CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-28804-0 — Applied Longitudinal Data Analysis for Medical Science
3rd Edition
Frontmatter
More Information

Applied Longitudinal Data Analysis for Medical Science

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-28804-0 — Applied Longitudinal Data Analysis for Medical Science
3rd Edition
Frontmatter
More Information

Applied Longitudinal Data Analysis for Medical Science

A Practical Guide

Jos W. R. Twisk
Amsterdam University Medical Centers

CAMBRIDGE UNIVERSITY PRESS

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE UNIVERSITY PRESS

Shaftesbury Road, Cambridge CB2 8EA, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India

103 Penang Road, #05-06/07, Visioncrest Commercial, Singapore 238467

Cambridge University Press is part of Cambridge University Press & Assessment, a department of the University of Cambridge.

We share the University's mission to contribute to society through the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781009288040

DOI: 10.1017/9781009288002

© Jos W. R. Twisk 2023

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press & Assessment.

First published 2023
A catalogue record for this publication is available from the British Library.

Library of Congress Cataloging-in-Publication Data
Names: Twisk, Jos W. R., 1962- author.
Title: Applied longitudinal data analysis for medical science : a practical guide / Jos W.R. Twisk.
Other titles: Applied longitudinal data analysis for epidemiology
Description: 3. | Cambridge, United Kingdom ; New York, NY : Cambridge University Press, 2023. | Preceded by Applied longitudinal data analysis for epidemiology / Jos W.R. Twisk. Second edition. 2013. | Includes bibliographical references and index.
Identifiers: LCCN 2022053460 (print) | LCCN 2022053461 (ebook) | ISBN 9781009288040 (hardback) | ISBN 9781009288033 (paperback) | ISBN 9781009288002 (epub)
Subjects: MESH: Longitudinal Studies | Data Analysis | Data Interpretation, Statistical | Models, Statistical
Classification: LCC RA652.2.M3 (print) | LCC RA652.2.M3 (ebook) | NLM WA 950 | DDC 614.4072/7–dc23/eng/20230125
LC record available at https://lccn.loc.gov/2022053460
LC ebook record available at https://lccn.loc.gov/2022053461

ISBN 978-1-009-28804-0 Hardback
ISBN 978-1-009-28803-3 Paperback

Cambridge University Press & Assessment has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

[Medical disclaimer:]

Every effort has been made in preparing this book to provide accurate and up-to-date information that is in accord with accepted standards and practice at the time of publication. Although case histories are drawn from actual cases, every effort has been made to disguise the identities of the individuals involved. Nevertheless, the authors, editors, and publishers can make no warranties that the information contained herein is totally free from error, not least because clinical standards are constantly changing through research and regulation. The authors, editors, and publishers therefore disclaim all liability for direct or consequential damages resulting from the use of material contained in this book. Readers are strongly advised to pay careful attention to information provided by the manufacturer of any drugs or equipment that they plan to use.

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-28804-0 — Applied Longitudinal Data Analysis for Medical Science
3rd Edition
Frontmatter
More Information

To Marjon, Mike and Nick

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-28804-0 — Applied Longitudinal Data Analysis for Medical Science
3rd Edition
Frontmatter
More Information

Content

Preface *page* xi
Acknowledgements xii

1. **Introduction** 1
   - **1.1 Introduction** 1
   - **1.2 Study Design** 1
   - 1.2.1 Observational Longitudinal Studies 2
   - **1.3 General Approach** 4
   - **1.4 Prior Knowledge** 4
   - **1.5 Example** 4
   - **1.6 Software** 5
   - **1.7 Data Structure** 6
   - **1.8 What is New in the Third Edition** 6
2. **Continuous Outcome Variables** 7
   - **2.1 Two Measurements** 7
   - 2.1.1 Example 7
   - **2.2 Non-parametric Equivalent of the Paired *t*-test** 8
   - 2.2.1 Example 9
   - **2.3 More than Two Measurements** 9
   - 2.3.1 The Univariate Approach: A Numerical Example 12
   - 2.3.2 The Shape of the Relationship between an Outcome Variable and Time 13
   - 2.3.2.1 A Numerical Example 14
   - 2.3.3 Example 15
   - **2.4 The Univariate or the Multivariate Approach?** 19
   - **2.5 Comparing Groups** 19
   - 2.5.1 The Univariate Approach: A Numerical Example 20
   - 2.5.2 Example 21
   - **2.6 Comments** 25
   - **2.7 Post-hoc Procedures** 25
   - **2.8 Different Contrasts** 26
   - 2.8.1 Example 26
   - **2.9 Non-parametric Equivalent of GLM for Repeated Measures** 29
   - 2.9.1 Example 30
3. **Continuous Outcome Variables: Regression-based Methods** 31
   - **3.1 Introduction** 31
   - **3.2 Longitudinal Regression Methods** 31
   - **3.3 Mixed Model Analysis** 32
   - 3.3.1 Introduction 32
   - 3.3.2 Mixed Models for Longitudinal Data Analysis 32
   - 3.3.3 Example 34
   - 3.3.4 Interpretation of the Regression Coefficient 38
   - 3.3.5 Comments 41
   - **3.4 Generalised Estimating Equations** 42
   - 3.4.1 Introduction 42
   - 3.4.2 Correlation Structures 42
   - 3.4.3 Example 44
   - 3.4.3.1 Different Correlation Structures 46
   - **3.5 Comparison between Mixed Model Analysis and GEE Analysis** 48
   - **3.6 The Adjustment for Covariance Method** 49
   - 3.6.1 Example 50
   - 3.6.2 Extension of Mixed Model Analysis 54
   - 3.6.3 Comments 54
4. **The Modelling of Time** 56
   - **4.1 Growth Curve Analysis** 56
   - **4.2 Comparing Groups** 60
   - **4.3 Adjustment for Time** 64
   - 4.3.1 Time versus Age 68
   - **4.4 Interaction with Time** 69
   - **4.5 Classification of Subjects with Different Growth Trajectories** 70
5. **Models to Disentangle the Between- and Within-subjects Relationship** 76
   - **5.1 Introduction** 76
   - **5.2 Hybrid Models** 76
   - 5.2.1 Example 76
   - 5.2.2 Direct Estimation of the Hybrid Model 78
   - 5.2.3 Hybrid Models with Categorical Time-dependent Covariates 80

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-28804-0 — Applied Longitudinal Data Analysis for Medical Science
3rd Edition
Frontmatter
More Information

5.2.4 Comments 82
5.3 Models to Estimate the Within-subjects Part of the Longitudinal Relationship 82
5.3.1. Introduction 82
5.3.2 Model of Changes 83
5.3.2.1 Example 83
5.3.2.2 Another Example 85
5.3.3 Autoregressive Model 88
5.3.3.1 Example 88
5.3.4 Comments 90

6. **Causality in Observational Longitudinal Studies** 92
6.1 Time-lag Models 92
6.1.1 Example 92
6.1.2 Comments 92
6.2 Longitudinal Mediation Models 94
6.2.1 Example 96
6.2.2 Comments 103
6.3 Other Methods that Claim to Estimate Causal Relationships 106
6.3.1 G-methods 107
6.3.2 Joint models 110
6.3.2.1 Example 111

7. **Dichotomous Outcome Variables** 116
7.1 Two Measurements 116
7.2 More than Two Measurements 117
7.3 Comparing Groups 117
7.4 Example 117
7.4.1 Introduction 117
7.4.2 Development over Time 117
7.4.3 Comparing Groups 119
7.5 Longitudinal Regression Methods 119
7.5.1 Introduction 119
7.5.2 Generalised Estimating Equations 121
7.5.3 Mixed Model Analysis 124
7.5.4 Comparison between GEE Analysis and Mixed Model Analysis 127
7.5.5 The Adjustment for Covariance Method 130
7.5.6 Models to Disentangle the Between- and Within-subjects Relationship 130
7.5.7 Comments 133

8. **Categorical and Count Outcome Variables** 134
8.1 Categorical Outcome Variables 134
8.1.1 Two Measurements 134
8.1.2 More than Two Measurements 135
8.1.3 Comparing Groups 135
8.1.4 Example 135
8.1.5 Regression-based Methods 136
8.1.5.1 Example 138
8.2 Count Outcome Variables 141
8.2.1 Example 143
8.2.1.1 Introduction 143
8.2.1.2 GEE Analysis 143
8.2.1.3 Mixed Model Analysis 146
8.2.2 Comparison between GEE Analysis and Mixed Model Analysis 147
8.2.3 Negative Binomial Regression Analysis 148
8.2.4 Comments 150

9. **Outcome Variables with Floor or Ceiling Effects** 152
9.1 Introduction 152
9.2 Tobit Mixed Model Analysis 153
9.2.1 Example 153
9.3 Longitudinal Two-part Models 159
9.3.1 Example 160
9.3.2 Comments 162

10. **Analysis of Longitudinal Intervention Studies** 164
10.1 Introduction 164
10.2 Continuous Outcome Variables 164
10.2.1 Randomised Controlled Trials with One Follow-up Measurement 165
10.2.1.1 Example 168
10.2.2 Randomised Controlled Trials with More than One Follow-up Measurement 172
10.2.2.1 Simple Analysis 175
10.2.2.2 Summary Statistics 177
10.2.2.3 Generalised Linear Model for Repeated Measures 178
10.2.2.4 Generalised Linear Model for Repeated Measures Adjusted for Baseline 178
10.2.2.5 Regression-based Methods 178
10.3 Dichotomous Outcome Variables 187
10.3.1 Introduction 187
10.3.2 Simple Analysis 188
10.3.3 Regression-based Methods 189
10.3.4 Other Methods 191
10.4 Stepped Wedge Designs 195
10.5 Comments 196

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-28804-0 — Applied Longitudinal Data Analysis for Medical Science
3rd Edition
Frontmatter
More Information

10.6 Beyond the Randomised Controlled Trial 197
10.6.1 Difference in Difference Method 198
10.6.1.1 Example 198
10.6.1.2 Comments 199

11. **Missing Data in Longitudinal Studies** 201
11.1 Introduction 201
11.2 Informative or Non-informative Missing Data 201
11.3 Example 202
11.3.1 Generating Datasets with Missing Data 202
11.3.2 Analysis of Determinants for Missing Data 203
11.4 Analysis Performed on Datasets with Missing Data 204
11.5 Imputation Methods 205
11.5.1 Continuous Variables 205
11.5.1.1 Cross-sectional Imputation Methods 205
11.5.1.2 Longitudinal Imputation Methods 206
11.5.1.3 Comments 206
11.5.1.4 Multiple Imputation 206
11.5.2 Dichotomous and Categorical Variables 207
11.5.3 Example 207
11.5.3.1 Continuous Variables 207
11.5.3.2 Multiple Imputation in Combination with Mixed Model Analysis? 209
11.5.3.3 Additional Analyses 210
11.5.3.4 Dichotomous Variables 210
11.5.4 Comments 212
11.6 Alternative Methods 213
11.7 GEE Analysis or Mixed Model Analysis for the Analysis of Datasets with Missing Data? 214
11.8 Conclusions 214

12. **Sample Size Calculations** 216
12.1 Introduction 216
12.2 Example 218
12.3 Comment 219

13. **Software for Longitudinal Data Analysis** 220
13.1 Introduction 220
13.2 GEE Analysis with a Continuous Outcome Variable 220
13.2.1 STATA 220
13.2.2 SAS 220
13.2.3 R 221
13.2.4 SPSS 222
13.2.5 Overview 223
13.3 GEE Analysis with a Dichotomous Outcome Variable 224
13.3.1 STATA 224
13.3.2 SAS 224
13.3.3 R 224
13.3.4 SPSS 224
13.3.5 Overview 224
13.4 Mixed Model Analysis with a Continuous Outcome Variable 226
13.4.1 Introduction 226
13.4.2 STATA 226
13.4.3 SAS 227
13.4.4 R 229
13.4.5 SPSS 231
13.4.6 Overview 235
13.5 Mixed Model Analysis with a Dichotomous Outcome Variable 235
13.5.1 Introduction 235
13.5.2 STATA 235
13.5.3 SAS 236
13.5.4 R 239
13.5.5 SPSS 240
13.5.6 Overview 242

References 243
Index 251

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-28804-0 — Applied Longitudinal Data Analysis for Medical Science
3rd Edition
Frontmatter
More Information

Preface

The most important feature of this book is the word applied in the title. This implies that the emphasis of this book lies more on the application of statistical methods for longitudinal data analysis and not so much on the mathematical background. In most other books on longitudinal data analysis, the mathematical background is the major issue, which may not be surprising since (nearly) all the books on this topic have been written by statisticians. Although statisticians fully understand the difficult mathematical material underlying longitudinal data analysis, they often have difficulty in explaining this complex material in a way that is understandable for the researchers who have to use the method or interpret the results. Therefore, this book is not written by a statistician, but by an applied medical researcher. In fact, an applied researcher is not primarily interested in the basic (difficult) mathematical background of the statistical methods, but in finding the answer to a specific research question; the applied researcher wants to know how to apply a statistical method and how to interpret the results. Owing to their different basic interests and different level of thinking, communication problems between statisticians and applied researchers are quite common. This, in addition to the growing interest in longitudinal studies, initiated the writing of this book: a book on longitudinal data analysis, which is especially suitable for non-statistical applied researchers. The aim of this book is to provide a practical guide on how to handle data from a longitudinal study. The purpose of this book is to build a bridge over the communication gap that exists between statisticians and applied researchers regarding the (complicated) topic of longitudinal data analysis.

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-28804-0 — Applied Longitudinal Data Analysis for Medical Science
3rd Edition
Frontmatter
More Information

Acknowledgements

I am very grateful to all my colleagues and students who came to me with (mostly) practical questions on longitudinal data analysis. This book is based on all those questions.

© in this web service Cambridge University Press & Assessment www.cambridge.org