Validation of a generic Bayesian method for prevalence estimation under misclassification

CURRENT STATUS: UNDER REVIEW

BMC Public Health BMC Series

Matthias Flor
German Federal Institute for Risk Assessment

matthias.flor@bfr.bund.de*Corresponding Author*
ORCiD: https://orcid.org/0000-0003-0429-7457

Michael Weiβ
German Federal Institute for Risk Assessment

Thomas Selhorst
German Federal Institute for Risk Assessment

Christine Müller-Graf
German Federal Institute for Risk Assessment

Matthias Greiner
Bundesinstitut fur Risikobewertung Standort Marienfelde

DOI:
10.21203/rs.2.23469/v1

SUBJECT AREAS
Statistical Epidemiology

KEYWORDS
prevalence estimation, imperfect diagnostic test, misclassification, Bayesian prevalence estimate, Rogan-Gladen estimate, diagnostic sensitivity, diagnostic specificity

Abstract

Background: Various methods exist for statistical inference about a prevalence that consider misclassifications due to an imperfect diagnostic test. However, traditional methods are known to suffer from censoring of the prevalence estimate and the confidence intervals constructed around the point estimate, as well as from under-performance of the confidence intervals' coverage.

Methods: In this study, we used simulated data sets to validate a Bayesian prevalence estimation method and compare its performance to frequentist methods, i.e. the Rogan-Gladen estimate for prevalence, RGE, in combination with several methods of confidence interval construction. Our performance measures are (i) bias of the point estimate against the simulated true prevalence and (ii) coverage and length of the confidence interval, or credible interval in the case of the Bayesian method.

Results: Across all data sets, the Bayesian point estimate and the RGE produced similar bias distributions with slight advanteges of the former over the latter. In addition, the Bayesian estimate did not suffer from the RGE's censoring problem at zero or unity. With respect to coverage performance of the confidence and credible intervals, all of the traditional frequentist methods exhibited strong under-coverage, whereas the Bayesian credible interval as well as a newly developed frequentist method by Lang and Reiczigel performed as desired, with the Bayesian method having a very slight advantage in terms of interval length.

Conclusion: The Bayesian prevalence estimation method should be prefered over traditional frequentist methods. An acceptable alternative is to combine the Rogan-Gladen point estimate with the Lang-Reiczigel confidence interval.

Full Text

Due to technical limitations, full-text HTML conversion of this manuscript could not be completed. However, the manuscript can be downloaded and accessed as a PDF.

Figures

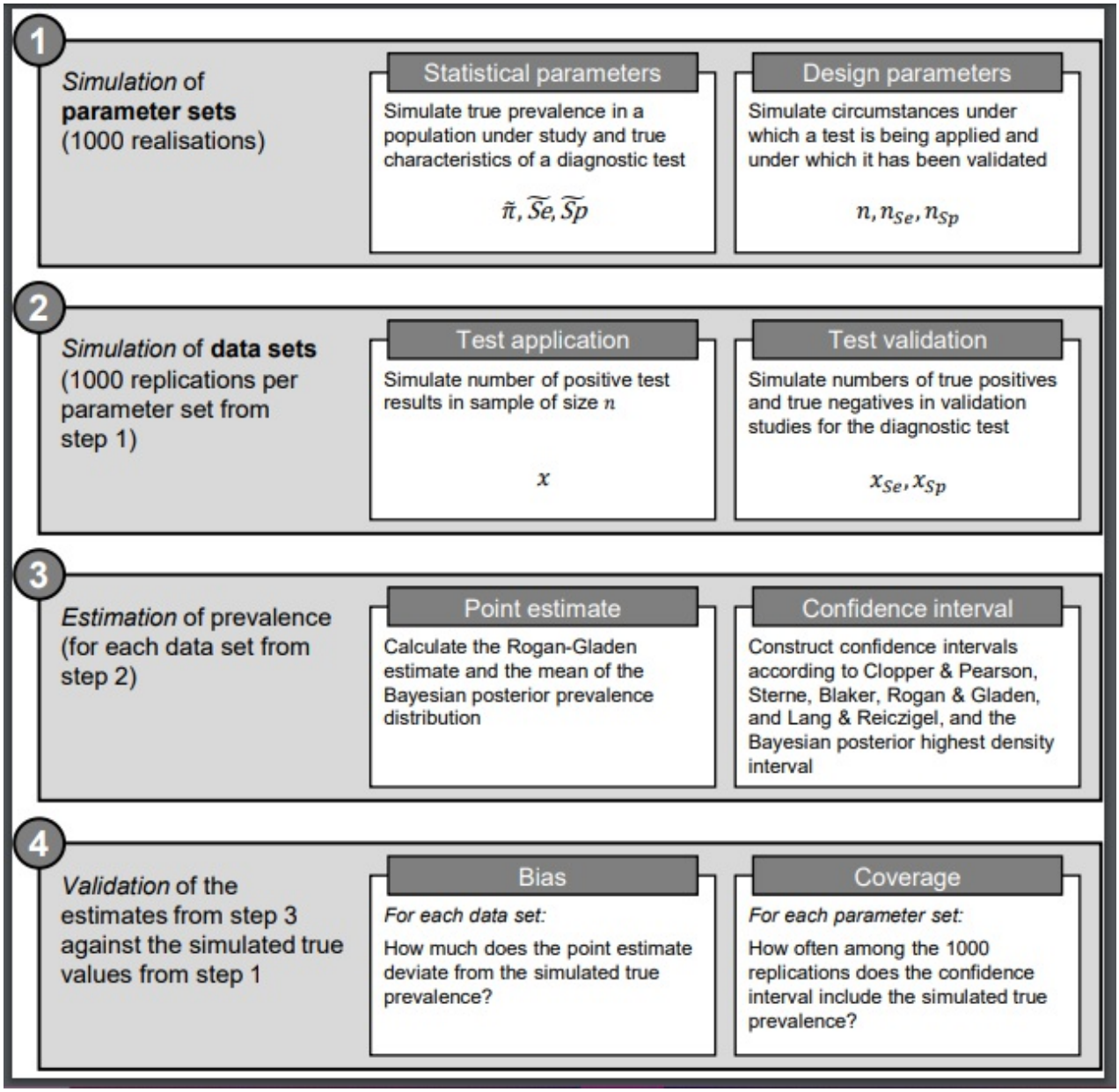

Figure 1

The four steps of the validation study. (1) Simulation of parameter sets to generate true values, (2) simulation of data sets, (3) estimation of prevalence and calculation of confidence intervals, and (4) validation of the estimates against simulated true values.

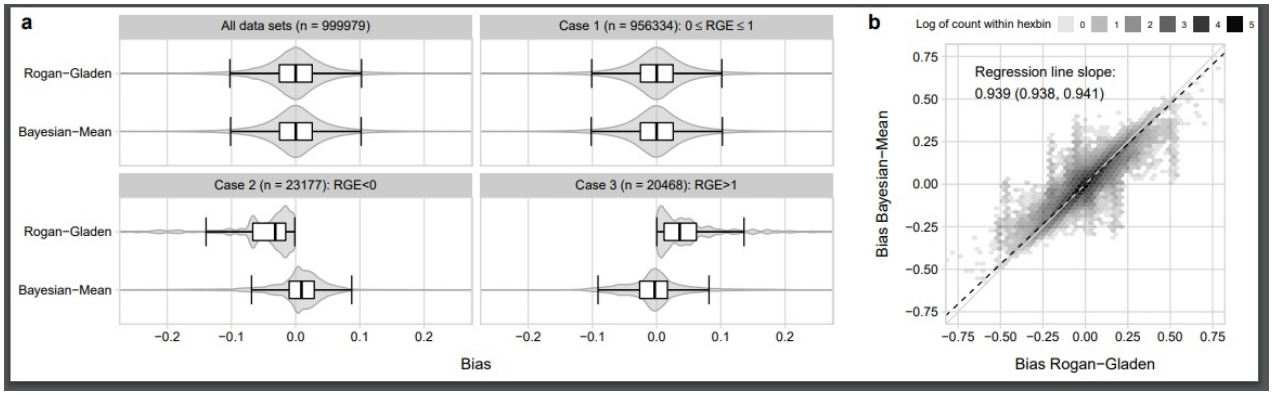

Figure 2

Point estimate bias. (a) Bias distributions of the Rogan-Gladen point estimate and the Bayesian estimate (MCMC mean) across all data sets (top left), and across data sets as classified according to the non-censored Rogan-Gladen estimate (case 1, top right). The Bayesian estimator shows adequate bias distributions for the data sets with a censored RGE (cases 2 and 3, bottom row). (b) Comparison of the biases of the Bayesian mean and the Rogan-Gladen estimate for all data sets. Hexagonal binning is used to deal with overplotting, and the hex gray scale codes for the number of data sets that fall within it. The dashed black line shows a Deming regression of the Bayesian bias on the Rogan-Gladen bias. Its slope is 0.939 with a confidence interval of (0.938, 0.941).

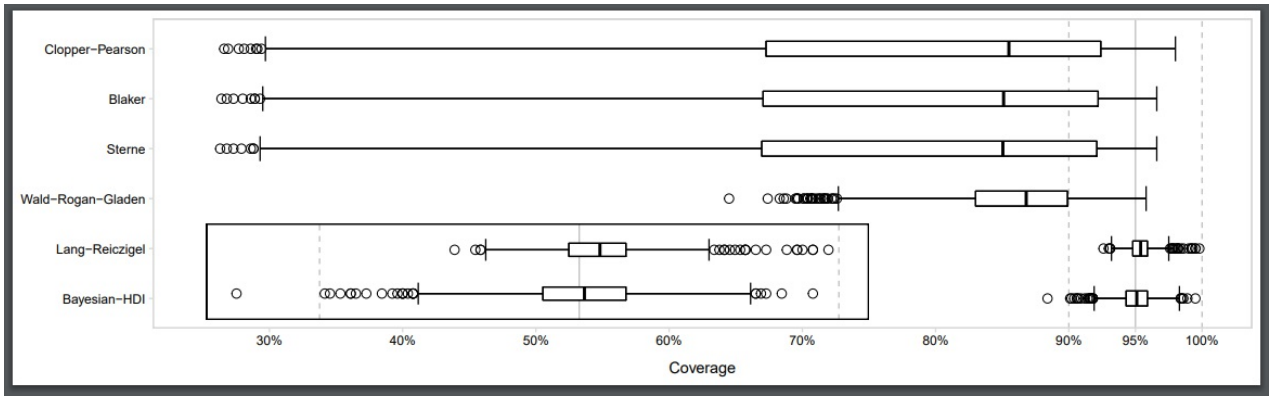

Figure 3

Confidence interval coverage. Coverage by parameter set for several methods of 95% confidence interval (CI) construction. The nominal coverage of 95% is marked by a solid gray line, and the dahed lines mark coverage values of 90% and 100%. The traditional CI's (Clopper-Pearson, Blaker, Sterne, and Wald-Rogan-Gladen) all exhibit significant under coverage. The bottom left inset shows the Lang-Reiczigel CI's and the Bayesian HDI's coverage distributions in more detail. It reveals the the Lang-Reiczigel CI tends to have some over coverage and the Bayesian HDI's coverage appears to be more more symmetrical around the nominal 95% value.

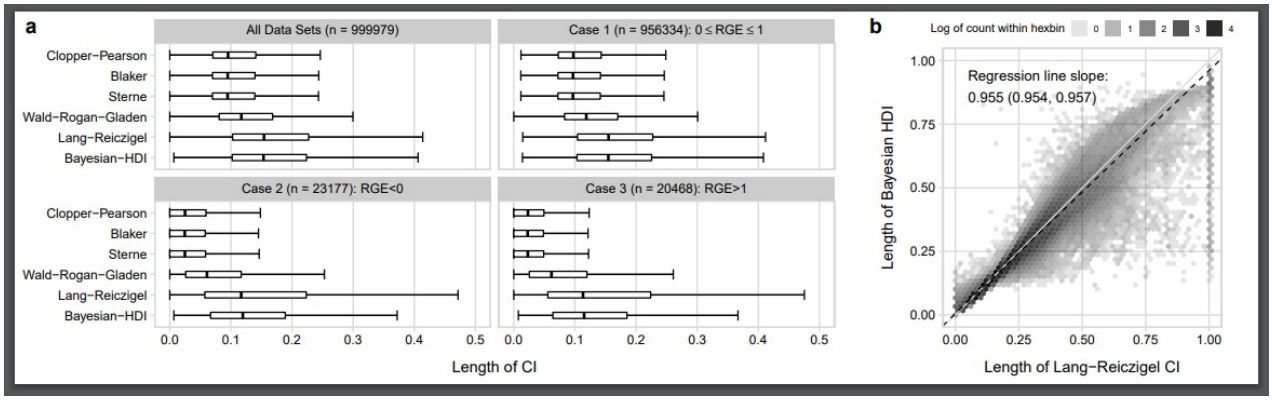

Figure 4

Confidence interval length. (a) Confidence (credible) interval length for several methods across all data sets (top left), and across data sets as classified according to the non-censored Rogan-Gladen estimate. Traditional CI's are generally shorter than the Lang-Reiczigel CI and the Bayesian HDI. (b) Comparison of the lengths of the Bayesian HDI and the Lang-Reiczigel CI for all data sets. Hexagonal binning is used to deal with overplotting, and the hex gray scale codes for the number of data sets that fall within it. A Deming regression is shown as a dashed black line, its slope is 0.955 with a confidence interval of (0.954, 0.957).