Estimating the number of usability problems affecting medical devices: modelling the discovery matrix

CURRENT STATUS: UNDER REVISION

BMC Medical Research Methodology BMC Series

Vincent Vandewalle
Inria Centre de recherche Lille Nord Europe

Alexandre Caron
Univ. Lille, CHU Lille, EA2694, F-59000 Lille, France

alexandre.caron2@univ-lille.fr*Corresponding Author*
ORCiD: https://orcid.org/0000-0002-9872-0633

Coralie Delettrez
CHU LILLE

Sylvia Pelayo
Centre d'Investigation Clinique - Innovation Technologique Lille

Alain Duhamel
Universite de Lille

Benoit Dervaux
Universite de Lille

DOI:
10.21203/rs.2.16958/v1

SUBJECT AREAS
Health Economics & Outcomes Research

KEYWORDS
usability study, medical device, missing data, Bayesian statistics, maximum likelihood

Abstract

Background Usability studies of medical devices are mandatory for market access. The studies' goal is to identify and eliminate usability problems that could cause harm the user or limit the device's effectiveness. In practice, human factor engineers study participants under actual conditions of use and list the problems encountered. This results in a binary discovery matrix in which each row corresponds to a participant, and each column corresponds to a usability problem. One of the main challenges in usability studies is estimating the total number of problems, in order to assess the completeness of the discovery process. Today's margin-based methods fit the column sums to a binomial model of problem detection. However, the discovery matrix actually observed is truncated because of undiscovered problems, which corresponds to fitting the marginal sums without the zeros. Margin-based methods fail to overcome the bias related to truncation of the matrix. The objective of the present study was to develop and test a matrix-based method for estimating the total number of usability problems.

Methods The matrix-based method models the likelihood of the discovery matrix, and allows one to account for all the available information. It also circumvents a drawback of margin-based methods by simultaneously estimating two unknown parameters: the probability of problem detection and the total number of problems. Furthermore, the matrix-based method takes account of a heterogeneous probability of detection, which better reflects a real-life setting. As suggested in the usability literature, a logit-normal prior for the probability of detection is selected.

Results We assessed the matrix-based method's performance in a range of settings reflecting real-life usability studies and with both homogeneous and heterogeneous probabilities of problem detection. In our simulations, the matrix-based method improved the estimation of the number of problems (in terms of bias, consistency, and coverage probability of the confidence interval) in a wide range of settings. We also applied our method to real data from a usability study of infusion pumps.

Conclusions Our method should be applied by regulators and device manufacturers to estimate the number of usability problems using the set of statistical routines provided.

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

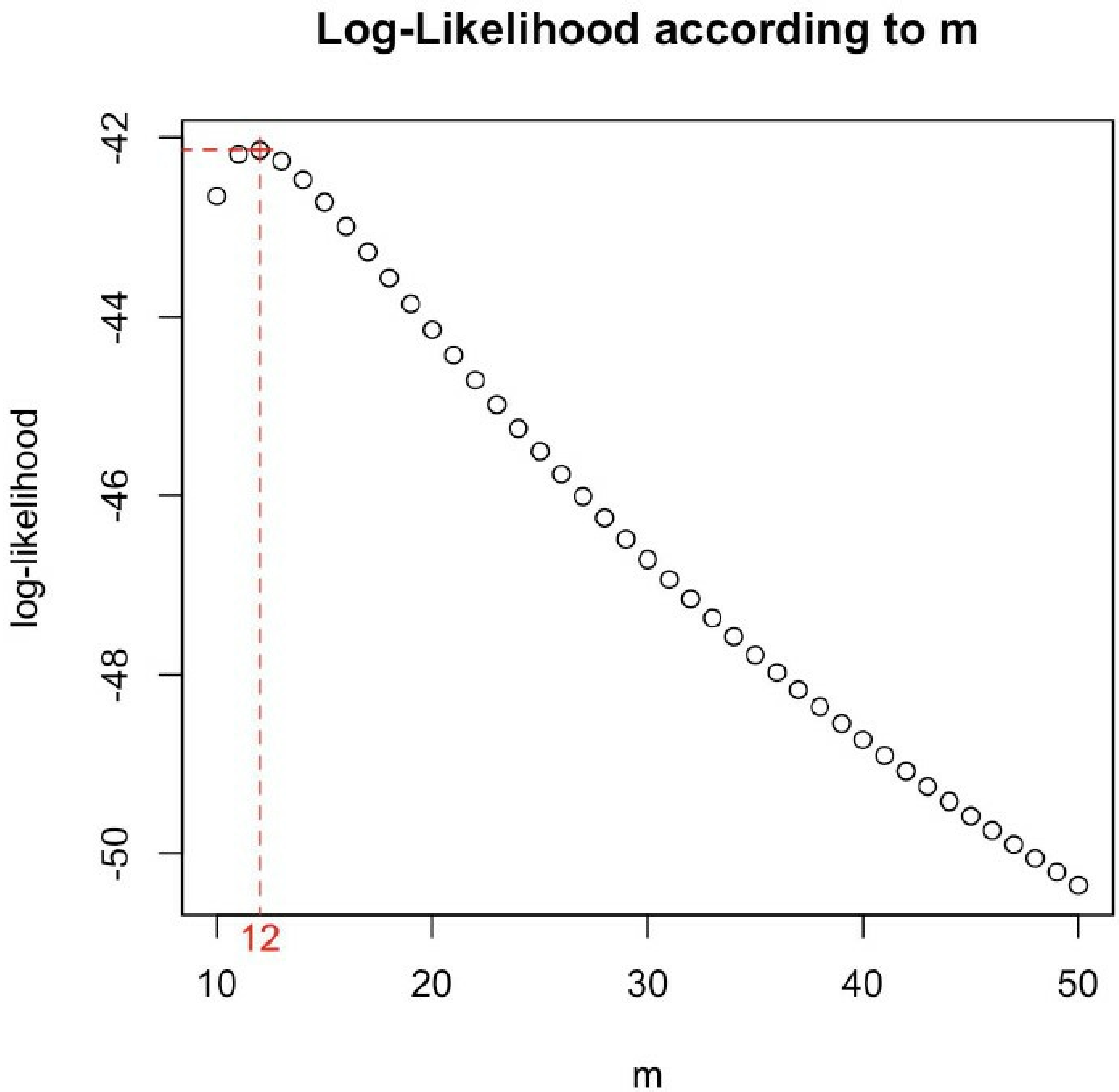

Figure 1

Log-likelihood according to the total number of problems (m), based on the observed discovery matrix (see d in page 6).. The maximum log-log likelihood is reached for m=12. Since the number of observed problems is j=10, the maximum likelihood estimate predicts that 2 usability problems remain undetected.

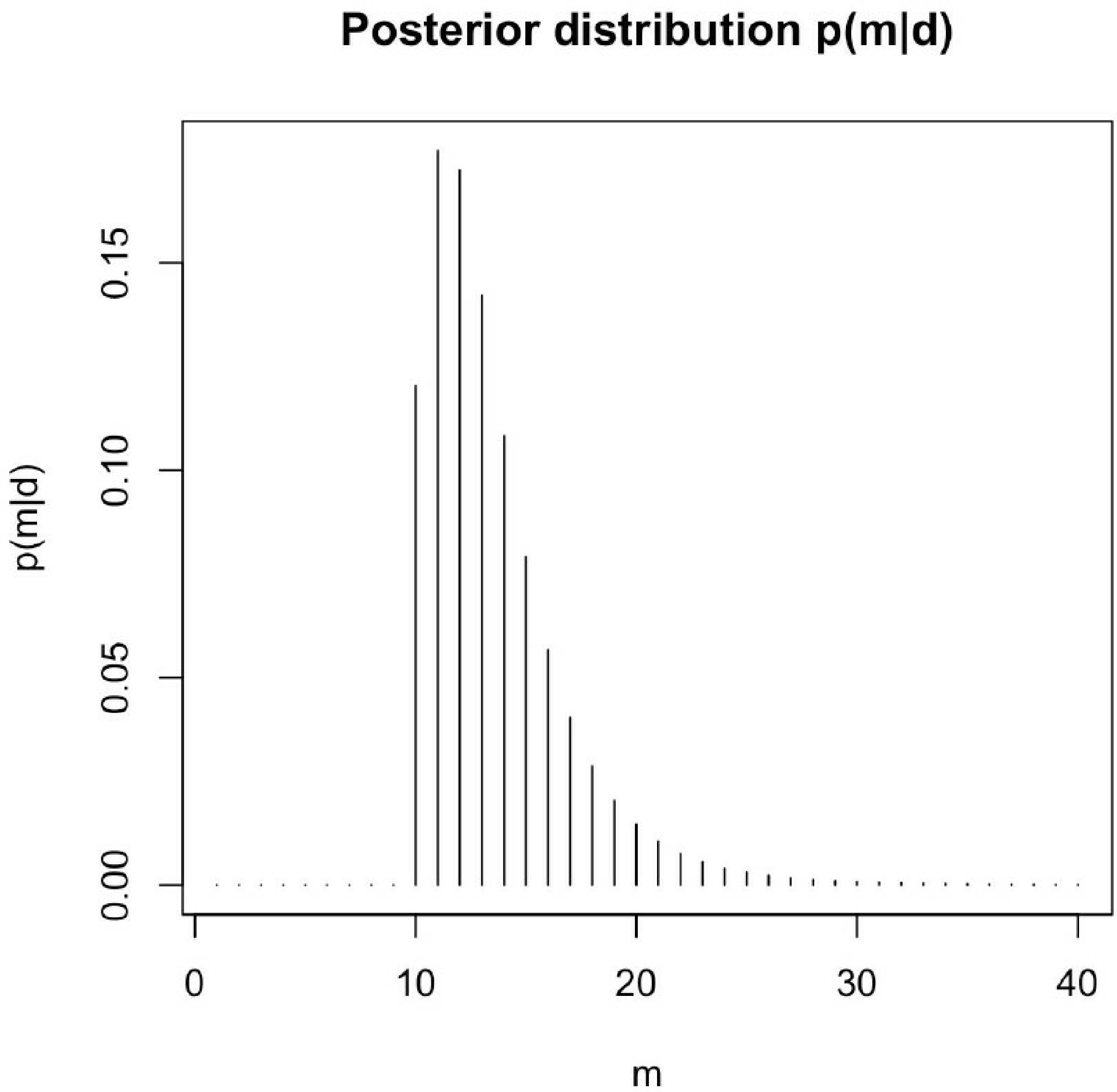

Figure 2

Posterior distribution of m, given the observed discovery matrix (see d in page 6). The distribution mode is at m=11. Since the number of observed problems is j=10, the Bayesian estimate predicts that 1 usability problem has yet to be detected.

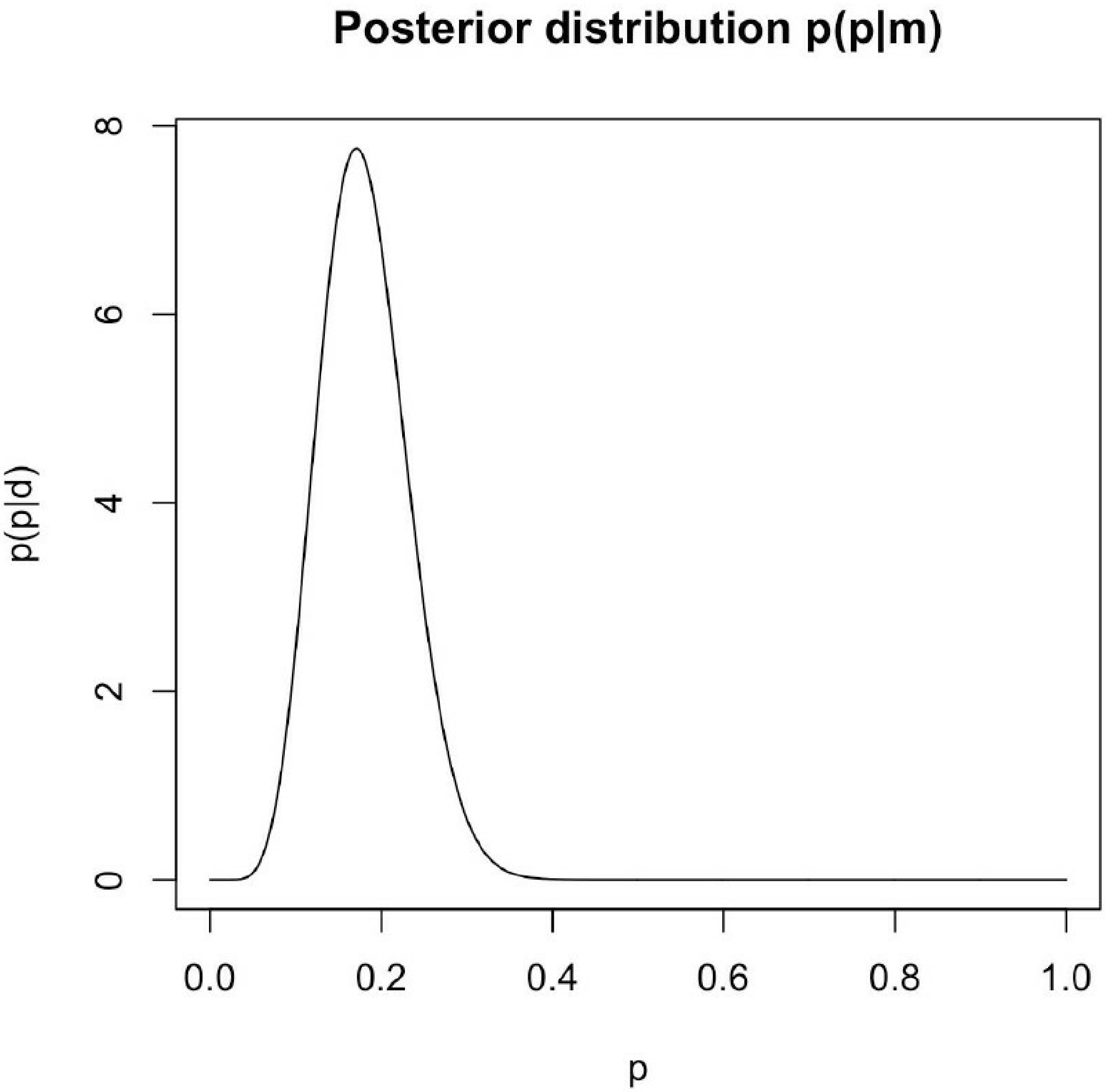

Figure 3

Posterior distribution of p, given the observed discovery matrix (see d in page 6).

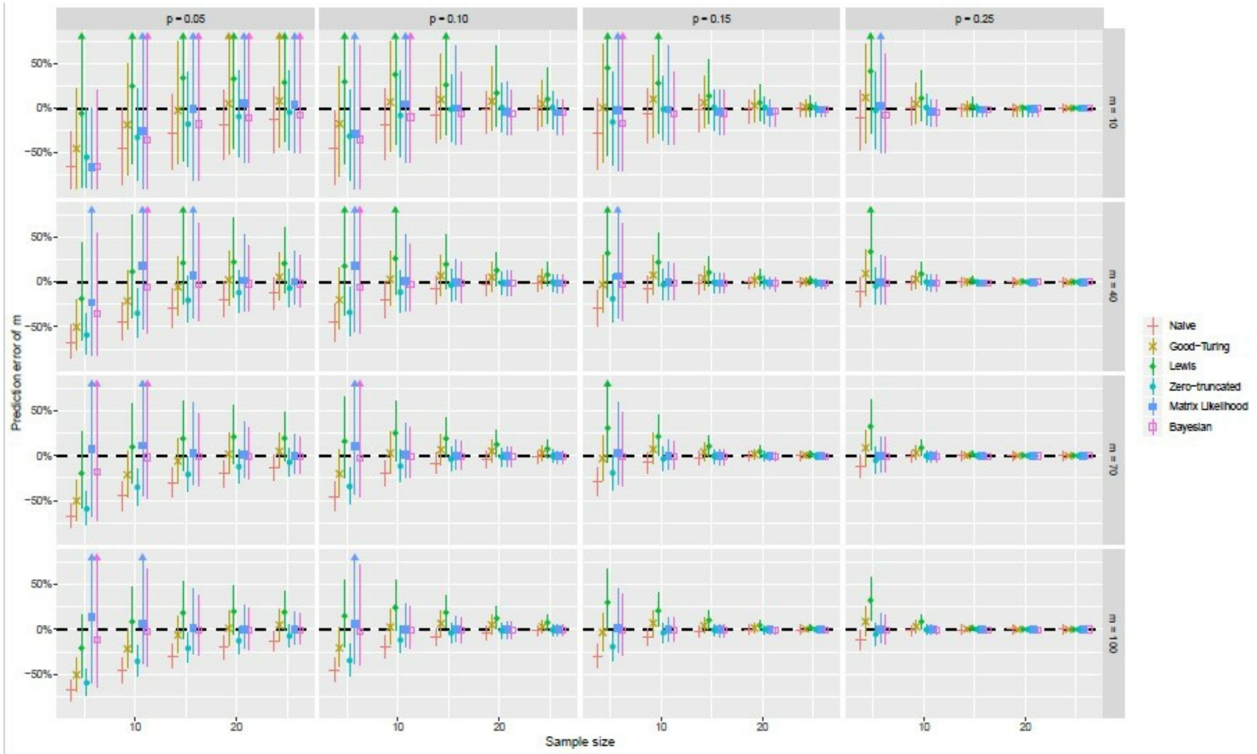

Figure 4

Prediction error for m (mean and 95% fluctuation interval, in %) as a function of the sample size (n), for 6 estimators. The results are presented for various probabilities of detection (p, columns) and various numbers of usability problems (m, rows). The dashed line represents the “true” m. The upper boundaries of credible intervals that exceed 100% are indicated by △.

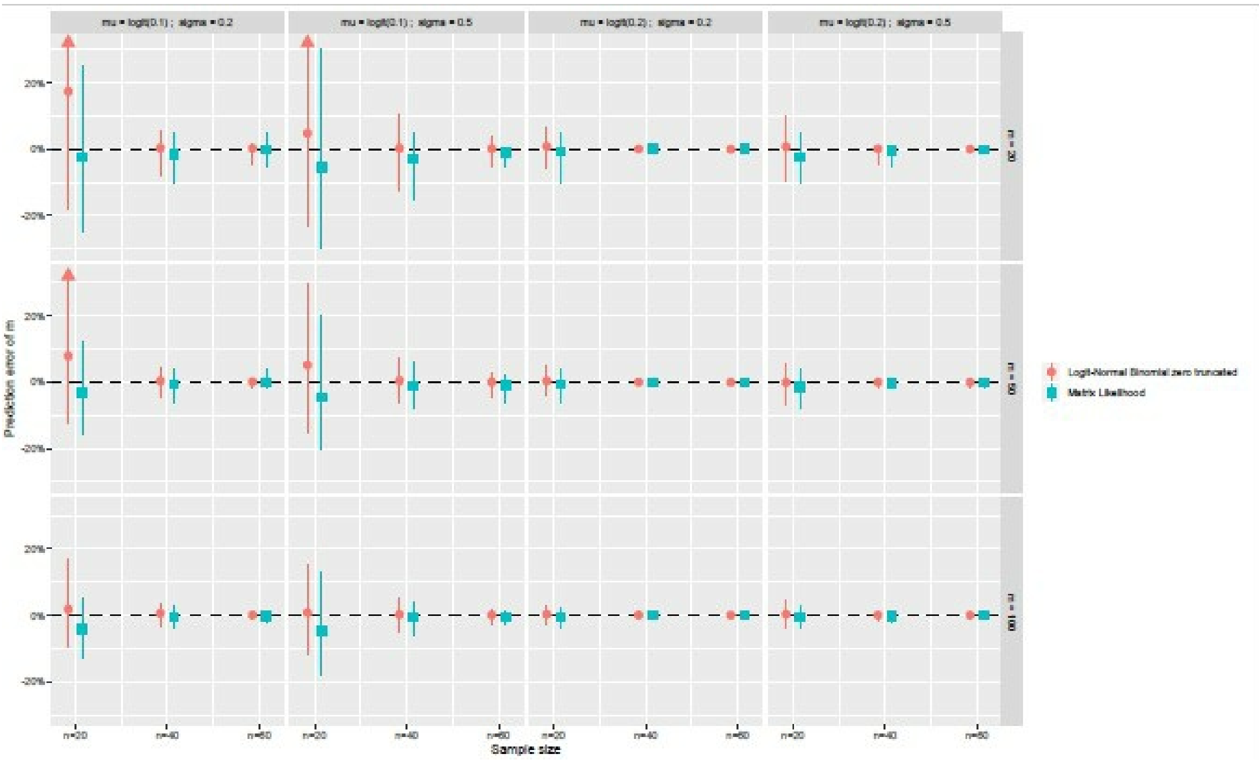

Figure 5

Prediction error of m (mean and 95% fluctuation interval, in %) as a function of the sample size (n) in the context of a heterogeneous probability of problem detection. The results are presented for various probabilities of problem detection ((μ,σ), columns) and various numbers of usability problems (m, rows). The dashed line represents the “true” m. The upper boundaries of the credible intervals that exceed 100% are indicated by △.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

Additional file 1.rar