

Preface

Introducing the “Tree of Evaluation”

This volume is a substantial revision and modernization of *Evaluation Methods in Biomedical Informatics, Second Edition*, which was originally published in 2006. Much in the world of informatics has changed since the appearance of the second edition, and this volume has changed accordingly. This is signaled by the addition of the important word “health” to the title of the book, to reflect the increasing role that personal health applications play in the informatics world and to emphasize that the methods described in the book include but also transcend the work of the medical profession.

This remains primarily a textbook with a focus on widely applicable evaluation methods, including both quantitative and qualitative approaches. The chapters on quantitative studies retain an emphasis on measurement which is very important in informatics and is not typically covered in statistics courses. The book is written for use in formal courses for students of informatics at all levels, and for use as a reference by informatics researchers and professionals. It makes only modest assumptions about knowledge of basic statistics.

Because the author team spans the Atlantic, readers will find examples and references that reflect experiences and work performed on two continents.

The most significant changes from the second edition include:

- Based on the authors’ additional years of teaching experience, a significant attempt at clarifying complex issues of methodology that have proven challenging for students.
- The addition of three important new topics: evidence-based informatics (Chap. 5), methods using “real-world” data (Chap. 13), and mixed-methods studies (Chap. 17).
- Reconstructing the book with a larger number of somewhat shorter chapters.
- An update of examples and references, while respecting the importance of older studies that for decades have been guideposts for work in the field. Several older studies are retained as examples because of their value as teaching cases.

- The use of hyperlinks to enable users of the digital version to navigate quickly between interdependent, and now numbered, chapters and sections.
- Additional self-tests, with answers, and “Food for Thought” questions.

Like the earlier volumes, the third edition uses “evaluation” in the title while recognizing that the methods introduced apply to both evaluation and research. For this reason, this book could have been called “Empirical Methods...” to cover both. “Evaluation” was retained in the title to reflect the nature of informatics as an applied science grounded in practice -- and the fact that many studies of information resources using the methods described in this book are intended to inform practical decisions as opposed to, or in addition to, resulting in publications. To help readers clarify this, the initial chapters devote significant attention to the distinction between evaluation and research.

The Tree of Evaluation

To illustrate the topics covered and how the key ideas of the book relate to one another, we have represented the contents of this volume as a tree in the figure below. In several places, the discussions in the chapters refer back to this Fig. 1.

The progression of topics in the book rises from the base of the tree embracing the full range of approaches to evaluation. The early chapters discuss a range of issues common to all approaches and trace the differences between quantitative and qualitative methods to their philosophical roots. As the discussion progresses up the tree, it follows a major branch on the left to quantitative methods, introducing there an all-important distinction between studies that emphasize measurement and those that emphasize demonstration. This distinction is important because measurement is under-appreciated in informatics. From there, the discussion follows another major branch on the right into the world of qualitative methods, subsequently rising to a central branch of mixed methods. At the highest level of elevation, the tree reaches a canopy of important topics emphasizing communication and ethics.

What Isn’t Included

Informatics is an increasingly vast and diverse field, and no single work can cover all aspects of it. In revising this volume, we became increasingly aware of topics to which only passing reference is made or in some cases were not included at all.

Most important, while the methods introduced in this book are widely applicable, the discussions make no attempt to describe all of the information resources and purposes of their use, to which these methods can be applied. The examples illustrated are chosen for their potential to illustrate specific methods. Inevitably, readers of this book will take away from it not so much a deeper sense of “what works and

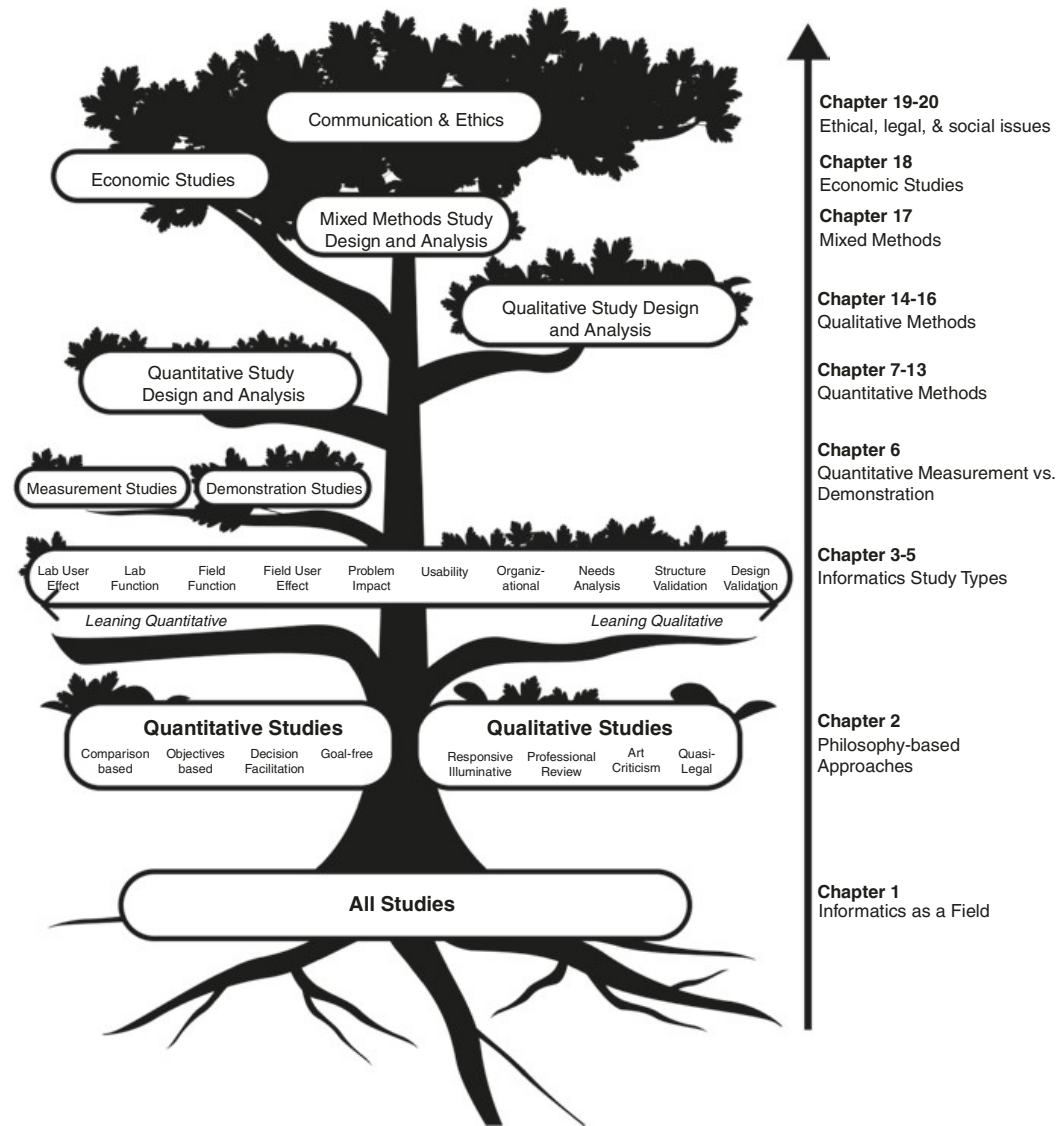


Fig. 1 The tree of evaluation

what doesn't" to improve health with information resources, but rather a much greater ability to determine rigorously what works and what doesn't—and why.

Because this is a basic methods textbook, scholars of evaluation methodology will not find cutting-edge techniques discussed here. We also omitted detailed discussion of what might be called intermediate level methods, such as principal component analysis and generalizability theory. For this reason, doctoral students who wish to employ these methods in their dissertation research will likely need further reading or course work.

While the evaluation methods addressed in this book are applicable to studies of information resources employed in bioinformatics and computational biology, the examples in the text include relatively few applications in these important fields.

In Sum

Returning to the Tree of Evaluation and extending the metaphor, we hope this volume serves for informatics the same purposes that a tree serves in nature, most notably endurance through its emphasis on the basic methods that will change more slowly than the resources to which they are applied. And like a tree that serves many purposes from cleaning the air to a play place for children, this volume will hopefully prove useful to everyone who seeks to improve health through the best possible use of the best possible information resources. Finally, those of us who work intimately with these methods appreciate them as a thing of beauty, much like a tree. And much like a tree that provides shelter for many forms of life, large and small—helping them to thrive and multiply—we hope our tree will both stimulate and sustain the growth of the field of informatics.

Ann Arbor, MI
Southampton, UK
Portland, OR

Charles P. Friedman
Jeremy C. Wyatt
Joan S. Ash

Acknowledgments

We are enormously grateful for the superb assistance of several colleagues in the creation of this book.

Monica Guo produced the artwork for many figures, including the “Tree of Evaluation,” and provided vital support to maintain consistency within and across chapters as they each went through several stages of review and revision. It is difficult to imagine how this volume—with authors spanning two continents, two versions of the English language, and eight time zones—could have come together without Monica’s creativity, dedication, and attention to detail. Her invariably cheerful “can do” approach to this work was a source of buoyancy for the entire team.

Joshua Rubin, as our work drew to a conclusion, provided enormous assistance in researching policy issues and in ensuring the consistency and completeness of references. We are equally appreciative of Josh’s creativity, dedication, and attention to detail.

Professor Mark Roberts of the University of Pittsburgh provided valuable insights and suggestions that supported the revision of the chapter on economic analysis.

We are grateful to Ted Shortliffe for writing a Foreword to this edition, as he so ably did for the previous two editions.

We also thank the students we collectively taught at the University of Michigan, Oregon Health & Science University, and several universities in the UK. Their curiosity and passion for learning helped us recognize which sections of the second edition required clarification and which additional topics should be addressed in this third edition. In particular, we appreciate Keaton Whittaker’s assistance with developing the new mixed-methods chapter.

And we are exceedingly grateful to our wonderful spouses—Patti, Sylvia, and Paul—for their encouragement to proceed through this work during a global pandemic, and their patience and understanding as we invaded what otherwise would have been family time to bring this work to conclusion.

Contents

Part I Landscape of Evaluation in Informatics

1	What Is Evaluation and Why Is It Challenging?	3
1.1	Introduction	3
1.2	The Place of Evaluation Within Informatics	4
1.3	Framing Evaluation as Answering Questions	6
1.4	Examples of Information Resources	9
1.5	Further Distinguishing Evaluation and Research.	11
1.6	Reasons for Performing Evaluations	12
1.7	Evaluation in Informatics as a Complex Socio-Technical Endeavor	13
1.7.1	Studying People	14
1.7.2	Professional Practice and Culture	14
1.7.3	Safety and Regulation	15
1.7.4	Domain Knowledge Complexity	16
1.7.5	Chains of Effect	17
1.7.6	Decision Making Processes	18
1.7.7	Data Availability and Quality	18
1.7.8	Performance Standards	19
1.7.9	Computer-Based Information Resources	19
1.8	Concluding Observations.	20
	Answers to Self-Tests	22
	Self-Test 1.1	22
	Self-Test 1.2	22
	References.	22
2	The Panorama of Evaluation Approaches	25
2.1	Introduction	25
2.2	Deeper Definitions of Evaluation.	26
2.3	The Common Anatomy of All Evaluation Studies.	27
2.4	Diverse Roles in Evaluation Studies	28
2.5	Some Philosophical Bases of Evaluation.	30

2.6	The Specific Anatomies of Quantitative and Qualitative Studies . .	32
2.7	Multiple Approaches to Evaluation	33
2.7.1	Approaches Rooted in Quantitative Assumptions	34
2.7.2	Approaches Rooted in Qualitative Assumptions	35
2.8	Conclusion: Why Are There so Many Approaches?	37
	Answers to Self-Tests	40
	References.	41
3	From Study Questions to Study Design: Exploring the Full Range of Informatics Study Types.	43
3.1	Introduction	43
3.2	The Key Role of Study Questions to Guide Study Design	44
3.2.1	Why Study Questions Are Important.	45
3.2.2	The Qualitative Process of Getting to a Final Set of Study Questions	46
3.2.3	What Is the Right Number of Study Questions?	46
3.2.4	How Should Study Questions Be Stated?	47
3.3	Ten Informatics Study Types	48
3.3.1	Needs Assessment Studies.	50
3.3.2	Design Validation Studies	50
3.3.3	Structure Validation Studies.	50
3.3.4	Usability Studies	51
3.3.5	Laboratory Function Studies	51
3.3.6	Field-Function Studies.	52
3.3.7	Laboratory-User Effect Studies	52
3.3.8	Field User Effect Studies	53
3.3.9	Problem Impact Studies.	54
3.3.10	Organizational and System Studies	54
3.4	Specific Features Distinguishing the Ten Informatics Study Types.	55
3.4.1	The Setting in Which the Study Takes Place.	55
3.4.2	The Status of the Information Resource Employed.	55
3.4.3	The Sampled Resource Users	55
3.4.4	The Sampled Tasks	57
3.4.5	What Is Observed or Measured	57
3.4.6	Sampling of Users and Tasks.	58
3.5	Conclusions: Approaching a Recipe for Study Design	59
	Answers to Self-Tests	60
	References.	61
4	Study Design Scenarios and Examples	63
4.1	Introduction	63
4.2	A Scenario Invoking a Needs Assessment Study.	64
4.2.1	Typical Study Questions	64
4.2.2	Performing Needs Assessments.	65
4.3	A Scenario Invoking a Usability and/or a Function Study	66
4.3.1	Typical Study Questions	67
4.3.2	Performing Usability and Function Studies.	67

4.4	A Scenario Invoking a Field User Effect Study	68
4.4.1	Typical Study Questions	69
4.4.2	Performing Effect Studies	69
4.5	Scenario Invoking Problem Impact and Organization and System Studies	70
4.5.1	Typical Study Questions	71
4.5.2	Performing Impact and Organizational and S ystem Studies	71
4.6	Exemplary Evaluation Studies from the Literature	72
4.6.1	Needs Assessment Study: “Using Ethnography to Build a Working System: Rethinking Basic Design Assumptions”	73
4.6.2	Usability Study: “Assessing Performance of an Electronic Health Record Using Cognitive Task Analysis”.	73
4.6.3	Lab Function Study: “Diagnostic Inaccuracy of Smartphone Applications for Melanoma Detection”	73
4.6.4	Field Function Study: “Evaluation of a Machine Learning Capability for a Clinical Decision Support System to Enhance Antimicrobial Stewardship Programs”	74
4.6.5	Lab User Effect Study: “Applying Human Factors Principles to Alert Design Increases Efficiency and Reduces Prescribing Errors in a Scenario-Base Simulation”	74
4.6.6	Field User Effect Study: “Reminders to Physicians from an Introspective Computer Medical Record: A Two-Year Randomized Trial”	75
4.6.7	Field User Effect Study: “Electronic Health Records and Health Care Quality Over Time in a Federally Qualified Health Center”	75
4.6.8	Problem Impact Study: “Effects of a Mobile Phone Short Message Service on Antiretroviral Treatment Adherence in Kenya (WelTel Kenya1): A Randomised Trial”	76
4.6.9	Organizational Study: “Electronic Health Record Adoption in US Hospitals: The Emergence of a Digital ‘Advanced Use’ Divide”	76
4.7	Summary and Conclusions	77
	References	77
5	Study Planning	79
5.1	Introduction	80
5.2	Adopting an Evaluation Mindset	81
5.3	Why It May Not Work Out as Planned	83
5.3.1	Sometimes Stakeholders Would Rather Not Know	84
5.3.2	Differences in Values	84

5.4	Planning a Study	85
5.4.1	Assessing Stakeholder Needs	85
5.4.2	Choosing the Level of Evaluation	87
5.4.3	Matching What Is Studied to the Stage in the Life Cycle	88
5.4.4	Considering Context: Local Resources and Organizational Knowledge	89
5.4.5	The Focus: Matching What Is Evaluated to the Type of Information Resource.	89
5.4.6	The Search for Evidence	90
5.5	The Value of Evidence-Based Health Informatics.	90
5.6	Searching the Literature.	91
5.6.1	Identifying a Search Topic	92
5.6.2	Frameworks and Theories	92
5.7	How to Conduct a Literature Search	93
5.8	Systematic Reviews and Narrative Meta-Analyses	95
5.8.1	Systematic Reviews	95
5.8.2	Meta-Narrative Reviews	95
5.9	Project Management	96
5.9.1	The Team	96
5.9.2	Project Management Tools.	97
5.10	Summary	98
	Answers to Self-Tests	98
	References.	99

Part II Quantitative Studies

6	The Structure of Quantitative Studies	103
6.1	Introduction	103
6.2	Elements of a Measurement Process	105
6.2.1	Measurement	106
6.2.2	Attribute	106
6.2.3	Object and Object Class.	106
6.2.4	Attribute–Object Class Pairs	107
6.2.5	Attributes as “Constructs”	107
6.2.6	Measurement Instruments	108
6.2.7	Independent Observations	109
6.3	Key Object Classes and Types of Observations for Measurement in Informatics.	110
6.3.1	Key Object Classes	110
6.3.2	Key Categories of Observations.	111
6.4	Levels of Measurement	112
6.5	Importance of Measurement in Quantitative Studies.	114
6.6	Measurement and Demonstration Studies	115
6.7	Goals and Structure of Measurement Studies	117
6.8	The Structure and Differing Terminologies of Demonstration Studies	120

6.9	Demonstration Study Designs	121
6.9.1	Prospective and Retrospective Studies.	121
6.9.2	Descriptive, Interventional, and Correlational Studies	122
6.10	Demonstration Studies and Stages of Resource Development	124
	Answers to Self-tests.	125
	References.	126
7	Measurement Fundamentals: Reliability and Validity	129
7.1	Introduction	130
7.2	The Classical Theory of Measurement: Framing Error as Reliability and Validity	131
7.3	Toward Quantifying Reliability and Validity	132
7.4	A Conceptual Analogy Between Measurement and Archery	133
7.5	Two Methods to Estimate Reliability.	135
7.6	Quantifying Reliability and Measurement Errors	137
7.6.1	The Benefit of Calculating Reliability.	137
7.6.2	Enabling Reliability Estimation Using Objects-by- Observations Matrices	138
7.6.3	The Standard Error of Measurement	139
7.7	Design Considerations in Measurement Studies to Estimate Reliability.	141
7.7.1	Dependence of Reliability on Populations of Objects and Observations	141
7.7.2	Effects on Reliability of Changing the Number of <i>Observations</i> in a Measurement Process	142
7.7.3	Value of the Prophecy Formula: How Much Reliability Is Enough?	142
7.7.4	Effects on Reliability of Changing the Number of <i>Objects</i> in a Measurement Process	143
7.8	Validity and Its Estimation.	144
7.8.1	Distinguishing Validity from Reliability	144
7.8.2	Content Validity	146
7.8.3	Criterion-Related Validity	147
7.8.4	Construct Validity	148
7.9	Concluding Observations About Validity.	150
7.10	Generalizability Theory: An Extension of Classical Theory	151
	Answers to Self-Tests	152
	References.	154
8	Conducting Measurement Studies and Using the Results.	155
8.1	Introduction	155
8.2	Process of Measurement Studies	156
8.3	How to Improve Measurement Using Measurement Study Results	158

8.4	Using Measurement Study Results to Diagnose Measurement Problems	159
8.4.1	Analyzing the Objects-by-Observations Matrix	159
8.4.2	Improving a Measurement Process	162
8.5	The Relationship Between Measurement Reliability and Demonstration Study Results	164
8.5.1	Effects of Measurement Error in Descriptive and Interventional Demonstration Studies	165
8.5.2	Effects of Measurement Error in Correlational Demonstration Studies	166
8.6	Demonstration Study Results and Measurement Error: A Detailed Example.	168
8.6.1	Reliability Estimate	171
8.6.2	Demonstration Study Correcting the Correlation for Attenuation.	171
8.6.3	Additional Considerations from This Example	172
8.7	Conclusion	174
	Answers to Self-Tests	174
	References.	176
9	Designing Measurement Processes and Instruments.	177
9.1	Introduction	178
9.2	When Judges' Opinions Are the Primary Observations.	180
9.2.1	Sources of Variation Among Judges	181
9.2.2	Number of Judges Needed for Reliable Measurement	182
9.2.3	Improving Measurements That Use Judges as Observations	182
9.3	When Tasks Are the Primary Observation.	184
9.3.1	Sources of Variation Among Tasks	185
9.3.2	Number of Tasks Needed for Reliable Measurement	187
9.3.3	Improving Measurements That Use Tasks as Observations	188
9.4	When Items Are the Primary Observation.	194
9.4.1	Sources of Variation Among Items	194
9.4.2	Number of Items Needed for Reliable Measurement	196
9.4.3	Improving Measurements That Use Items as Observations	196
9.4.4	The Ratings Paradox	199
9.5	Conclusion	200
	Answers to Self-Tests	201
	References.	202
10	Conducting Demonstration Studies.	205
10.1	Introduction	206
10.2	Overview of Demonstration Studies	207
10.2.1	Types of Demonstration Studies.	207

10.2.2	Terminology for Demonstration Studies	208
10.2.3	Demonstration Study Types Further Distinguished	210
10.2.4	Internal and External Validity of Demonstration Studies	212
10.3	The Process of Conducting Demonstration Studies.	213
10.4	Threats to Internal Validity That Affect All Demonstration Studies, and How to Avoid Them	214
10.4.1	Assessment Biases and the Hawthorne Effect	214
10.4.2	Incomplete Response and Missing Data.	215
10.4.3	Insufficient Data to Allow a Clear Answer to the Study Question	216
10.4.4	Inadvertently Using Different Data Items or Codes.	217
10.5	Threats to the External Validity and How to Overcome Them . . .	217
10.5.1	The Volunteer Effect.	218
10.5.2	Age, Ethnic Group, Gender and Other Biases	218
10.5.3	Study Setting Bias	219
10.5.4	Implementation Bias.	219
10.5.5	Training Set Bias and Overfitting.	219
10.6	Analysis Methods for Descriptive Studies That Also Apply to All Study Types	221
10.6.1	Graphical Portrayal of the Results	222
10.6.2	Indices of Central Tendency.	222
10.6.3	Indices of Variability.	222
10.7	Conclusions	224
	Answers to Self-Tests	225
	References.	227
11	Design of Interventional Studies	229
11.1	Introduction	229
11.2	Challenges When Designing Interventional Studies to Attribute Causation	230
11.2.1	Secular Trends	231
11.2.2	Regression to the Mean	232
11.2.3	Association Is Not Causation	233
11.3	Control Strategies for Interventional Studies.	233
11.3.1	Uncontrolled Study.	233
11.3.2	Historically Controlled or Before-After Studies	234
11.3.3	Simultaneous External Controls.	235
11.3.4	Internally and Externally Controlled Before-After Studies	236
11.3.5	Simultaneous Randomized Controls	237
11.3.6	Cluster Randomized and Step Wedge Designs.	238
11.3.7	Matched Controls as an Alternative to Randomization, and the Fallacy of Case-Control Studies	239
11.3.8	Summary of Control Strategies	240

11.4	Biases and Challenges to the Internal Validity of Interventional Studies	241
11.4.1	The Placebo Effect	241
11.4.2	Allocation and Recruitment Biases	242
11.4.3	Biases Associated with Data Collection	243
11.4.4	Carryover Effect	244
11.4.5	Complex Interventions and Associated Biases	244
11.5	Statistical Inference: Considering the Effect of Chance	248
11.5.1	Effect Size and “Effectiveness”	248
11.5.2	Statistical Inference	249
11.5.3	Study Power	250
11.5.4	Data Dredging	251
11.6	Conclusions	251
	Answers to Self-Tests	253
	References	256
12	Analyzing Interventional Study Results	259
12.1	Introduction	259
12.2	Grand Strategy for Analysis of Study Results	260
12.3	Analyzing Studies with Discrete Independent and Discrete Dependent Variables: Basic Techniques	262
12.3.1	Contingency Tables	262
12.3.2	Using Contingency (2×2) Tables: Indices of Effect Size	263
12.3.3	Chi-Square and Fisher’s Exact Test for Statistical Significance	265
12.3.4	Cohen’s Kappa: A Useful Effect Size Index	266
12.4	Analyzing Studies with Discrete Independent and Discrete Dependent Variables: ROC Analysis	267
12.4.1	Estimating the Performance of a Classifier Using a Single Cut Point	267
12.4.2	ROC Analysis	268
12.5	Analyzing Studies with Continuous Dependent Variables and Discrete Independent Variables	270
12.5.1	Two Group Studies with Continuous Outcome Variables	270
12.5.2	Alternative Methods for Expressing Effect Size	271
12.5.3	Logic of Analysis of Variance (ANOVA)	272
12.5.4	Using ANOVA to Test Statistical Significance	273
12.5.5	Special Issues with ANOVA	275
12.6	Analyzing Studies with Continuous Independent and Dependent Variables	277
12.6.1	Studies with One Continuous Independent Variable and One Continuous Dependent Variable	277
12.6.2	Studies with Multiple Independent Variables and a Continuous Dependent Variable	278

12.6.3	Relationships to Other Methods	279
12.7	Choice of Effect Size Metrics: Absolute Change, Relative Change, Number Needed to Treat	280
12.8	Special Considerations when Studying Predictive Models	281
12.8.1	Use of Training and Testing Datasets	281
12.8.2	Some Cautionary Tales About Studying Predictive Models	282
12.9	Conclusions	283
	Answers to Self-Tests	284
	References	286
13	Designing and Carrying Out Correlational Studies	
	Using Real-World Data	289
13.1	Introduction	289
13.2	Reasons for Carrying Out Correlational Studies	290
13.3	Types of Data Used in Correlational Studies	292
13.4	Advantages and Disadvantages of Correlational Studies Using Routine Data	293
13.4.1	Advantages	293
13.4.2	Disadvantages	294
13.5	How to Carry Out Correlational Studies Using Person-Level Datasets	295
13.5.1	Agreeing to the Study Questions, Cohorts, and Study Variables	296
13.5.2	Identify the Datasets, Code Sets, and Data Analysis Routines	296
13.5.3	Develop and Finalize the Data Analysis Plan	297
13.5.4	Applying for Permissions	298
13.5.5	Virtual Data Enclave (Virtual Safe Haven)	298
13.5.6	Data Linkage	298
13.5.7	Data Cleaning	299
13.5.8	Estimate the Size and Direction of Potential Biases	300
13.5.9	Test and Refine the Data Analysis Plan and Carry Out the Data Analysis	300
13.5.10	Carry Out Sensitivity Analyses	300
13.5.11	Check for Statistical Disclosure Risk	301
13.5.12	Disseminate the Study Results	301
13.6	The Challenge of Inferring Causation from Association	301
13.6.1	Spurious Correlation	302
13.6.2	Bradford Hill's Criteria to Support Causality in Correlational Studies	303
13.7	Factors Leading to False Inference of Causation	304
13.7.1	Selection Bias	306
13.7.2	Information Bias	307
13.7.3	Confounding	309

13.7.4	Some General Methods to Detect or Control for Confounding	313
13.8	Advanced Analysis Methods for Correlational Studies	314
13.8.1	The Case-Crossover Design	314
13.8.2	Instrumental Variable Analysis	315
13.8.3	Regression Discontinuity Design	316
13.9	Conclusions	317
	Answers to Self-Tests	319
	References	323

Part III Qualitative Studies

14	An Introduction to Qualitative Evaluation Approaches	329
14.1	Introduction	329
14.2	Definition of the Qualitative Approach	330
14.3	Motivation for Qualitative Studies: What People Really Want to Know	332
14.4	Support for Qualitative Approaches	333
14.5	Why Are Qualitative Studies Especially Useful in Informatics?	335
14.6	When Are Qualitative Studies Appropriate?	336
14.7	Rigorous, but Different, Methodology	337
14.8	Qualitative Approaches and Their Philosophical Premises	338
14.9	Natural History of a Qualitative Study	340
14.9.1	Negotiation of the “Ground Rules” of the Study	340
14.9.2	Immersion into the Environment and Initial Data Collection	340
14.9.3	Iterative Qualitative Loop	341
14.9.4	Preliminary Report	342
14.9.5	Final Report	343
14.10	Summary: The Value of Different Approaches	343
14.11	Two Example Summaries	344
	Answers to Self-Tests	345
	References	346
15	Qualitative Study Design and Data Collection	347
15.1	Introduction	347
15.2	The Iterative Loop of Qualitative Processes	349
15.3	Theories, Frameworks, and Models that Can Guide the Process.	350
15.4	Strategies for Study Rigor	353
15.4.1	Reflexivity	353
15.4.2	Triangulation	354
15.4.3	Member Checking	356
15.4.4	Data Saturation	356
15.4.5	Audit Trail	357
15.5	Specifics of Conducting Qualitative Studies	359
15.5.1	Site and Informant Selection	360

15.5.2	The Team	361
15.6	Techniques for Data Collection	362
15.6.1	Entering the Field	362
15.6.2	Interviews	363
15.6.3	Focus Groups	367
15.6.4	Observations	368
15.6.5	Field Notes	370
15.6.6	Gathering Naturally Occurring Data	372
15.7	Toolkits of Complementary Techniques	374
15.7.1	Case Studies	374
15.7.2	Action Research	375
15.7.3	Rapid Assessment Process	375
15.7.4	Virtual Ethnography Techniques	376
15.8	The Future of Qualitative Data Gathering Methods	377
	Answers to Self-Tests	377
	References	379
16	Qualitative Data Analysis and Presentation of Analysis Results	381
16.1	Introduction	381
16.2	What Is Qualitative Analysis?	382
16.3	Approaches to Analysis	382
16.4	Qualitative Data Management	385
16.5	Qualitative Analysis Software	386
16.6	How to Code	387
16.7	How to Interpret	391
16.8	Other Analysis Approaches	392
16.8.1	Narrative Analysis	392
16.8.2	Analysis of Naturally Occurring Text or Visual Data. . . .	393
16.9	Using Graphics in the Analysis and Reporting Processes	394
16.10	Evaluating the Quality of Qualitative Analysis	396
16.11	The Future of Qualitative Analysis	397
	Answers to Self-Tests	397
	References	398

Part IV Special Study Types

17	Mixed Methods Studies	403
17.1	Introduction	403
17.2	The Ultimate Triangulation	404
17.3	Why Use Mixed Methods	405
17.4	Designing Mixed Methods Studies	406
17.5	Data Collection Strategies	410
17.6	Data Analysis Strategies	410
17.7	Presenting Results of Mixed Methods Studies	411
17.8	Usability Studies as Exemplars of Mixed Methods	412

17.8.1	What Is Usability?	412
17.8.2	Importance of Usability	413
17.8.3	Usability Testing Methods	413
17.8.4	Examples of Usability Studies in Informatics	414
17.9	Organizational Systems-Level Studies	416
17.9.1	What Are Organizational Systems-Level Studies?	416
17.9.2	Methods Used in Organizational and System-Level Studies	416
17.10	Special Ethical Considerations in Mixed Method Studies.	418
17.11	The Future of Mixed Methods Inquiry	418
	Answers to Self-Tests	419
	References.	420
18	Principles of Economic Evaluation and Their Application to Informatics.	423
18.1	Introduction to Economic Studies	423
18.1.1	Motivation for Economic Analysis.	424
18.2	Principles of Economic Analysis	425
18.2.1	Types of Cost Studies	426
18.3	Conducting an Economic Analysis	428
18.3.1	The Perspective of the Analysis	428
18.3.2	Time Frame of the Analysis	430
18.4	Definition and Measurement of Costs	431
18.4.1	Why Costs Do Not Equal Charges.	431
18.4.2	Mechanics of Cost Determinations	433
18.5	Definition and Measurement of Outcomes and Benefits	434
18.5.1	Matching the Correct Outcome Measure to the Problem.	434
18.5.2	Adjusting for Quality of Life	435
18.6	Cost-Minimizing Analysis	435
18.7	Cost-Effectiveness Analysis.	437
18.7.1	The Incremental Cost-Effectiveness Ratio.	438
18.7.2	Cost-Benefit Analyses: The Cost-Benefit Ratio	440
18.7.3	Discounting Future Costs and Benefits	442
18.8	Sensitivity Analysis	444
18.8.1	Structural Sensitivity Analyses.	444
18.8.2	Parameter Estimate Sensitivity Analyses	445
18.9	Special Characteristics of Cost-Effectiveness Studies in Biomedical and Health Informatics.	447
18.9.1	System Start Up and Existing Infrastructure Costs	447
18.9.2	Sharing Clinical Information System Costs	448
18.10	Critically Appraising Sample Cost-Effectiveness Studies in Biomedical and Health Informatics	448
18.11	Conclusion	450
	Answers to Self-Tests	451
	References.	452

Part V Ethical, Legal and Social Issues in Evaluation

19 Proposing Studies and Communicating Their Results	457
19.1 Introduction	457
19.2 Writing Evaluation Proposals	458
19.2.1 Why Proposals Are Necessary and Difficult	458
19.2.2 Format of a Study Proposal	460
19.2.3 Suggestions for Expressing Study Plans	460
19.2.4 Special Issues in Proposing Evaluations	464
19.3 Refereeing Evaluation Studies	465
19.4 Communicating the Results of Completed Studies	467
19.4.1 What Are the Options for Communicating Study Results?	467
19.4.2 What Is the Role of the Study Team: Reporter or Change Agent?	469
19.4.3 Role of the Evaluation Contract	470
19.5 Specific Issues of Report Writing	471
19.5.1 The IMRaD Format for Writing Study Reports	471
19.5.2 Writing Qualitative Study Reports	472
19.6 Conclusion	473
References	474
20 Ethics, Safety, and Closing Thoughts	475
20.1 Introduction	475
20.2 Broad Ethical Issues in ICT Evaluation	476
20.2.1 Standard Practices for the Protection of Human Subjects	476
20.2.2 The Safety of Implementing ICT	479
20.2.3 Encouraging Diversity	480
20.3 Ethical Issues During Each Phase of a Study	481
20.3.1 Planning for Ethical Evaluation	481
20.3.2 The Data Gathering and Storage Phase	484
20.3.3 Analysis and Interpretation Phase	487
20.3.4 Reporting Results Phase	487
20.4 Ethical Considerations for the Future	488
20.5 Closing Thoughts About the Future of Informatics Evaluation	489
20.5.1 What Will Be Studied	489
20.5.2 Future Trends in Evaluation Methods	492
Answers to Self-Tests	493
References	494
Glossary	497
Index	521