Statistical methods for the meta-analysis of full ROC curves

Oliver Kuss (joint work with Stefan Hirt and Annika Hoyer)

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

May 4, 2016

DDZ
Deutsches Diabetes-Zentrum

Outline

Deutsches Diabetes-Zentrum

Introduction

Diagnostic study

Diagnostic Test T	**Gold Standard** D +	-
+	TP	FP
-	FN	TN

- TP: True Positive, FP: False Positive, FN: False Negative, TN: True Negative, $N = TP + FP + FN + TN$
- Sensitivity: $Se = P(T^+ \mid D^+) = \frac{TP}{TP+FN}$
- Specificity: $Sp = P(T^- \mid D^-) = \frac{TN}{FP+TN}$

DDZ Deutsches Diabetes-Zentrum

Diagnostic study

Diagnostic Test T	**Gold Standard** D +	-
+	392	1130
-	267	5015

- Sensitivity: $Se = P(T^+ \mid D^+) = \frac{392}{392+267} = 59\%$
- Specificity: $Sp = P(T^- \mid D^-) = \frac{5015}{1130+5015} = 82\%$

DDZ Deutsches Diabetes-Zentrum

Meta-analysis

- 'Meta-analysis is a quantitative, systematic summary of a collection of separate studies for the purpose of obtaining information that cannot be derived from any of the studies alone' (Boissel et al., 1988)

Deutsches Diabetes-Zentrum

Meta-analysis of diagnostic studies

- Meta-analysis for intervention studies is well established today
- In contrast, meta-analysis of diagnostic studies 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 of diagnostic studies

Study	Sensitivity	Specificity
1	392/659	5015/6145
2	207/252	2704/5865
...	...	...
	Se_{MA}=...	Sp_{MA}=...

generally negatively correlated

DDZ Deutsches Diabetes-Zentrum

Meta-analysis of diagnostic studies

- 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 studies - An example

Study	Threshold	Sensitivity	Specificity
1	5.0	596/659	2062/6145
1	5.5	392/659	5015/6145
1	6.0	137/659	5999/6145
2	5.3	207/252	2704/5865
2	5.4	196/252	3308/5865
...	...	...	...

DDZ Deutsches Diabetes-Zentrum

Meta-analysis of diagnostic studies - 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

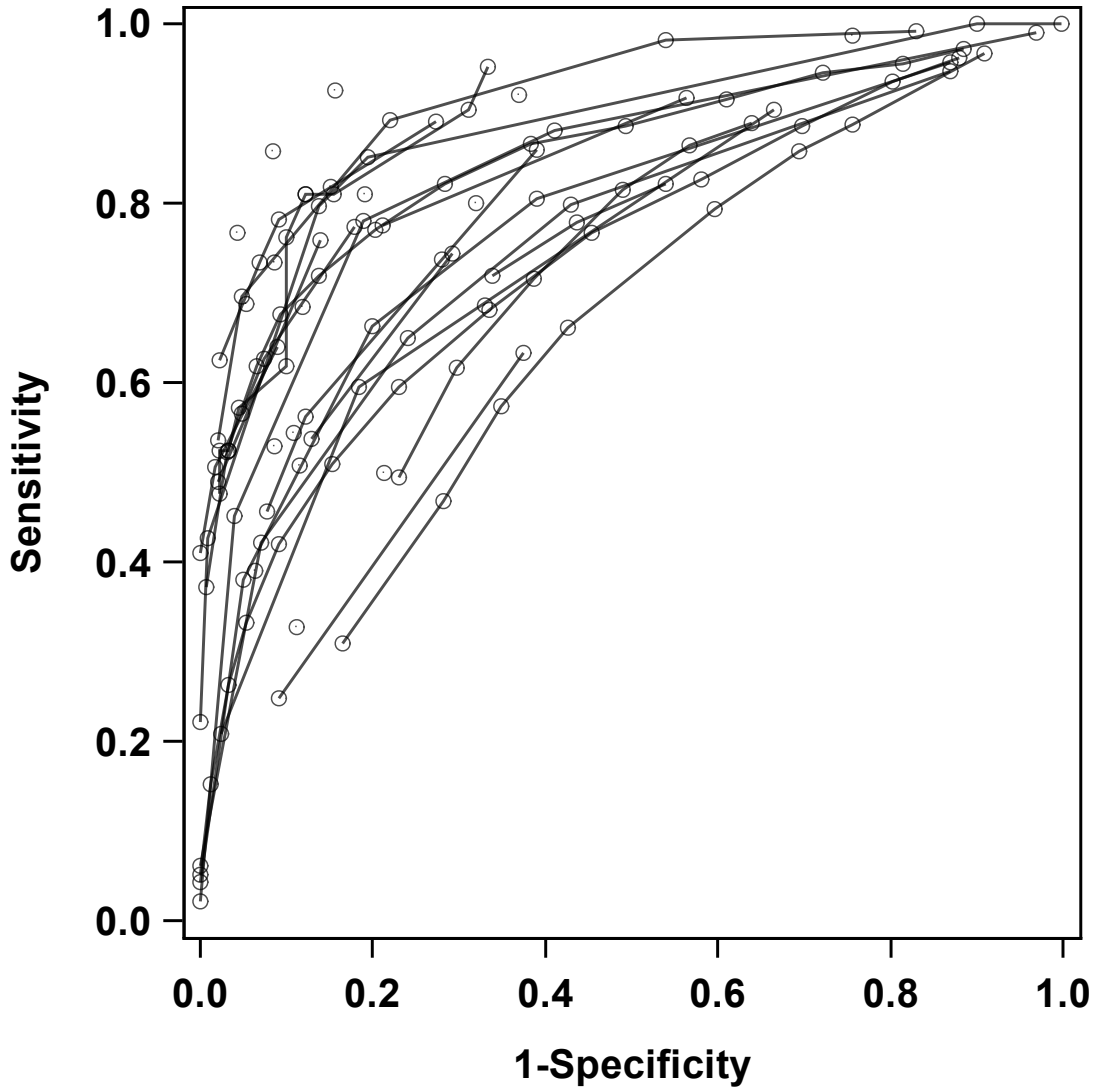
Deutsches Diabetes-Zentrum

Meta-analysis of diagnostic studies - 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 studies - An example



Deutsches Diabetes-Zentrum

Meta-analysis of diagnostic studies - 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

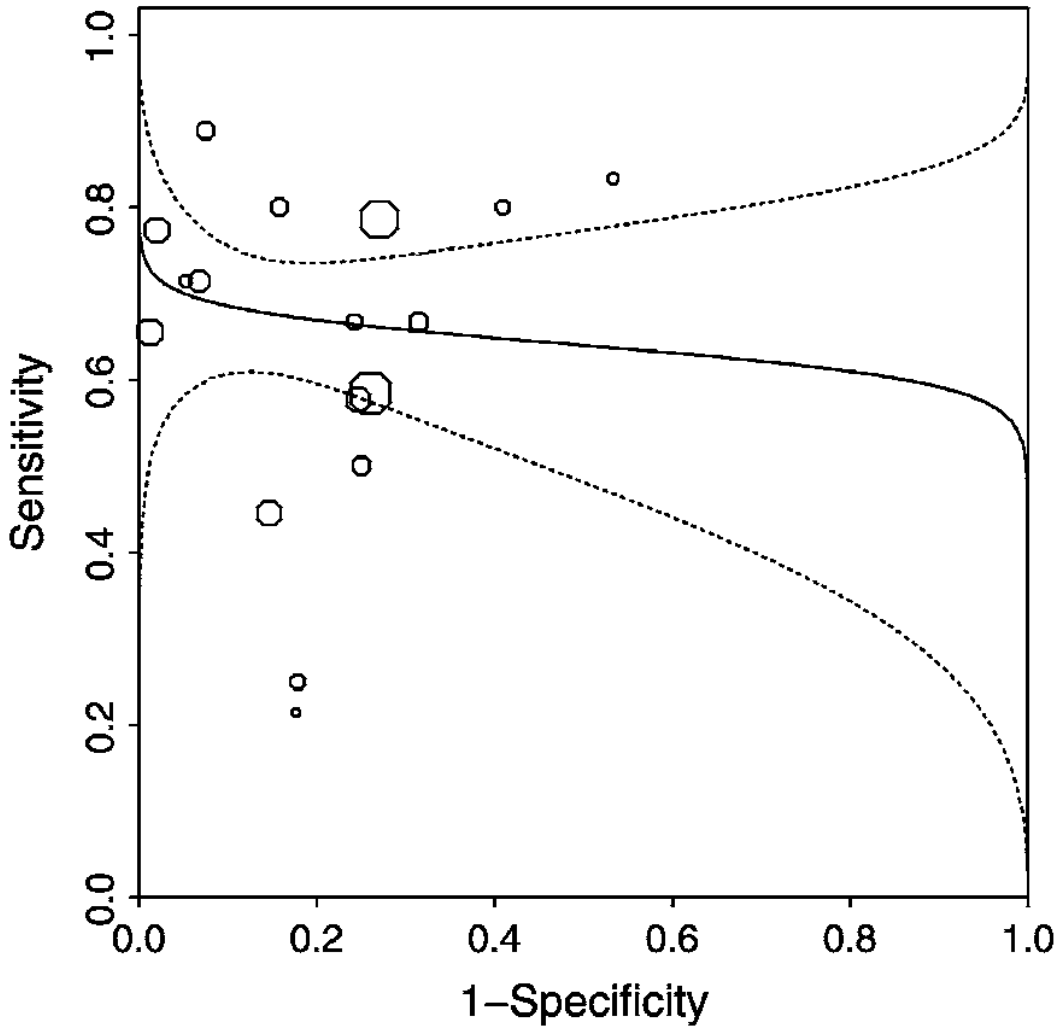
DDZ Deutsches Diabetes-Zentrum

Meta-analysis of diagnostic studies - 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]

DDZ Deutsches Diabetes-Zentrum

Meta-analysis of diagnostic studies - The paradox of the standard SROC



[Chu/Guo, 2009]

Meta-analysis of diagnostic studies - 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)

Deutsches Diabetes-Zentrum

Meta-analysis of diagnostic studies - 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 studies - 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

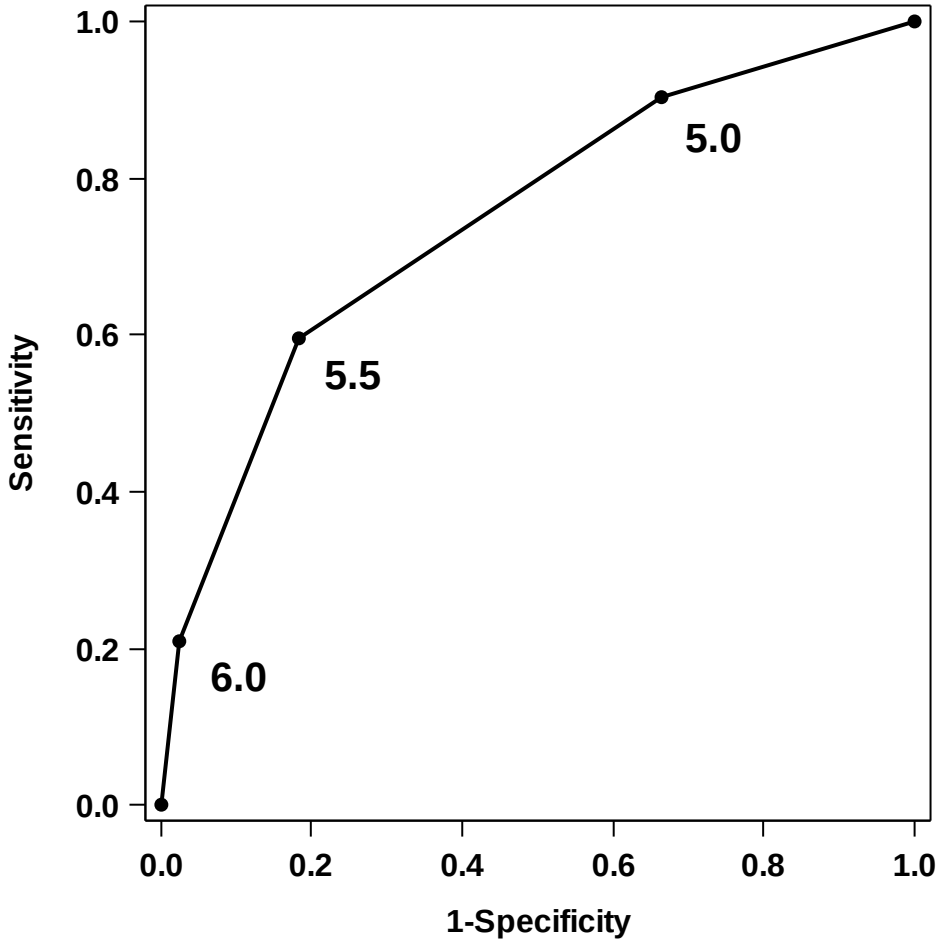
Meta-analysis of diagnostic studies - A boring solution

Extend current standard model (Bivariate logistic regression model with random effects) to a meta-regression by adding a covariate for the threshold (under review at 'Statistical Methods in Medical Research')

Deutsches Diabetes-Zentrum

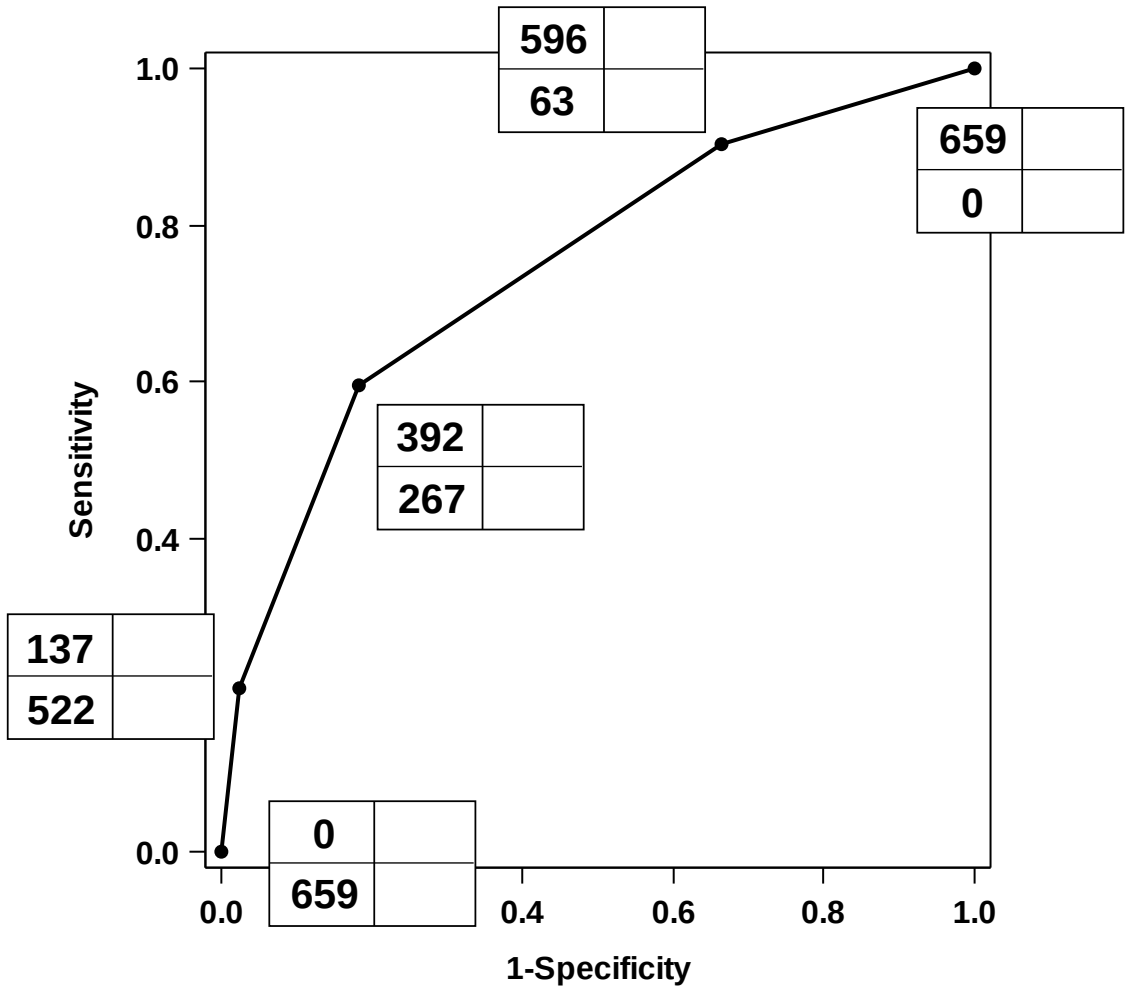
Bivariate time-to-event model for interval-censored data

ROC from Sato et al.

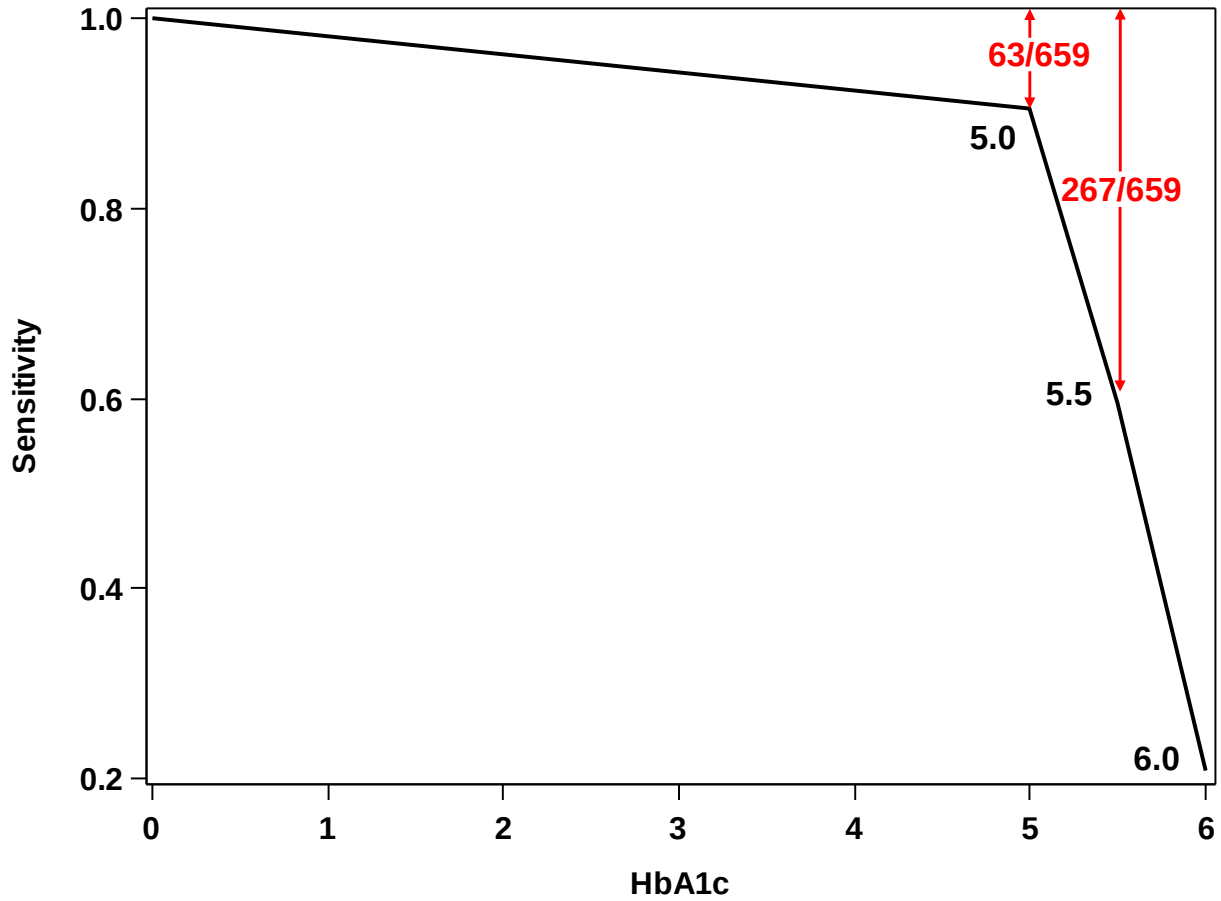


Deutsches Diabetes-Zentrum

ROC from Sato et al.



’Life-table’ estimator from Sato et al.



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

After a log transformation, the diagnostic test values Y_0 and Y_1 for non-diseased and diseased probands are allowed to vary between studies and are linked by a bivariate normal distribution with correlation ρ.

$$\log(y_0) = b_0 + \epsilon_0 + u_{0i}, \tag{1}$$

$$\log(y_1) = b_1 + \epsilon_1 + u_{1i}, \tag{2}$$

and

$$\begin{pmatrix} u_{0i} \\ u_{1i} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_0^2 & \rho\sigma_0\sigma_1 \\ \rho\sigma_0\sigma_1 & \sigma_1^2 \end{pmatrix}\right]. \tag{3}$$

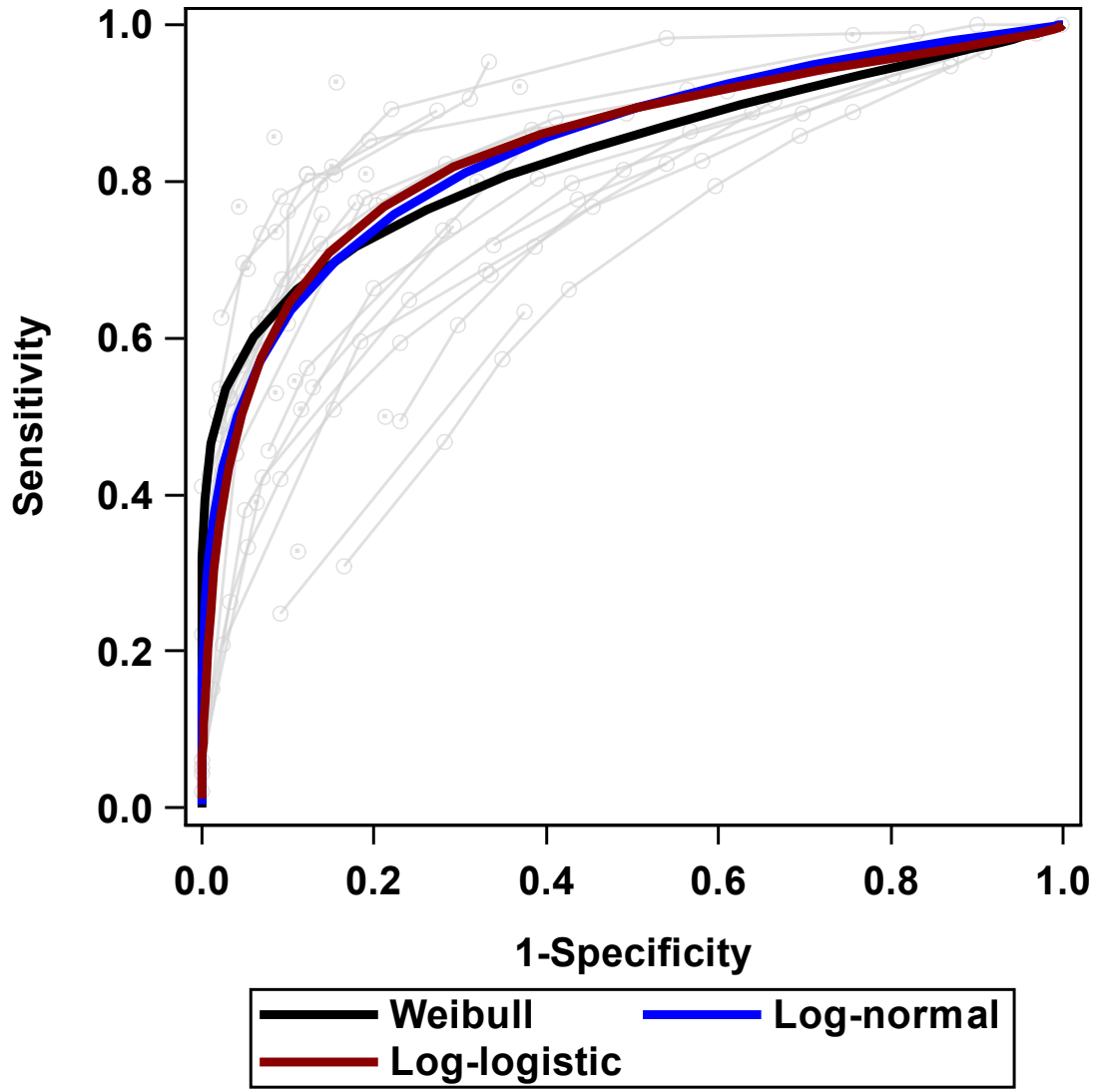
DDZ Deutsches Diabetes-Zentrum

Parameter estimation

- The following parameters have to be estimated: $b_0, b_1, \epsilon_0, \epsilon_1, \sigma_0^2, \sigma_1^2$, and ρ.
- Estimation method: Gaussian quadrature (SAS PROC NLMIXED)
- 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)

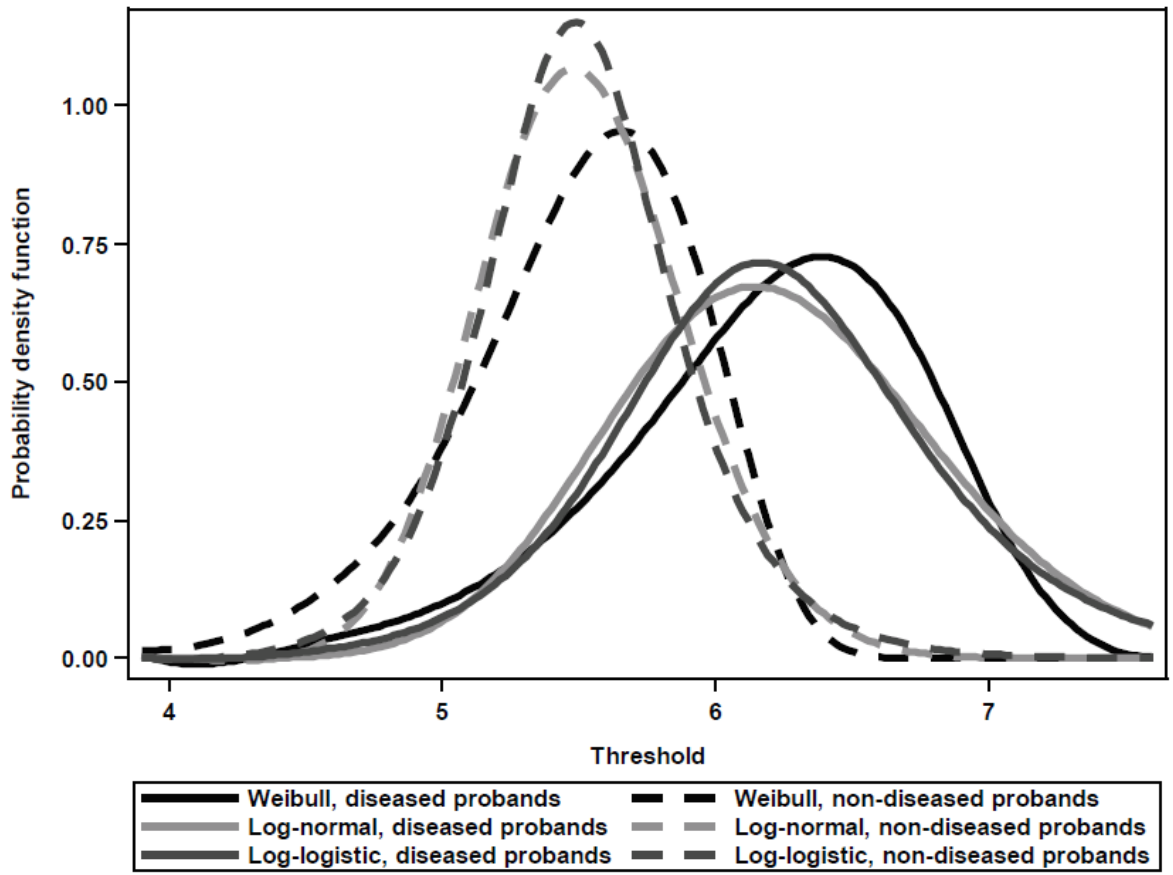
DDZ Deutsches Diabetes-Zentrum

Results for the HbA1c example



DDZ
Deutsches Diabetes-Zentrum

Results for the HbA1c example - Estimated densitites



DDZ
Deutsches Diabetes-Zentrum

Discussion

Summary

We proposed an approach for the meta-analysis of full ROC curves that use the information for all thresholds.
It avoids the problems of previous methods and comes with the following advantages:

- Constitutes a one-step approach
- Allows varying numbers and varying values of thresholds
- Explicitly models the value of the diagnostic test
- Allows heterogeneity of sensitivity and specificity across studies
- Accounts for correlation within studies

DDZ Deutsches Diabetes-Zentrum

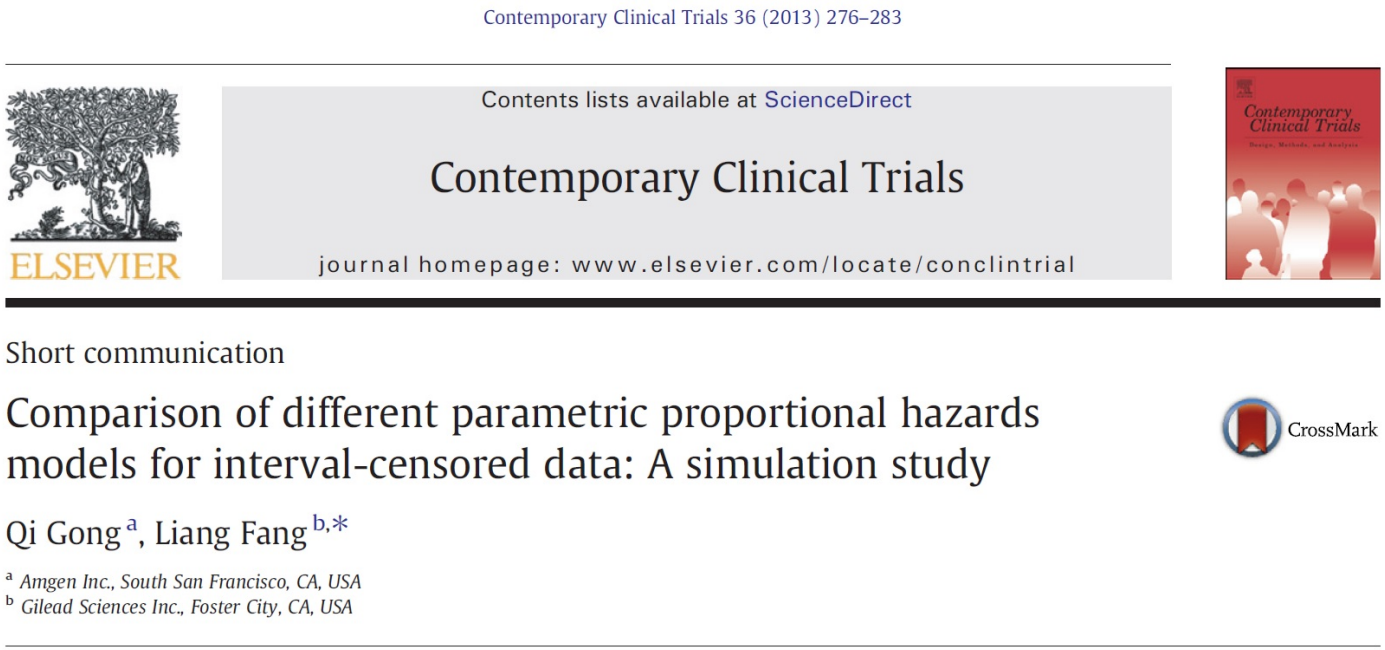
Summary

- Ensures proper weighting of studies' sample size
- Works 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)



Contemporary Clinical Trials 36 (2013) 276–283

Contents lists available at ScienceDirect

Contemporary Clinical Trials

journal homepage: www.elsevier.com/locate/conclintrial

Short communication

Comparison of different parametric proportional hazards models for interval-censored data: A simulation study

Qi Gong [a], Liang Fang [b,*]

[a] Amgen Inc., South San Francisco, CA, USA
[b] Gilead Sciences Inc., Foster City, CA, USA

- Robustness

Thank you!

View from the IBE on ”Rheinisches Braunkohlerevier”

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.


Gong Q, Fang L. Comparison of different parametric proportional hazards models for interval-censored data: a simulation study.
Contemporary Clinical Trials. 2013;36(1):276-83.

Deutsches Diabetes-Zentrum