Statistical methods for the meta-analysis of full ROC curves

Oliver Kuss, Annika Hoyer

German Diabetes Center,
Leibniz Institute for Diabetes Research at Heinrich Heine University Düsseldorf,
Institute for Biometry and Epidemiology

October 15, 2015

DDZ
Deutsches Diabetes-Zentrum

Outline

Introduction

Idea I: Meta-Regression in the standard bivariate model

Idea II: Bivariate time-to-event model for interval-censored data

Discussion

Deutsches Diabetes-Zentrum

Introduction

Meta-analysis for diagnostic accuracy trials

- Meta-analysis for intervention studies is well established today
- In contrast, meta-analysis for diagnostic accuracy trials has been a vivid research area in recent years
- Reason: Increased complexity of diagnostic trials with their bivariate outcome of sensitivity and specificity

DDZ Deutsches Diabetes-Zentrum

Meta-analysis for diagnostic accuracy trials

- Additional challenge: Single studies report a full ROC curve with several pairs of sensitivity and specificity, each one for a different threshold
- Still greater challenge: Values and numbers of thresholds can vary between single studies

Deutsches Diabetes-Zentrum

Meta-analysis of diagnostic accuracy trials - An example

Study	Cutpoint	Sensitivity	Specificity
1	5.3	89/115	1382/1684
1	5.5	72/115	1558/1684
1	5.6	65/115	1602/1684
2	5.3	207/252	2704/5865
2	5.4	196/252	3308/5865
...	...	...	...

Deutsches Diabetes-Zentrum

Meta-analysis of diagnostic accuracy trials - An example

DIABETICMedicine

DOI: 10.1111/j.1464-5491.2007.02106.x

HbA_{1c} as a screening tool for detection of Type 2 diabetes: a systematic review

C. M. Bennett, M. Guo and S. C. Dharmage

Department of Public Health, School of Population Health, The University of Melbourne, Australia

DIABETES/METABOLISM RESEARCH AND REVIEWS
Diabetes Metab Res Rev 2013; **29**: 680–692.
Published online in Wiley Online Library (wileyonlinelibrary.com) **DOI:** 10.1002/dmrr.2445

RESEARCH ARTICLE

Use of high-normal levels of haemoglobin A_{1C} and fasting plasma glucose for diabetes screening and for prediction: a meta-analysis

DDZ Deutsches Diabetes-Zentrum

Meta-analysis of diagnostic accuracy trials - An example

- Population-based screening for type 2 diabetes mellitus
- Two systematic reviews (Kodama, 2013 and Bennett, 2007) report on 38 single studies (one pair of sensitivity and specificity from each study) to assess HbA1c as diagnostic marker
- **However:** Intensified search yields 124 pairs of sensitivity and specificity for 26 different HbA1c thresholds

DDZ Deutsches Diabetes-Zentrum

Meta-analysis of diagnostic accuracy trials - An example

- An analysis using only the original 38 pairs discards more than 70% of the available observations
- Even worse: Using only the original 38 pairs would also ignore the different HbA1c values that were used as thresholds

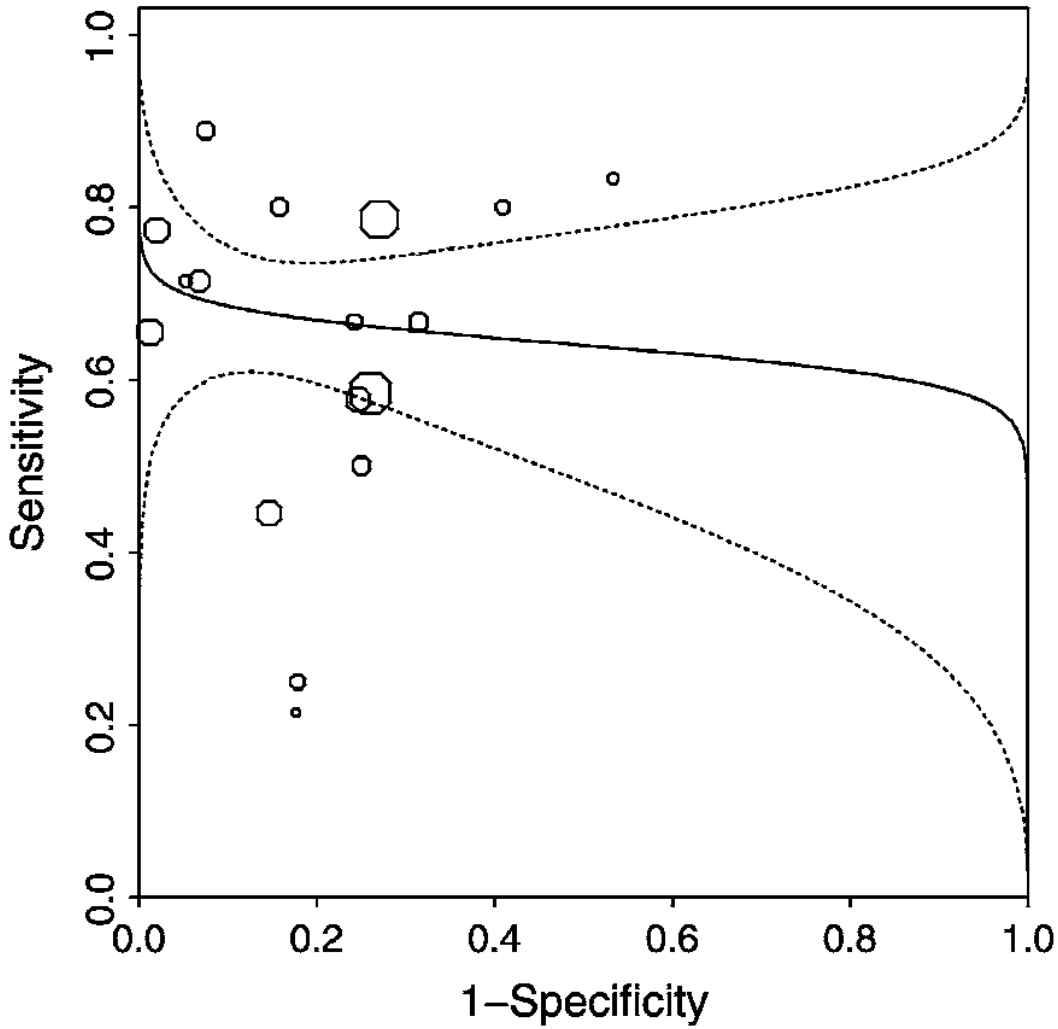
Deutsches Diabetes-Zentrum

Meta-analysis of diagnostic accuracy trials - The paradox of the standard SROC

- A summary ROC curve from the original 38 pairs explicitly ignores the ROC information from the single studies and ...
- '...is in principle unidentifiable' [Rücker/Schumacher, 2010]'
- '...cannot be interpreted as a kind of average curve or a curve typical for the study-specific ROC curves. It can have a shape that is very different from the study-specific shapes.' [Arends et al., 2008]

Deutsches Diabetes-Zentrum

Meta-analysis of diagnostic accuracy trials - The paradox of the standard SROC



[Chu/Guo, 2009]

DDZ Deutsches Diabetes-Zentrum

Meta-analysis of diagnostic accuracy trials - Recommendations

- '... although several of these methods allow for different test thresholds to be used across the primary studies, none have been used to incorporate threshold values explicitly, a notable limitation.' (Sutton, 2008)
- 'If more than one threshold is reported per study, this has to be taken into account in the quantitative analyses.' (Trikalinos, 2012)

DDZ Deutsches Diabetes-Zentrum

Meta-analysis of diagnostic accuracy trials - Current approaches using the full information from each study

- Good news: Methods for the meta-analysis of full ROC curves have been proposed [Kester, 2000; Dukic, 2003; Poon, 2004; Bipat, 2007; Hamza, 2009; Putter, 2010; Martinez-Camblor, 2014; Riley, 2014].

DDZ Deutsches Diabetes-Zentrum

Meta-analysis of diagnostic accuracy trials - Current approaches using the full information from each study

Not so good news: Each of these methods has at least one of the following disadvantages:

- The method constitutes a two-step approach and estimation uncertainty from the first step is ignored in the second step
- The number of thresholds has to be identical across all studies
- The concrete values of the thresholds are ignored
- The method assumes a fixed-effect model
- The method is not applicable with extreme values

DDZ Deutsches Diabetes-Zentrum

Idea I: Meta-Regression in the standard bivariate model

Bivariate logistic regression model with random effects

Sensitivity $Se = \frac{TP}{TP+FN}$ and specificity $Sp = \frac{TN}{TN+FP}$

(TP, FN, TN, FP represent the number of true positives, false negatives, true negatives and false positives, respectively)

$$TP_i \mid Se_i \sim Binomial(TP_i + FN_i, Se_i),\ logit(Se_i) = \mu + \phi_i$$

$$TN_i \mid Sp_i \sim Binomial(TN_i + FP_i, Sp_i),\ logit(Sp_i) = \nu + \psi_i$$

$$\begin{pmatrix} \phi_i \\ \psi_i \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\phi^2 & \rho\sigma_\phi\sigma_\psi \\ \rho\sigma_\phi\sigma_\psi & \sigma_\psi^2 \end{pmatrix}\right]$$

$logit(p) = log(p/(1-p)), \mu, \nu$ intercepts for $logit(Se_i), logit(Sp_i)$

DDZ Deutsches Diabetes-Zentrum

Extend to meta-regression model by including covariates

Modify the GLMM including covariates:

$$logit(Se_i) = \mu + \alpha x_j + \phi_i$$

$$logit(Sp_i) = \nu + \beta x_j + \psi_i$$

where the x_j are the values of the covariate (in our case: HbA1c) that denotes the (potentially different but maybe equal) values of the threshold j, where the number of thresholds can be different between studies $j = 1, \ldots, J_i$.

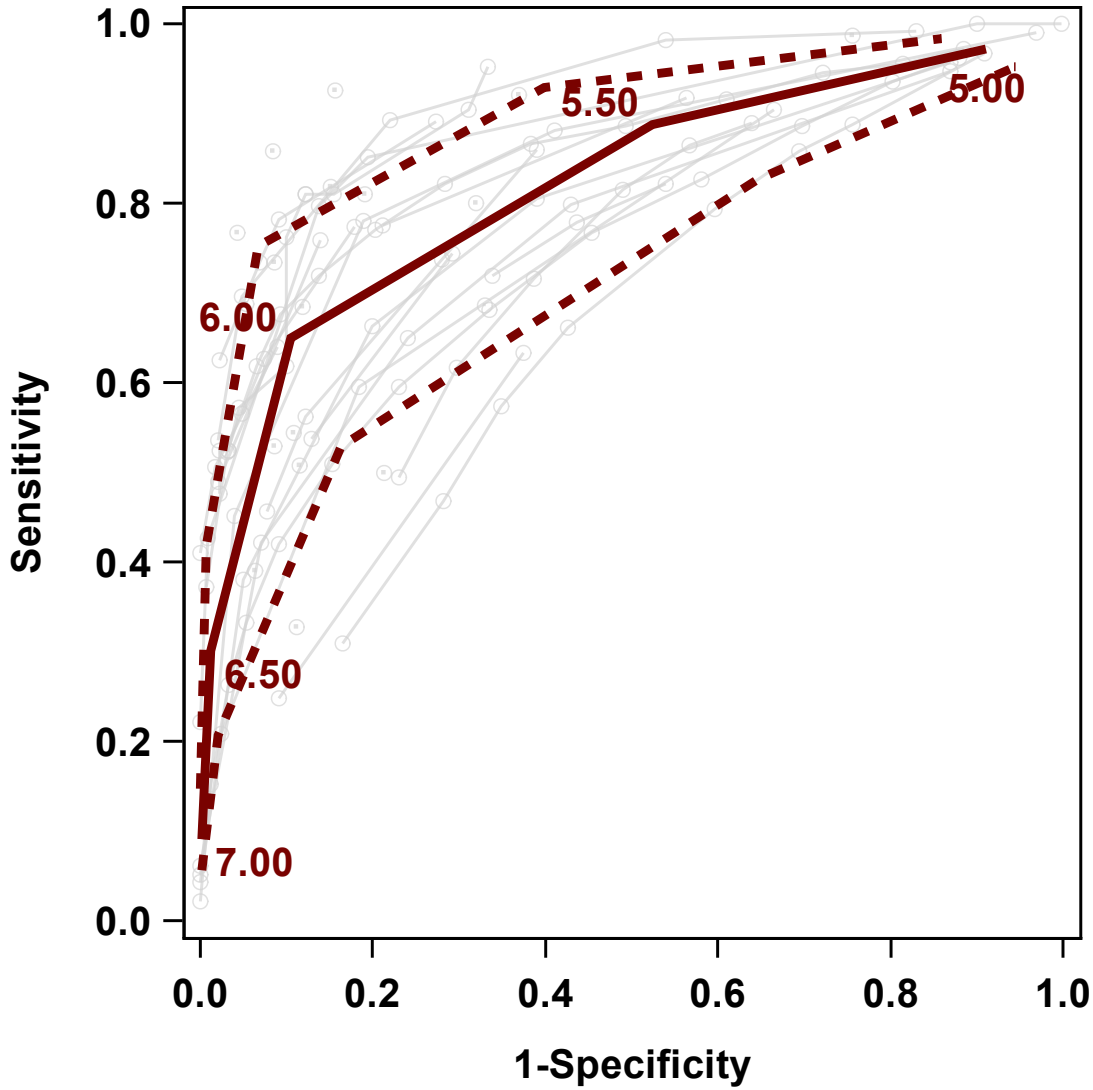
DDZ Deutsches Diabetes-Zentrum

Parameter estimation

- The following parameters have to be estimated:
  $\mu, \nu, \sigma^2_\phi, \sigma^2_\psi, \rho, \alpha, \beta$
- Estimation method: PQL (SAS PROC GLIMMIX)
- However, we are not interested in the model parameters, but in the summary ROC curve
- That is, predict sensitivity and specificity at given thresholds by the BLUP principle (LSMESTIMATE statement, GLIMMIX)

DDZ Deutsches Diabetes-Zentrum

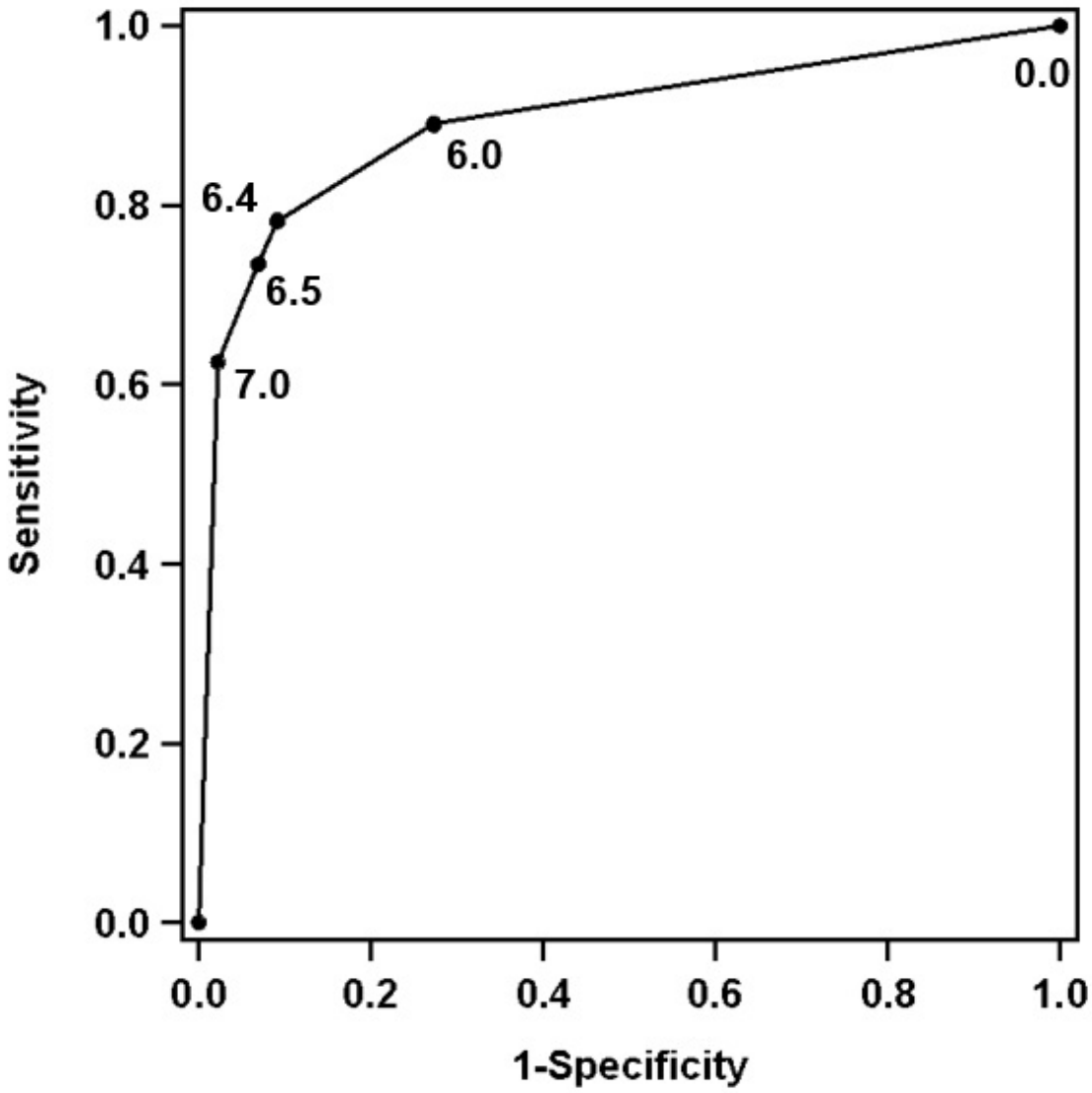
Results for the HbA1c example



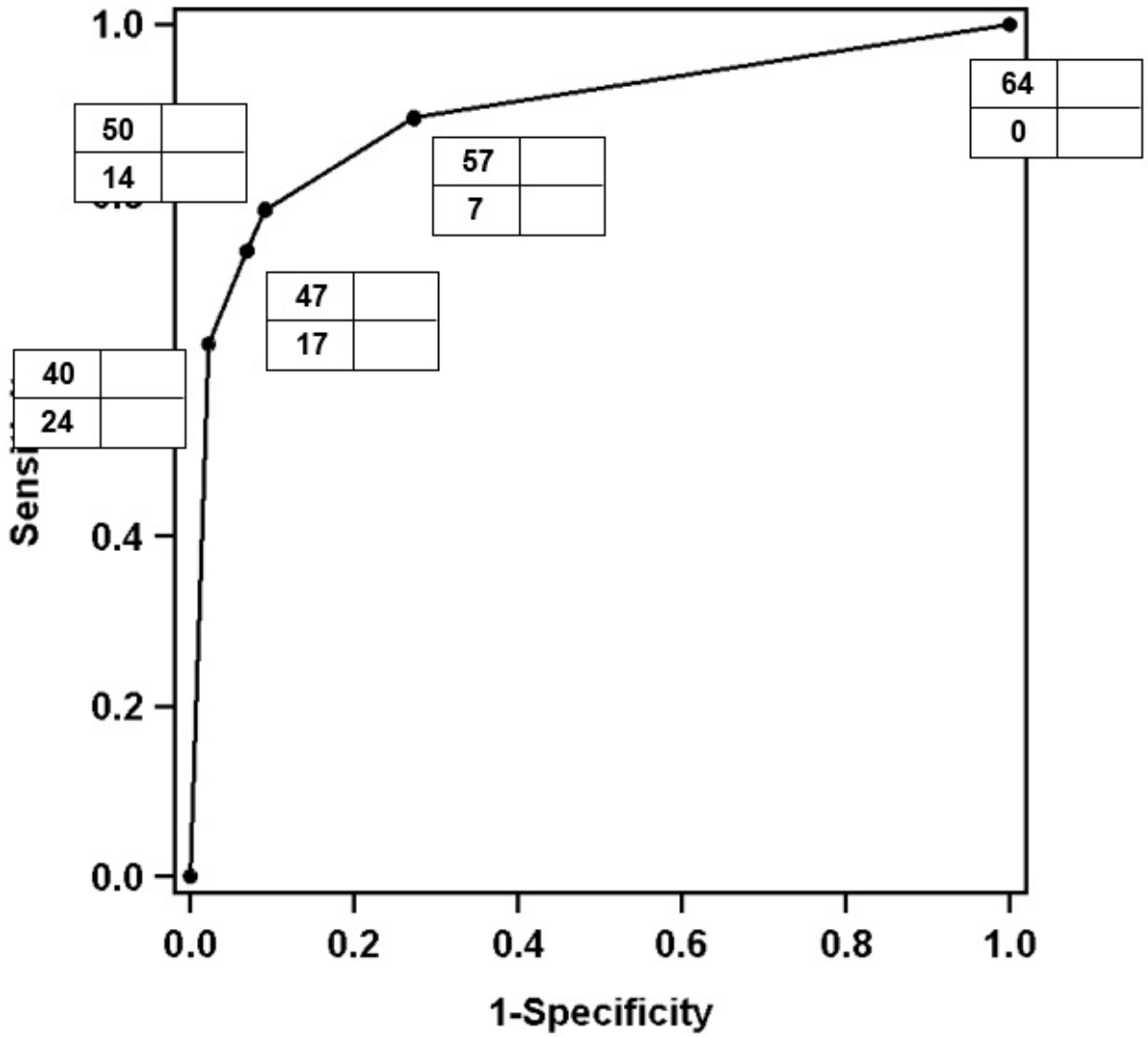
Deutsches Diabetes-Zentrum

Idea II: Bivariate time-to-event model for interval-censored data

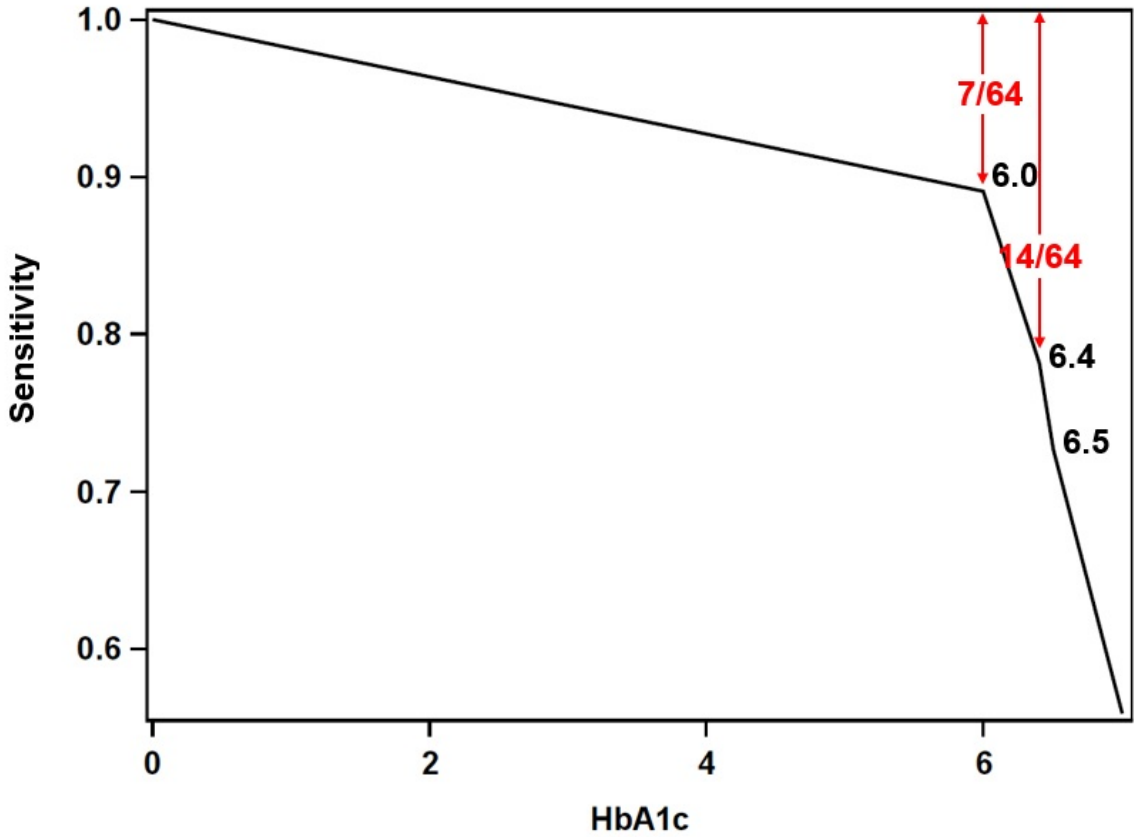
ROC curve from Zhang, 2010



ROC curve from Zhang, 2010



’Life-table’ from Zhang, 2010



Fundamental insight

An ROC curve can be considered a bivariate time-to-event model for interval-censored data!

Deutsches Diabetes-Zentrum

Bivariate time-to-event model for interval-censored data - Notation

- Given K observations from a continuous variable Y (in our case: HbA1c) as lying in intervals $(y_k^L, y_k^R]$ with $\Delta_k = y_k^R - y_k^L$
- Y has density $f(y, \mu, \theta)$ with corresponding cdf $F(y, \mu, \theta)$; μ is a location and ϕ a scale parameter
- In general, three types of censoring can occur:

Type of Censoring	Interval	Contribution to LogL
Left	$y_k^L = 0, y_k^R \neq \infty$	$F(y_k^R; \mu, \phi)$
Interval	$y_k^L \neq 0, y_k^R \neq \infty$	$[f(y_k; \mu, \phi)\Delta_k]$
Right	$y_k^L \neq 0, y_k^R = \infty$	$[1 - F(y_k^L; \mu, \phi)]$

DDZ Deutsches Diabetes-Zentrum

Bivariate time-to-event model for interval-censored data - Densities

We consider three densities for Y:

Weibull $$f(y;\mu,\phi) = \frac{\phi y^{\phi-1} exp(-(y/\mu)^{\phi})}{\mu^{\phi}}$$

Log-normal $$f(y;\mu,\phi) = \frac{exp(-[log(y)-\mu]^2/(2\phi))}{y\sqrt{2\pi\phi}}$$

Log-logistic $$f(y;\mu,\phi) = \frac{\pi exp(-\pi[log(y)-\mu]/(\sqrt{3}\phi))}{y\sqrt{3}\phi(1+exp(-\pi[log(y)-\mu]/(\sqrt{3}\phi)))}$$

(See Lindsey, 1998 for more densities)

DDZ Deutsches Diabetes-Zentrum

Bivariate time-to-event model for interval-censored data - Final model

Model two location parameters μ_{D+}, μ_{D-} one for the diseased ($D+$) and one for the non-diseased ($D-$) simultaneously, and link the groups from the same study by a bivariate random effect

$$log(\mu_{D+}) = b_{D+} + \phi_{D+},\ log(\mu_{D-}) = b_{D-} + \phi_{D-}$$

$$\begin{pmatrix} \phi_{D+} \\ \phi_{D-} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{D+} & \rho\sigma_{D+}\sigma_{D-} \\ \rho\sigma_{D+}\sigma_{D-} & \sigma^2_{D-} \end{pmatrix} \right]$$

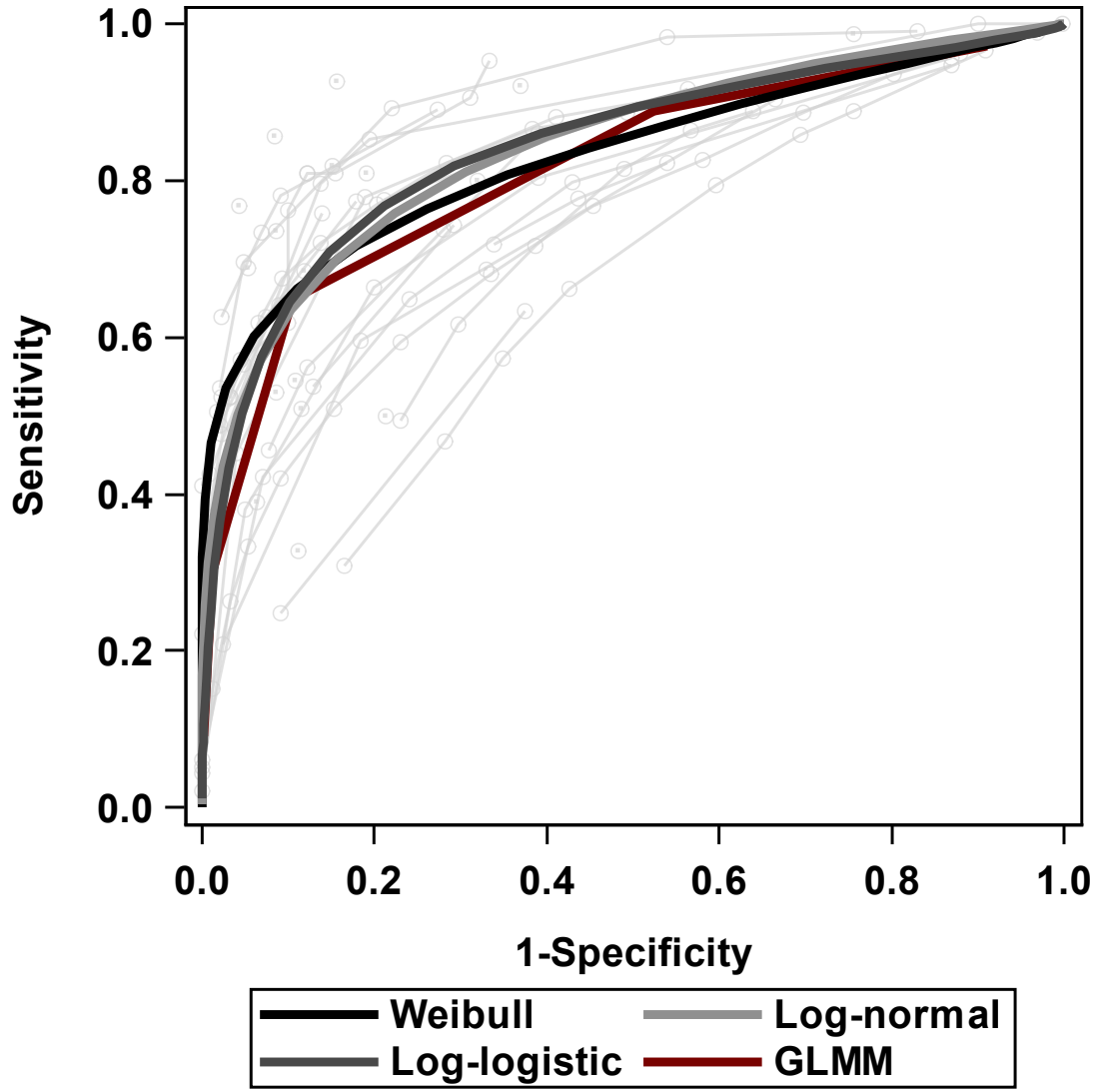
DDZ Deutsches Diabetes-Zentrum

Parameter estimation

- The following parameters have to be estimated: $b_{D+}, b_{D-}, \phi_{D+}, \phi_{D-}, \sigma^2_{D+}, \sigma^2_{D-}, \rho$
- Estimation method: Gaussian quadrature (SAS PROC NLMIXED)
- Again, we are not interested in the model parameters, but in the summary ROC curve
- That is, predict sensitivity and specificity at given thresholds by the BLUP principle (PREDICT statement, NLMIXED)

DDZ Deutsches Diabetes-Zentrum

Results for the HbA1c example



Deutsches Diabetes-Zentrum

Discussion

Summary

We propose two approaches for the meta-analysis of full ROC curves that use the information for all thresholds.
Both models avoid the problems of previous methods and come with the following advantages:

- Constitute a one-step approach
- Allow varying numbers and varying values of thresholds
- Explicitly model the value of the diagnostic test
- Allow heterogeneity of sensitivity and specificity across studies
- Account for correlation within studies

DDZ Deutsches Diabetes-Zentrum

Summary

- Ensure proper weighting of studies' sample size
- Work well in simulations
- Can be generalized to compare two (or even more) diagnostic tests to a common gold standard

Deutsches Diabetes-Zentrum

Limitations

- Distributional assumptions for the values of the diagnostic test (Idea: Relax this by using a piecewise constant function)
- Robustness
- Idea II: How to choose the minimal value for the diagnostic test?

Deutsches Diabetes-Zentrum

Thank you!



View from the IBE on Düsseldorf

DDZ Deutsches Diabetes-Zentrum

Bibliography

Bennett CM, Guo M, Dharmage SC. HbA1c as a screening tool for detection of Type 2 diabetes: a systematic review.
Diabetic Medicine. 2007;24:333-343.

Kodama S, Horikawa C, Fujihara K, Hirasawa R, Yachi Y, Yoshizawa S, Tanaka S, Sone Y, Shimano H, Iida KT, Saito K, Sone H. Use of high-normal levels of haemoglobin A_1 *C* and fasting plasma glucose for diabetes screening and for prediction: a meta-analysis.
Diabetes/ Metabolism Research and Reviews. 2013;29(8):680-692.

Rücker G, Schumacher M. Summary ROC curve based on a weighted Youden index for selecting an optimal cutpoint in meta-analysis of diagnostic accuracy.
Statistics in Medicine. 2010; 29(30):3069-3078.

Arends LR, Hamza TH, van Houwelingen JC, Heijenbrok-Kal MH, Hunink MG, Stijnen T. Bivariate random effects meta-analysis of ROC curves.
Medical Decision Making. 2008;28(5):621-38.

Chu H, Guo H. Letter to the editor.
Biostatistics. 2009;10(1):201-203.

Kester AD, Buntinx F. Meta-analysis of ROC curves.
Medical Decision Making. 2000;20(4):430-9.

Dukic V, Gatsonis C. Meta-analysis of diagnostic test accuracy assessment studies with varying number of thresholds.
Biometrics. 2003;59:936-46.

Poon WY. A latent normal distribution model for analysing ordinal responses with applications in meta-analysis.
Statistics in Medicine. 2004;23(14):2155-72.

DDZ Deutsches Diabetes-Zentrum

Bibliography


Bipat S, Zwinderman AH, Bossuyt PM, Stoker J. Multivariate random-effects approach: for meta-analysis of cancer staging studies.
Academic Radiology. 2007;14(8):974- 84.


Hamza TH, Arends LR, van Houwelingen HC, Stijnen T. Multivariate random effects meta-analysis of diagnostic tests with multiple thresholds.
BMC Medical Research Methodology. 2009;9:73.


Putter H, Fiocco M, Stijnen T. Meta-analysis of diagnostic test accuracy studies with multiple thresholds using survival methods.
Biometrical Journal. 2010;52(1):95-110.


Martinez-Camblor P. Fully non-parametric receiver operating characteristic curve estimation for random-effects meta-analysis.
Statistical Methods in Medical Research. 2014 May 28.


Riley RD, Takwoingi Y, Trikalinos T, Guha A, Biswas A, Ensor J, Morris RK, Deeks JJ. Meta-Analysis of Test Accuracy Studies with Multiple and Missing Thresholds: A Multivariate-Normal Model.
Journal of Biometrics & Biostatistics. 2014;5:196.


Zhang J, Yang Q, Zhang JA, Chen Y, Qin Q, Yang J, Yan N. The value of HbA1c for diagnosing type 2 diabetes in high-risk population.
Journal of Xian Jiaotong University (Medical Sciences). 2010;31(4):512514.


Lindsey JK. A study of interval censoring in parametric regression models.
Lifetime Data Analysis. 1998;4(4):329-54.

Deutsches Diabetes-Zentrum