\{t : Z_{ijt} > d_{jt}^{(g)}\} > t) = S_{jt}^{(g)}$. Unfortunately, this equation seems to defy simple solution for general $d_{jt}^{(g)}$ and the correlation structure (3), although the relationship has been established for certain special cases as noted by Aalen and Gjessing (2001). We suggest modeling by picking the $d_{jt}^{(g)}$ values, checking that the implied survival functions $S_{jt}^{(g)}$ are reasonable, then adjusting the $d_{jt}^{(g)}$ as needed. The existing system we describe estimates the $S_{jt}^{(g)}$ using a quick simulation so that the user can modify the $d_{jt}^{(g)}$ interactively.

As far as correlation inputs go, the values in (3) can be suggested by historical progression data when such is available. Otherwise, the inputs (3) are not much of a concern, as there is only one survival outcome $Y_{ij}^{(g)}$, not repeated measures indexed by t. However, the relationship between $d_{jt}^{(g)}$ and $S_{jt}^{(g)}$ depends on the values chosen in (3), so again, the implied values of $S_{jt}^{(g)}$ should be examined. Note that the $S_{jt}^{(g)}$ values might be determined from historical analyses using Cox proportional hazards models.

Correlations between the latent "progression" and other endpoints (as given in (4)) can possibly be estimated from maximum likelihood analysis of data using the "first hitting time model," or simply assigned *a priori*.

6. DROPOUTS AND NONCOMPLIANCE

The mechanisms in Section 5 create an "ideal data set." However, in the real world of clinical trials, data sets are rarely ideal: patients skip visits, they drop out of the study altogether, and they do not comply with their assigned regimen, by failing to take doses, or by mixing their medication with other medications and/or nutraceuticals. Thus the actual data $D_{ijt}^{(g)}$ differ from the ideal data $Y_{ijt}^{(g)}$ that are defined above. The following sections describe ways to simulate actual, realistic data. Inputs to the simulation system may include historical data or assumed dropout and noncompliance mechanisms; sensitivity analysis using a variety of plausible mechanisms is recommended in either case.

6.1. Dropouts and Missing Data

The dropout and missing data mechanisms are complex, and generally cannot be assumed to be missing at random (MAR) or missing completely at random (MCAR). More realistically, the missing values should depend on patient experience in the trial. The following models simultaneously aim for realism and simplicity, using the framework of the previous sections.

First we make a distinction: a "dropout" refers to a patient's discontinuation in the trial. "Missing data" refer to a patient's missing of a visit. In some cases (e.g., for $t = T$) it is impossible to distinguish the two simply from the missing value

in the clinical record; nevertheless, it makes sense to model the data generating mechanisms as distinct.

6.1.1. Missing Data. A simplifying assumption is that the missing values (not the dropouts) are MCAR, reflecting a sporadic tendency to miss visits that is consistent across all treatment groups. If $U_{ijt}^{(g)} \sim U(0, 1)$ are generated independently of the $Y_{ijt}^{(g)}$, and if the missing value rate is β, then assign $D_{ijt}^{(g)}$ to missing when $U_{ijt}^{(g)} < \beta$, else $D_{ijt}^{(g)} = Y_{ijt}^{(g)}$.

6.1.2. Dropouts. The dropout mechanism we assume is specifically not MAR. In some cases patient dropouts are sporadic, with MCAR mechanism; in other cases dropout may be related to lack of safety and/or efficacy of the patient's experience. There are several possible ways to model the dropout mechanism; some examples and further references are contained in O'Brien et al. (2005). We have found it convenient to define a "misery index" in terms of the basic quantities $Z_{ijt}^{(g)}$, allowing group-specific dropout rates that are defined by quantiles of the distribution of the misery index. The misery index is a combination of efficacy and safety: if both efficacy and safety are very bad, then the patient drops out. On the other hand, if safety is somewhat bad but efficacy is very good, the patient might stay; conversely, if efficacy is somewhat bad but safety is good the patient might stay. This model is illustrated in Figure 3:

To be specific, define {Safety} to be the set of safety endpoints, and {Efficacy} to be the set of efficacy endpoints. Then, for $j \in$ {Safety} let $c_j = -1(1)$ if lower(higher) values of $Z_{ijt}^{(g)}$ indicate greater safety; for $j \in$ {Efficacy} let $c_j = -1(1)$ if lower(higher) values of $Z_{ijt}^{(g)}$ indicate greater efficacy. Define

$$S_{it}^{(g)} = \sum_{j \in \{\text{Safety}\}} c_j Z_{ijt}^{(g)}, \quad E_{it}^{(g)} = \sum_{j \in \{\text{Efficacy}\}} c_j Z_{ijt}^{(g)},$$

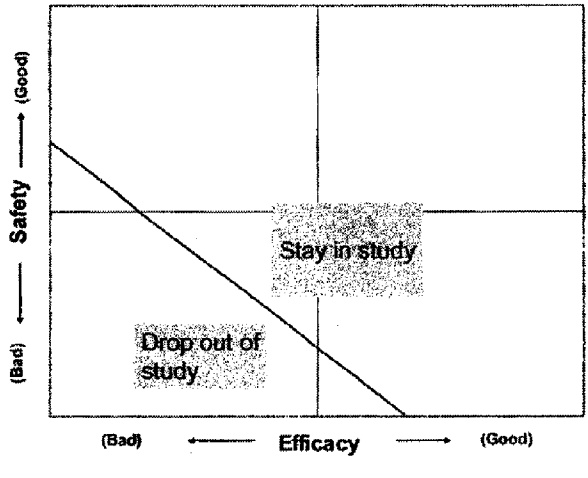


Figure 3 Dropout mechanism.

the combined safety and efficacy scores respectively. The dropout mechanism shown in Figure 3 is related to these quantities, but we assume further that the effects may be cumulative, and that the cumulative effect may be more local (depending more upon the patient's recent experience) or global (depending more on the patient's experience through the entire trial history). Thus we define an exponential smooth parameter $s \in [0, 1]$, where $s = 1$, denotes a completely local effect, $s = 0$ denotes a cumulative effect, and values intermediate denote intermediate cumulative effect of misery. Our "misery index" that allows differential weight $w \in [0, 1]$ to safety and efficacy is defined as follows:

$$I_{it}^{(g)} = w\frac{S_{it}^{(g)}}{Var^{1/2}(S_{it}^{(g)})} + (1-w)\frac{E_{it}^{(g)}}{Var^{1/2}(E_{it}^{(g)})},$$

$$I_{it}^{(g)} \leftarrow I_{it}^{(g)} + (1-s)I_{i,t-1}^{(g)} + (1-s)^2 I_{i,t-2}^{(g)} + \cdots,$$

to allow exponentially smoothed carryover effects of the misery index, and finally we have our index

$$I_{it}^{(g)} \leftarrow \frac{I_{it}^{(g)}}{Var^{1/2}(I_{it}^{(g)})},$$

a $N(0, 1)$ quantity. Patient i is assigned to leave at time t (and subsequently) if $I_{it}^{(g)} > \Phi^{-1}(1 - r_t^{(g)})$, where $r_t^{(g)}$ is the group-specific dropout rate.

If variables are designated as survival (see Section 5.6), the described system does not censor the data as a result. If a survival variable (such as death) censors observations, this is handled in the analysis stage; irrelevant data generated by the system can be ignored.

6.2. Noncompliance

Realistically, patients will not adhere to their regimens. Our model assumes that, within each between-visit interval (indexed by t), a patient's compliance $C_{it}^{(g)}$ is a continuous measurement on the $[0, 1]$ scale. If compliance is meant to refer to actual percentage of doses taken during the interval, then true compliance is actually discrete, taking values such as 14/14, 13/14, etc., and the value $C_{it}^{(g)}$ may be considered as a continuous approximation to this discrete value. On the other hand, "true compliance" might more actually reflect a variety of behaviors including proper timing of doses, and compliance with food and additional medication requirements. In such case the use of a continuous distribution for $C_{it}^{(g)}$ has better justification.

It is sensible that there are subject-specific and carryover effects for the compliance data. Thus, the same CS + AR(1) model of (3) might therefore be used as above, by first generating CS + AR(1) normally distributed fundamental quantities Z_{ict} (i denotes patient, c denotes compliance endpoint, t denotes time interval), and then convert to proportions through the normal probability transform. Although the Z_{ict} can be thought to be dependent on the endpoints, it simplifies inputs to make them independent. Dependence of the patient responses upon compliance can then be modeled directly in terms of effects of noncompliance

on the patient-specific mean response, which is more consistent with PK/PD models (Holford and Peace, 1992; Lee et al., 2003).

The $C_{it}^{(g)}$ are determined in terms of the Z_{ict} using group-specific target compliance rates. Assuming compliance is stationary over time, the median compliance at any time might be prespecified as $C_{.5}^{(g)}$, e.g., $C_{.5}^{(0)} = 0.95$ and $C_{.5}^{(1)} = 0.90$ specifies median compliances of 95% and 90% in the control and dosed groups, respectively. An additional constraint in the form of another percentile is needed; let us suppose the tenth percentile is specified. Then the $C_{it}^{(g)}$ are determined as

$$C_{it}^{(g)} = \Phi(a^{(g)} + b^{(g)} Z_{ict}),$$

where

$$a^{(g)} = \Phi^{-1}(C_{.5}^{(g)})$$

and

$$b^{(g)} = \frac{\Phi^{-1}(C_{.1}^{(g)}) - \Phi^{-1}(C_{.5}^{(g)})}{\Phi^{-1}(C_{.1}^{(g)})}.$$

6.2.1. Effect of Noncompliance on Treatment Response. Motivated by models used by Holford and Peace (1992), and by Lee et al. (2003) respectively, we assume that noncompliance regresses the patient response toward natural history or placebo, depending on user preference. No matter which distribution structure is assumed, the parameters $\mu_{jt}^{(g)}$ are closely tied to drug or placebo effects. We assume noncompliance alters these values by "regressing" them towards baseline values $\mu_{jt}^{(b)}$ as shown in the following replacement operation:

$$\mu_{jt}^{(g)} \leftarrow \mu_{jt}^{(b)} + CE_{it}^{(g)}(\mu_{jt}^{(g)} - \mu_{jt}^{(b)}),$$

where $CE_{it}^{(g)}$ denotes "compliance effect," which like compliance, is between 0 and 1, and is explained further below. Note that this is just a mean effect: individual patient outcomes do not necessarily regress since the model assumes there is patient-specific error.

In some cases it is sensible to take $\mu_{jt}^{(b)} = \mu_{jt}^{(0)}$; i.e., noncompliance regresses effects towards the placebo response (Lee et al., 2003). In other cases it may be sensible to input the $\mu_{jt}^{(b)}$ as "natural history" values, and assume that noncompliance regresses all treatment groups (and perhaps even the placebo group) toward the natural history (Holford and Peace, 1992).

The compliance effect $CE_{it}^{(g)}$ is assumed to be related to actual compliance depending on the study – some studies model compliance effects as cumulative, suggesting

$$CE_{it}^{(g)} = (C_{i1}^{(g)} + \cdots + C_{it}^{(g)})/t.$$

In other cases, such as with antibacterials, the effect of compliance might be purely local:

$$CE_{it}^{(g)} = C_{it}^{(g)}.$$

In general, one might suppose both local and global effects via an exponential smooth specified as follows:

$$CE_{i1}^{(g)} = C_{i1}^{(g)}$$

$$\cdots$$

$$CE_{it}^{(g)} = C_{it}^{(g)} + (1-s)CE_{i,t-1}^{(g)}$$

Converting to averages so that all compliance effects are between 0 and 1, we perform the following substitution:

$$CE_{it}^{(g)} \leftarrow \frac{s\{CE_{it}^{(g)}\}}{1-(1-s)^t}, \quad \text{for } s \neq 0$$

$$CE_{it}^{(g)} \leftarrow \frac{CE_{it}^{(g)}}{t}, \quad \text{for } s = 0$$

6.2.2. Correlation with Misery Index. The local compliance values logically might be related to the misery index. If the correlation desired between local compliance and cumulative misery index is $\tau^{1/2}$, then the local compliance may be defined via the following substitution:

$$C_{it}^{(g)} \leftarrow \tau^{1/2} I_{it}^{(g)} + (1-\tau)^{1/2} C_{it}^{(g)}.$$

7. AN EXAMPLE

Gatekeeping strategies allow both determination of clinical success and determination of a collection of secondary indications, all with strong control of the familywise error rate (or FWE; see Westfall and Krishen, 2001; Wiens and Dmitrienko, 2005). The complex multivariate nature of these procedures makes mathematical determination of their operating characteristics (true FWE and various power functions) difficult. Adding issues of noncompliance, nonnormality (including survival and other types of endpoints), and informative dropouts, makes the problem completely intractable in finite samples.

We consider a hypothetical trial with two arms, a primary endpoint that normal with slight positive kurtosis, and a collection of three equally important secondary endpoints, two of which are measured on 5-point Likert scales, and one of which is a survival endpoint. We suppose that the (latent) endpoint correlations, subject correlation, and AR(1) correlation are all .5; as well as cumulative noncompliance that regresses the response toward the placebo. Compliance is assumed to be good, with median and 10th percentiles equal to 95% and 80% respectively in both groups. We assume there are five binary safety variables, and that the dropout mechanism is equally weighted by safety and efficacy, where the

Table 1 Simulated power for primary endpoints

Design	AOV	ANCOV (Mean)	ANCOV (Med)	Diff (Mean)	Diff (Med)	KW	KWdiff (Mean)	KWdiff (Med)
12 wks/30,30	0.41	0.55	0.54	0.48	0.47	0.58	0.67	0.65
12 wks/50,50	0.57	0.73	0.72	0.67	0.64	0.80	0.87	0.86
12 wks/100,100	0.83	0.94	0.93	0.90	0.89	0.97	0.99	0.99
8 wks/30,30	0.36	0.49	0.48	0.43	0.41	0.51	0.59	0.57
8 wks/50,50	0.51	0.67	0.66	0.59	0.58	0.73	0.82	0.80
8 wks/100,100	0.78	0.90	0.90	0.86	0.84	0.95	0.98	0.98

safety is determined by averaging all safety measures and efficacy is determined by averaging all efficacy measures. The dropout rates are low; 1% and 1.5% at each timepoint as determined by the misery index in the control and treatment groups, respectively and that the random missing value rate is 1%. Missing values will be imputed using last observation carried forward (LOCF). Other imputation methods may be more appropriate, but are less well-studied, making CTS especially useful to evaluate their operating characteristics.

The study is considered for either 12 weeks or 8 weeks; part of the analysis will be to determine pros and cons of each. This example is a reflection of a project where the toxicological coverage was 12 weeks, and it is important to find out if 8 weeks of treatment produces the same efficacy as the 12 weeks. If 8 weeks is enough with regard to efficacy that would have been great, since the compound potentially could cause side effect if used for longer than 8 weeks.

Mean response functions are assumed to increase rapidly, with 90% of the efficacy at 8 weeks and levelling off at 12 weeks for all efficacy measures. Only one of the safety measures is assumed to have a drug-related effect; all others are assumed null. Interesting questions are (a) What is the best test for the primary endpoint – The two sample t-test on the raw scores ("AOV" in Table 1 below)? The two-sample t-test on the changes from baseline average, or from baseline median ("Diff mean"; Diff median")? The Kruskal–Wallis test on the raw scores (KW)? The Kruskal–Wallis test on the change scores, mean or median ("KWdiff mean"; "KWdiff median")? Covariate-adjusted comparisons, using baseline mean or median ("ANCOV mean"; ANCOV median")? and (b) Among studies where the primary endpoint has been found significant by the best test found in the answer to (a), how many secondary endpoints will be found significant using the proportional hazards model for the survival endpoint and Kruskal–Wallis tests for the ordinal scores? The gatekeeping strategy chosen here uses Hochberg's procedure for all secondary endpoints. Hochberg's method is particularly attractive for its simplicity: because it requires only the p-values for the marginal tests, it can be used easily for endpoints with mixed data types, such as survival and metric. The sequential nature of the Hochberg procedure as well as the conditional nature of the gatekeeping structure make this problem virtually intractable mathematically.

Table 1 summarizes the results of various tests on the primary endpoint for a variety of designs.

Clearly, the power of the test is very sensitive to the primary analysis.

Table 2 shows the results for the secondary endpoints, using Hochberg's method, among cases where the KW difference test (from the mean) is significant.

Table 2 Simulated power for individual secondaries, for any secondary, and for all secondaries

Design	Sec 1	Sec 2	Sec 3	Any	All
12 wks/30,30	0.23	0.54	0.54	0.69	0.18
12 wks/50,50	0.49	0.78	0.78	0.89	0.42
12 wks/100,100	0.83	0.98	0.98	0.99	0.82
8 wks/30,30	0.04	0.40	0.40	0.54	0.03
8 wks/50,50	0.16	0.62	0.61	0.75	0.14
8 wks/100,100	0.41	0.91	0.91	0.97	0.39

Thus, the 8 weeks/50,50 design offers acceptable power (82%) for the primary endpoint for the KWdiff test. If it is acceptable to get at least one significance among the secondaries following a primary significance, then the 8 weeks/50,50 design remains possibly acceptable, with 75% conditional power. However, if difference among all secondaries is desired, then the 12 weeks/100,100 design is needed. It is worth noting that the Survival test (secondary endpoint 1) behaves poorly for small n, this is a failure of asymptotics of the usual proportional hazards model test. The simulation tool suggests that an alternative test should be sought if smaller sample sizes are to be used.

There were 20,000 simulations per trial configuration to reduce Monte Carlo error; this was accomplished in reasonable time by using the grid computing environment of SAS (Bremer et al., 2004) with 40 machines; total compute time for the three designs was ~1 hour using the grid but would have been ~35 hours on a stand-alone machine.

8. CONCLUSION

We have presented details concerning a system for simulation of typical clinical trials that can help to answer fundamental, yet complex, questions that are often mathematically intractable. While the system offers a rich flexible structure, simplifying modeling assumptions are made concerning data generation processes, including various homogeneity assumptions and compliance and dropout modeling assumptions, that enhance its usability but simultaneously limit its generality. Our hope is that pharmaceutical statisticians find the system information presented here to be not only useful in its own right, but also a useful reference point for developing other models for use in improved systems. The system is currently the proprietary property of Vertex Pharmaceuticals, Inc.

REFERENCES

Aalen, O. O., Gjessing, H. K. (2001). Understanding the shape of the hazard rate: A process point of view. *Statistical Science* 16:1–22.

Anderson, J. J., Bolognese, J. A., Felson, D. T. (2003). Comparison of rheumatoid arthritis clinical trial outcome measures. *Arthritis and Rheumatism* 48:3031–3038.

Bremer, R., Perez, J., Smith, P., Westfall, P. (2004). Grid computing at Texas Tech University using SAS. Proceedings of the 14th Annual South-Central SAS User's Group Regional Conference, 64–72.

Day, N. E. (1969). Estimating the components of a mixture of normal distributions. *Biometrika* 56:463–474.

Frison, L., Pocock, S. J. (1992). Repeated measures in clinical trials: Analysis using mean summary statistics and its implications for design. *Statistics in Medicine* 11:1685–1704.

Gao, F., Thompson, P., Xiong, C., Miller, J. P. (2006). Analyzing multivariate longitudinal data Using SAS©. Paper 187-31 of Proceedings of the Thirty-first Annual SAS® Users Group International Conference, pp. 1–9.

Gieschke, R., Steimer, J.-L. (2000). Pharmacometrics: modeling and simulation tools to improve decision making in clinical dug development. *European Journal of Drug Metabolism and Pharmacokinetics* 25(1):49–58.

Hale, M. D. (1997). Using population pharmacokinetics for planning a randomized concentration-controlled trial with binary response. In: Aarons, L., Balant, L. P., Danhof, M., eds. *European Cooperation in the Field of Scientific and Technical Research.* Geneva: European Communites, pp. 227–235.

Hale, M., Gillespie, D. W., Gupta, S. K., Holford, N. H. (1996). Clinical trial simulation streamlining your drug development process. *Applied Clinical Trials* 15:35–40.

Hale, M. D., Nicholls, A. J., Bullingham, R. E. S., Hene, R., Hoitsman, A., Squifflet, J. P., Weimar, W., Vanrenterghem, Y., Van de Woude, F. J., Verpooten, G. A. (1998). The pharmacokinetic-pharmacodynamic relationship for mycophenolate mofetil in renal transplantation. *Clinical Pharmacology and Therapeutics* 64:672–683.

Holford, N. H. G., Peace, K. E. (1992). Methodologic aspects of a population pharmacodynamic model for cognitive effects in Alzheimer patients treated with Tacrine. *Proceedings of the National Academy of Sciences of the United States of America* 89:11466–11470.

Holford, N., Kimko, H., Montelone, J., Peck, C. (2000). Simulation of clinical trials. *Annual Review of Pharmacology and Toxicology* 40:209–234.

Kimko, H. C., Duffull, S. B. (2003). *Simulation for Designing Clinical Trials: A Pharmacokinetic–Pharmacodynamic Modeling Perspective.* New York: Marcel Dekker Inc.

Lee, M.-L. T., Whitmore, G. A. (2004). First hitting time models for lifetime data. In: Balakrishnan, N., Rao, C., eds. *Handbook of Statistics 23, Advances in Survival Analysis.* Elsevier: Amsterdam, pp. 537–543.

Lee, H., Kimko, H. C., Rogge, M., Wang, D., Nestorov, I., Peck, C. C. (2003). Population pharmacokinetic and pharmacodynamic modeling of Etanercept using logistic regression analysis. *Clinical Pharmacology and Therapeutics* 73:348–365.

Machado, S. C., Miller, R., Hu, C. (1999). A regulatory perspective on pharmacokinetic/ pharmacodynamic modeling. *Statistical Methods in Medical Research* 8:217–245.

Mahmood, I. (1999). Allometric issues in drug development. *Journal of Pharmaceutical Sciences* 88:1101–1106.

Maxwell, C., Domenet, J. G., Joyce, C. R. R. (1971). Instant experience in clinical trials: a novel aid to teaching by simulation. *Journal of Clinical Pharmacology* 11(5):323–331.

O'Brien, P. C., Zhang, D., Bailey, K. R. (2005). Semi-parametric and non-parametric methods for clinical trials with incomplete data. *Statistics in Medicine* 24:341–358.

Peck, C., Desjardins, R. (1996). Simulation of clinical trials – encouragement and cautions. *Applied Clinical Trials* 5:30–32.

Peck, C. C., Rubin, D. B., Sheiner, L. B. (2003). Hypothesis: A single clinical trial plus causal evidence of effectiveness is sufficient for drug approval. *Clinical Pharmacology and Therapeutics* 73:481–490.

Poland, B., Wadda, R. (2001). Combining drug–disease and economic modelling to inform drug development decisions. *Drug Discovery Today* 6:1165–1170.

Reigner, B., William, P., Patel, I., Steimer, J.-L., Peck, C. van Brummelen, P. (1997). An evaluation of the integration of pharmacokinetic and pharmacodynamic principles in clinical drug development: experience within Hoffman La Roche. *Clinical Pharmacokinetics* 33:142–152.

Sheiner, L., Steimer, J.-L. (2000). Pharmacokinetic/pharmacodynamic modeling in drug development. *Annual Review of Pharmacology and Toxicology* 40:67–95.

U.S. Department of Health and Human Services, Food and Drug Administration (1999). Guidance for Industry: Clinical Development Programs for Drugs, Devices, and Biological Products for the Treatment of Rheumatoid Arthritis (RA); http://www.fda.gov/cder/guidance/index.htm.

Westfall, P. H., Krishen, A. (2001). Optimally weighted, fixed sequence, and gatekeeping multiple testing procedures. *Journal of Statistical Planning and Inference* 99:25–40.

Wiens, B. L., Dmitrienko, A. (2005). The fallback procedure for evaluating a single family of hypotheses. *Journal of Biopharmaceutical Statistics* 15:929–942.