

Statistics for Biology and Health

Niel Hens · Ziv Shkedy · Marc Aerts · Christel Faes · Pierre Van Damme · Philippe Beutels

Modeling Infectious Disease Parameters Based on Serological and Social Contact Data

A Modern Statistical Perspective

Mathematical epidemiology of infectious diseases usually involves describing the flow of individuals between mutually exclusive infection states. One of the key parameters describing the transition from the susceptible to the infected class is the hazard of infection, often referred to as the force of infection. The force of infection reflects the degree of contact with potential for transmission between infected and susceptible individuals. The mathematical relation between the force of infection and effective contact patterns is generally assumed to be subjected to the mass action principle, which yields the necessary information to estimate the basic reproduction number, another key parameter in infectious disease epidemiology.

It is within this context that the Center for Statistics (CenStat, I-Biostat, Hasselt University) and the Centre for the Evaluation of Vaccination and the Centre for Health Economic Research and Modelling Infectious Diseases (CEV, CHERMID, Vaccine and Infectious Disease Institute, University of Antwerp) have collaborated over the past 15 years. This book demonstrates the past and current research activities of these institutes and can be considered to be a milestone in this collaboration.

This book is focused on the application of modern statistical methods and models to estimate infectious disease parameters. We want to provide the readers with software guidance, such as R packages, and with data, as far as they can be made publicly available.

Statistics / Life Sciences,
Medicine, Health Sciences

ISBN 978-1-4614-4071-0



9 781461 440710



► springer.com



Part I Introducing the Concept of the Book

1	Why This Book? An Introduction	3
1.1	Terms for Germs	3
1.2	Models of Infectious Diseases	4
1.3	Where Does This Book Fit in the Field?	5
1.4	A Road Map for This Book	6

Part II Mathematical Models for Infectious Diseases: An Introduction

2	A Priori and A Posteriori Models for Infectious Diseases	11
2.1	The Theory of Happenings	11
2.2	An Example: A Basic Model for HIV/AIDS	13
2.2.1	A Mathematical Model for HIV/AIDS	13
2.2.2	A Statistical Model for the Initial HIV/AIDS Outbreak ..	17
2.3	The Mechanism that Generates the Data	20
2.4	The Basic Epidemiological Parameters	21
2.5	Discussion	23
3	The SIR Model	25
3.1	Introduction	25
3.1.1	The SIR Model	25
3.1.2	The Basic Model Dynamics	27
3.1.3	The Basic Model in R	29
3.1.4	Vaccination in the Basic Model	31
3.1.5	The Basic SIR Model with Vaccination in R	34
3.1.6	The Critical Vaccination Coverage	37
3.2	The SIR Model in Endemic Equilibrium	37
3.2.1	Compartments in the Time Homogeneous SIR Model ...	40

3.2.2	The SIR Model with Constant Force of Infection at Endemic State in R.....	42
3.2.3	The Critical Proportion of Vaccination in the Time Homogeneous SIR Model.....	44
3.3	The Time Homogeneous SIR Model and Serological Data.....	45
3.4	Models with Maternal Antibodies and Latent Periods.....	48
3.5	Transmission Within Multiple Subpopulations.....	51
3.5.1	An SIR Model with Interacting Subpopulations.....	52
3.5.2	Transmission Over Age and Time.....	55
3.5.3	Estimating the WAIFW Matrix.....	57
3.6	Discussion.....	58

Part III Data Sources

4	Data Sources for Modeling Infectious Diseases.....	61
4.1	Serological Data.....	61
4.1.1	Hepatitis A.....	62
4.1.2	Hepatitis B.....	64
4.1.3	Hepatitis C.....	65
4.1.4	Mumps.....	66
4.1.5	Parvovirus B19.....	67
4.1.6	Rubella.....	69
4.1.7	Tuberculosis.....	69
4.1.8	Varicella.....	70
4.1.9	Multisera Data.....	71
4.2	Hepatitis B Incidence Data.....	72
4.3	Belgian Contact Survey.....	75

Part IV Estimating the Force of Infection

5	Estimating the Force of Infection from Incidence and Prevalence....	79
5.1	Serological Data.....	79
5.2	Age-Dependent Force of Infection.....	81
5.3	Modeling Issues.....	82
5.4	Incidence Data.....	84
6	Parametric Approaches to Model the Prevalence and Force of Infection.....	89
6.1	Modeling the Force of Infection: A Historical Perspective.....	89
6.1.1	Polynomial Models.....	89
6.1.2	Nonlinear Models.....	93
6.1.3	Discussion.....	98
6.2	Fractional Polynomial Models.....	99
6.2.1	Motivating Example.....	99
6.2.2	Model Selection.....	101

6.2.3	Constrained Fractional Polynomials	101
6.2.4	Selection of Powers and Back to Nonlinear Models	102
6.2.5	Application to the Data	102
6.2.6	Influence of the Link Function	105
6.2.7	From Fractional Polynomials Back to Nonlinear Models	108
6.3	Discussion	109
7	Nonparametric Approaches to Model the Prevalence and Force of Infection	111
7.1	Nonparametric Approaches	111
7.1.1	The First Nonparametric Approaches	112
7.1.2	Local Estimation by Polynomials	114
7.2	Application to UK Mumps Data	117
7.3	Concluding Remarks	119
8	Semiparametric Approaches to Model the Prevalence and Force of Infection	121
8.1	Semiparametric Approaches	121
8.2	Penalized Splines	122
8.2.1	Penalized Likelihood Framework	125
8.2.2	Generalized Linear Mixed Model Framework	129
8.3	Covariate Effects	132
8.4	Adaptive Spline Smoothing	136
8.5	Synthesis	137
8.6	Non-, Semi- and Parametric Methods to Estimate the Prevalence and Force of Infection: A Case Study	138
9	The Constraint of Monotonicity	141
9.1	Introduction	141
9.2	Piecewise Constant Forces of Infection	142
9.3	Isotonic Regression	143
9.3.1	The Pool Adjacent Violator Algorithm	143
9.3.2	Keiding (1991)	145
9.3.3	“Smooth Then Constrain”	146
9.4	P-spline Regression with Shape Constraints	146
9.5	Concluding Remarks	148
10	Hierarchical Bayesian Models for the Force of Infection	149
10.1	Introduction	149
10.2	Exploratory Data Analysis	150
10.3	Hierarchical Bayesian Models for the Force of Infection	151
10.3.1	Nonlinear Hierarchical Model	151
10.3.2	Fitting Farrington’s Model in Winbugs	152
10.3.3	Hierarchical Log-Logistic Model	154
10.3.4	Fitting the Log-Logistic Model in Winbugs	155
10.3.5	Model Selection	155
10.3.6	Application to the Data	156

10.4	Hierarchical Nonparametric Model	158
10.4.1	Hierarchical Beta/Binomial Model	158
10.4.2	Fitting the Model in Winbugs	160
10.4.3	Application to the Data	162
10.5	Discussion	163
11	Modeling the Prevalence and the Force of Infection	
	Directly from Antibody Levels	167
11.1	Serological and Current Status Data	167
11.2	The Threshold Approach	170
11.3	A Direct Approach	171
11.4	Application to VZV Data	173
11.4.1	Fitting a Mixture	173
11.4.2	Fitting the Mean Antibody Levels as a Function of Age	174
11.4.3	Combining the Mixture with the Least Squares Regression Fit	174
11.5	Modeling the Force of Infection Directly from Antibody Titers Using Hierarchical Mixture Models	175
11.5.1	Model Formulation	176
11.5.2	Age-Dependent Mixture Probability	177
11.5.3	Application to the Data	178
11.5.4	Determining the Current Infection Status	179
11.5.5	Implementation in Winbugs	181
11.6	Concluding Remarks	183
12	Modeling Multivariate Serological Data	185
12.1	Introduction	185
12.2	Marginal and Conditional Models	186
12.2.1	The Bivariate Dale Model Applied to Airborne Infections	188
12.2.2	The Marginal, Conditional and Joint Force of Infection ..	191
12.3	Individual Heterogeneity	194
12.4	Concluding Remarks	199
13	Estimating Age-Time Dependent Prevalence and Force of Infection from Serial Prevalence Data	201
13.1	Introduction	201
13.2	Proportional Hazards Model	204
13.2.1	Semiparametric Model	204
13.2.2	Non-parametric Model	204
13.2.3	Example I: Tuberculosis	205
13.3	Non-proportional Hazards Model	207
13.4	The Tuberculosis Data Revisited	207
13.5	Example II: Hepatitis A	210
13.6	Monotonicity	214
13.7	Concluding Remarks	215

Part V Estimating Mixing Patterns and the Basic Reproduction Number

14	Who Acquires Infection from Whom? The Traditional Approach ...	219
14.1	Who Acquires Infection from Whom?	219
14.2	Estimation from Serological Data	220
14.2.1	The Discretized Mass Action Principle.....	223
14.2.2	Imposing Mixing Patterns: The Traditional Approach of Anderson and May (1991).....	224
14.2.3	Exploiting an Underlying Continuous Mixing Surface ...	228
14.3	Topics in Estimating WAIFW Matrices	230
14.3.1	Model Selection	230
14.3.2	Lifelong Immunity	231
14.4	Concluding Remarks	232
15	Informing WAIFW with Data on Social Contacts	233
15.1	Estimation from Serological Data and Data on Social Contacts ...	233
15.1.1	The Social Contact Hypothesis	233
15.1.2	Refinements to the Social Contact Data Approach	237
15.2	The Bootstrap and Multi-model Inference	239
15.2.1	The Nonparametric Bootstrap.....	240
15.2.2	Multi-model Inference.....	240
15.3	Concluding Remarks	242

Part VI Integration of Statistical Models into the Mathematical Model Framework

16	Integrating Estimated Parameters in a Basic SIR Model	247
16.1	Introduction.....	247
16.2	Application	249
16.2.1	Model Comparison	251
16.2.2	Uncertainty.....	252
16.2.3	Vaccination.....	253
16.3	Discussion	255
A	A Note on the R Environment	257
A.1	Data Structures in R.....	257
A.2	Reading and Writing Data	258
A.3	Graphical Procedures	258
A.4	Statistical Models in R	259
A.5	Packages	259
A.6	References	260
B	Statistical Inference	261
B.1	Maximum Likelihood Estimation and Likelihood Inference.....	261

B.2	Generalized Linear Models	268
B.3	Profile Likelihood and Other Likelihoods.....	271
B.3.1	Profile Likelihood Estimation	271
B.3.2	Profile Likelihood Confidence Intervals	273
B.4	The Bootstrap.....	275
B.4.1	Bootstrap Confidence Intervals	276
B.4.2	Nonparametric Bootstrap.....	276
B.4.3	Parametric Bootstrap	278
B.5	Bayesian Methodology.....	279
References.....		285
Index.....		295