Validation by Design®

The Statistical Handbook for Pharmaceutical Process Validation

Lynn Torbeck

PDA
Bethesda, MD, USA
DHI Publishing, LLC
River Grove, IL, USA

www.pda.org/bookstore

CONTENTS

Preface		xv
Introduction		xix
1	**PROCESS VALIDATION: GENERAL PRINCIPLES AND PRACTICES**	**1**
	Background	1
	Self-Audit Questions with Commentaries	2
2	**ANALYSIS OF VARIANCE**	**19**
	Other Names	19
	Acronyms	19
	Definition	19
	Related topics	20
	Calculation	20
	Illustration	20
	Cautions	21
	Advice	21
	Variations	21
	Historical	21
	References	21

www.pda.org/bookstore

3	**AVERAGE**	**23**
	Other Names	23
	Definition	23
	Related Topics	23
	Assumed	24
	Cautions	24
	References	24
4	**CAUSE AND EFFECT RELATIONSHIP**	**25**
	Definition	25
5	**CONFIDENCE, ASSURANCE CONSISTENCY AND SIGNIFICANCE**	**27**
	Other Names	27
	Definitions	27
	Discussion	28
	References	29
6	**CONTROL CHARTS**	**31**
	Other Names	31
	Definition	31
	Question	31
	Related topics	32
	Philosophy	32
	Discussion	32
	Goals	32
	Assumed	32
	Calculation	32
	Comments	33
	Cautions	33
	Implications	34
	Advice	34
	Variations	34
	Historical	35
	References	35
7	**COEFFICIENT OF CORRELATION**	**37**
	Other Names	37
	Acronyms	37
	Definition	37
	Related Topics	38

www.pda.org/bookstore

Advice 39
Assumed 39
Implications 39
Variations 39
Cautions 39
Comments 39
Interpretation 40
Formulas 40
References 40

8 Cpk 41
Other names 41
Definition 41
Related Topics 41
Question 41
Assumed 42
Discussion 42
Controversy 42
Cautions 43
Observation 43
Advice 43
In the language of mathematics 44
Variations 45
References 45

9 Cpk CONFIDENCE INTERVALS 47
Definition 47
Related Topics 47
Philosophy 47
Interpretation 47
Cautions 48
Implications 48
Advice 49
In Math 49
References 49

10 DATA 51
Other Names 51
Related Topics 51
Qualitative Analysis 51
Quantitative Measurement 52

www.pda.org/bookstore

	Distributions of Data	53
	References	54
11	**DATA ANALYSIS**	**55**
	Other Names	55
	Philosophy	55
	Data Presentations	55
	Data Summaries	56
	Data Inference	56
	Data Decisions	56
12	**DATA COLLECTION**	**57**
	Good Data Collection	57
	Poor Data Collection	58
	References	59
13	**DEFINING QUALITY**	**61**
	Defining Quality for the Pharmaceutical Industry	62
	References	63
14	**DESIGNED EXPERIMENTS**	**65**
	Other Names	65
	Acronyms	65
	Definitions	65
	Question	66
	Discussion	66
	Observation	67
	Comments	67
	Implications	67
	Variations	67
	Historical	67
	References	68
15	**HISTOGRAMS**	**69**
	Other Names	69
	Definition	69
	Question	69
	Assumptions	69
	Related Topics	70
	Comments	71

www.pda.org/bookstore

Cautions 71
Variations 71
Advice 71
Reference 71

16 INTERQUARTILE RANGE 73
Other Names 73
Acronyms 73
Definitions 73
Related Topics 73
Recommendations 74
Implications 74
In Math 74
Variations 76
References 77

17 MEDIAN 79
Other Names 79
Definition 79
Related Topics 79
Implications 80
Recommendations 80
Reference 80

18 MODE 81
Other Names 81
Definitions 81
Related Topics 81
Data Values 82
Data Sets 82
In Math 82
References 82

19 NORMAL PROBABILITY PLOTS 83
Other Names 83
Acronyms 83
Definition 83
Related Topics 83
Assumed 84
Discussion 84
Comments 85

www.pda.org/bookstore

	Variations	85
	Implications	85
	Advice	85
	References	86
20	**OUTLIERS (Barr Case)**	**87**
	Other Names	87
	Definition	87
	Philosophy	87
	Discussion	88
	Interpretation	88
	Observation	88
	Advice	89
	Comments	89
	Cautions	89
	Implications	89
	References	89
21	**PLACKETT-BURMAN DESIGNS**	**91**
	Other Names	91
	Definition	91
	Related Topics	91
	Interpretation	92
	Variations	92
	Generating Designs	93
	Cautions	93
	Observation	93
	Implications	93
	When to Use	94
	In Math	94
	Quotable Quote	94
	References	94
22	**PROCESS CAPABILITY**	**95**
	Definitions	95
	Question	95
	Related Topics	95
	Assumed	96
	Measurement	96
	Calculation	96
	Procedure	96

www.pda.org/bookstore

Implications 97
Advice 99
Variations 99
Quotable Quote 99
References 99

23 PROCESS MAPPING 101
Other Names 101
Related Topics 101
Philosophy 101
Definition 101
Goals 103
Cautions 104
Variations 104
References 104

24 RANGE 105
Other Names 105
Definition 105
Related Topics 105
Comments 106
Cautions 106
Implications 106
In Math 107
References 107

25 %RELATIVE STANDARD DEVIATION 109
Other Names 109
Acronyms 109
Definition 109
Related Topics 109
Assumed 109
Cautions 110
Uses 110
Variations 111
Implications 111
In Math 111
Historical 111
References 111

www.pda.org/bookstore

26	**REPRESENTATIVE SAMPLING**	**113**
	Definition	113
	Related Topics	113
	Assumed	113
	Random Selection Methods	113
	Non-Random Selection Methods	114
	Discussion	114
	Bergum's Method References	118
27	**ROOT CAUSE ANALYSIS**	**119**
	Other Names	119
	Definition	119
	Discussion	119
	References	123
28	**SAMPLE SIZE**	**125**
	Definition	125
	Related Topics	125
	Assumed	125
	Discussion	125
	Cautions	127
	A Very Rough Guide	127
	Implications	128
	References	128
29	**SAMPLES**	**129**
	Definition	129
	Antonym	129
	Related Topics	129
	Goal	130
	Assumed	130
	Implications	130
	References	130
30	**SAMPLING PLANS**	**131**
	Other Names	131
	Definition	131
	Related Topics	131
	Assumed	132
	Discussion	132
	Published Sampling Plans	132

www.pda.org/bookstore

Cautions 132
References 133

31 SCATTER PLOTS 135
Other Names 135
Definition 135
Related Topics 135
Goal 135
Assumed 135
Variations 136
Cautions 137
Implications 137
Advice 137
References 137

32 SHORT RUN SPC 139
Definition 139
Question 139
Related Topics 139
Techniques 140
Advice 140
Historical 141
References 141

33 SQC/SPC 143
Question 143
Acronyms 143
Related Topics 143
Definitions 143
Historical 144
Philosophy 144
References 144

34 STANDARD DEVIATION, STANDARD ERROR VARIANCE 147
Other Names 147
Definitions 147
Related Topics 148
Assumed 148
Cautions 148
Implications 148

www.pda.org/bookstore

In Math 149
References 149

35 STATE OF CONTROL 151
Other Names 151
Definitions 151
Other Forms of Control 152
Related Topics 152
Discussion 152
Observation 153
Measurement 153
Cautions 153
Historical 153
References 153

36 SUMMARY STATISTICS 155
Definition 155
Other Names 155
Philosophy 155
Reference 156

37 TIME PLOTS 157
Other Names 157
Definition 157
Question 157
Related Topics 157
Comments 159
Cautions 159
Variations 159
Advice 159
Quotable Quote 159
Regulatory 160
References 160

38 TOLERANCE INTERVAL — PARAMETRIC (One-sided, Standard Deviation Unknown) 161
Other Names 161
Related Topics 161
Definition 161
Assumed 161
Variations 162

www.pda.org/bookstore

Comments 163
References 163

39 TOLERANCE INTERVALS — PARAMETRIC (Two-sided, Standard Deviation Unknown) 165
Other Names 165
Related Topics 165
Definition 165
Assumed 165
Variations 166
Comments 167
References 167

40 TRENDING DATA 169
Abstract 169
Background 169
Scope of the Topic 170
Philosophical Basis for Trending 171
Changes that can Occur in a Data Series 173
Discussion 173
Time Series and Trending 174

41 VARIABILITY 177
Other Names 177
Definitions 177
Measurement
Philosophy

APPENDIX 181
Guidance for Industry
Process Validation: General Principles and Practice

www.pda.org/bookstore

PREFACE

The need for this book is illustrated by the many inquires and questions that I have received over a number years, and still receive, about how to use and implement basic statistics and designed experiments, DOE, for pharmaceutical process validation. There is clearly confusion and concern about meeting general regulatory requirements in a competitive business setting.

Typical questions include:

- "How big should my sample be?"
- "How many samples should I take?"
- "What sampling plan should I use?"
- "What is a 'high degree of assurance'?"
- "What level of statistical significance should I use?"
- "What is process capability and how do I measure it?"
- "How do I handle aberrant values?"
- "What is a 'state of control' and how do I know when I have it?
- "How do I measure and control variability?"
- "How do I conduct a 'statistical analysis'?"
- "How do I use designed experiments for validation?"

www.pda.bookstore

I gave my first presentation recommending designed experiments for validation in May 1978 at the First Annual Midwest Biopharmaceutical Statistics Workshop held at Ball State University in Muncie IN, in the "Roundtable Discussion of Statistics and GMPs."

In March 1996, I published a paper titled, "Ruggedness and Robustness with Designed Experiments," in *Pharmaceutical Technology* showing how to do assay validation with DOE.

Ron Branning and I wrote a journal article titled "Designed Experiments — A Vital Role in Validation" and published it in *Pharmaceutical Technology*, June 1996. It laid the philosophical basis and argument for using designed experiments for pharmaceutical validation.

For years I presented a training course on introductory designed experiments for GMP validation titled "Validation by Design®."

In 2007, I edited the book *Pharmaceutical and Medical Device Validation by Experimental Design*, published by Informa Health Care.

This then is a continuation of several publications I have written over time.

This book is for those engaged in meeting the requirements of the FDA process validation guidance. I have formatted it such that it is suitable for both those new to using statistics for validation and those tasked with teaching and managing teams implementing the guidance. The book forms a minimum expectation for the degree of implementation by mid and upper management and provides a common language for discussions with regulatory agencies.

I envision the book being used in several ways. First, the questions highlight statistical topics that can be further explored by studying the modules and the references. Second, the questions can be used to write new validation protocols. Third, the questions can be used to audit validation standard operating procedures (SOPs), protocols and reports.

What can this book do for you? I believe that it will address directly some of the more confusing issues about using basic statistics for process and product validation. I believe that I have given direct specific answers to direct specific questions where possible. Where specific answers are not possible, I have discussed the ramifications of the issue. Hopefully this will provide a basis for discussion with other team members and regulatory agencies.

www.pda.bookstore

This book cannot answer non-statistical questions that are in the area of subject matter, basic science, management and regulatory affairs. It cannot answer such questions as:

- "What will the FDA say if I use _____ ."
- "We can't afford to do _____ ."
- "Does the FDA still expect us to do X number of commercial lots at target?"
- "Our product is unique. Do we have to do everything in the guidance?"
- "Can we use data from other manufacturing plants in India and China?"

It is with real pleasure that I can acknowledge those people that I have been privileged to work with and have helped me with this book — Ronald Branning, Chris Burgess, Kevin Charrier, Jim DeMuth, Robert Dillard, T. Lynn Eudey, Jeff Field, Matt Hayat, Robert Kieffer, Sourav Kundu, David Lansky, Jacks Lee, Thomas Murphy, Jason Orloff, Dan Pilipauskas, Denise Spellman and Wayne Taylor.

This book is my interpretation of the process validation guideline and the applicable statistics. Other interpretations can be made. The recommendations are just that recommendations. Other statisticians can propose equally valid approaches. Also, the recommendations are the minimum expected. More would be done by most companies.

Last, all mistakes are mine and mine alone. This book should be considered a work in progress. I would be most grateful to hear from any readers that have questions, comments or content to share.

Lynn Torbeck

Note that Validation by Design® is a registered service mark of Lynn D. Torbeck.

www.pda.bookstore

INTRODUCTION

This book addresses the statistical issues expressed and implied in the Food and Drug Administration's FDA Guidance for Industry Process Validation: General Principles and Practices, Draft, November 2008.

As noted in the introduction:

> "This guidance outlines *the general principles and approaches* that the FDA considers to be appropriate elements of process validation for the manufacture of human and animal drug and biological products, including active pharmaceutical ingredients (API or drug substance), collectively referred to in this guidance as *drugs or products*. This guidance incorporates principles and approaches that *all manufacturers can use* in validating a manufacturing process (emphasis added.)

The Agency has not provided specific prescriptions or activities for implementing the general principles and approaches given in the guidance. It is the responsibility of each company to develop its own methods and procedures for meeting the requirements. However, each person reading the guidance has a different education, background and set of experiences. There is a need for more detailed discussion about specific tools and techniques, particularly statistical topics.

This book is unique in several ways. First, it presents and discusses only the statistical issues expressed or implied in the guidance. It is the author's experience

www.pda.bookstore

that there is less industry unity about statistical issues than for engineering, managerial or regulatory topics. There is a need to set a minimum threshold for interpretation and expectations. This will also enable discussions between departments within a company and between companies and regulatory agencies. Disagreements and misunderstandings can be resolved in the most scientific way relying on accepted statistical theory and accepted common practice.

The second unique feature is rewriting selected sentences from the guidance into questions. This focuses the readers attention toward answering the questions. This is an active mindset as opposed to the passive attitude experienced when reading a declarative sentence. The change in orientation is revealing and stimulates a precise response.

Third, the statistical issues are presented as self-contained stand-alone modules rather than an extended narrative. In many modules, worked examples are given with references for further study and reading. Users can turn directly to the specific topic to find the needed information.

Fourth, the book provides only information helpful in meeting the guidance. This is not a general purpose textbook for introductory statistics. Only statistical issues expressed or implied are included. The more important topics include:

- state of control
- process capability
- sampling, sample size and sampling plans
- statistical tolerance intervals
- variability
- summary statistics
- design of experiments (DOE)
- data and data collection.

Fifth, the statistical and quality topics necessary and presented here are derived from the body of knowledge (BOK) needed to pass the American Society for Quality (ASQ) exam for the Certified Quality Engineer (CQE) certification. Thus, the orientation is to continuous improvement in the quality of processes and products. Readers looking for a solid professional approach can use the ASQ BOK as a guide to further study and development.

Finally, the National Institute for Standards and Technology (NIST) e-Handbook for Engineering Statistics (NIST, 2009) is highly recommended for

www.pda.bookstore

study and reference. It is an excellent source of information particularly for the purposes of this book. It has, as the name implies, an engineering approach to quality improvement for processes and products. Further, since it is a government document it is free, copyright free and available on line.

A second government reference is the classic *Experimental Statistics* (Natrella, 1963). Look for a used copy on the internet. It is copyright free as well.

REFERENCES

NIST/SEMATECH (2009) e-Handbook of Statistical Methods. *http://www.itl.nist.gov/div898/handbook*, February.

Natrella, M.G. (1963) *Experimental Statistics*. Government Printing Office.

www.pda.bookstore