Program dataset challenge:

05 Sep 2018

11:15 – 11:30	Introduction and presentation of the example meta-analysis (A. Zapf)
11:30 – 12:15	Meta-analysis of diagnostic test accuracy studies with multiple cutoffs: The R package diagmeta (G. Rücker)
12:15 – 12:45	Meta-analysis of diagnostic test accuracy: a flexible Bayesian model for multiple and explicit thresholds (H. Jones)
12:45 – 13:30	Lunch break
13:30 – 14:00	Bivariate time-to-event models for the meta-analysis of diagnostic tests accounting for multiple thresholds (A. Hoyer)
14:00 – 14:30	Nonparametric meta-analysis of diagnostic accuracy studies with multiple cutoffs (C. Frömke)
14:30 – 15:00	Comparison of the results from the different approaches and conclusion (A. Zapf)

Meta-analysis of the accuracy of Neutrophil Gelatinase-Associated Lipocalin (NGAL) in diagnosis of acute kidney injury

Antonia Zapf [1], Judith Heinz [2], Christian Albert [3], Michael Haase [3]

[1] Department of Medical Biometry and Epidemiology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

[2] Department of Medical Statistics, University Medical Center Göttingen, Göttingen, Germany

[3] Diaverum, Rangsdorf, Germany

The need for clinically applicable cutoff values of neutrophil gelatinase-associated lipocalin (NGAL) for the diagnosis of acute kidney injury (AKI) has been a limiting factor in clinical nephrology research and practice.

PubMed, Web of Science, Cochrane Library, Scopus databases and congress abstracts published up to May 2016 were searched for diagnostic test studies reporting the value of urine or plasma NGAL measured on clinical laboratory platforms to diagnose AKI. To reduce heterogeneity, authors of identified original studies recalculated predictive values and cutoffs of NGAL using standardized AKI definition and timing of NGAL measurement in relation to kidney injurious event. The aim was to obtain threshold values at Youden index, given 95% sensitivity and 95% specificity per individual study allowing for complete ROC curve construction.

Meta-analysis of diagnostic test accuracy:

a flexible Bayesian model for multiple and explicit thresholds

Hayley E Jones[1], Constantine A Gatsonis[2,3], Thomas A Trikalinos[3], Nicky J Welton[1], A E Ades[1]

1. Population Health Sciences, Bristol Medical School, University of Bristol, UK

2. Center for Statistical Sciences, School of Public Health, Brown University, USA

3. Center for Evidence Synthesis in Health, School of Public Health, Brown University, USA

The optimum threshold at which to operate a clinical test is usually a key question for clinical practice. Standard methods for meta-analysis of test accuracy don't facilitate answering this question, since they (i) do not provide summary estimates of test accuracy across the full range of thresholds, and (ii) can only synthesise a single pair of sensitivity and specificity from each study, despite studies often reporting data at more than one threshold.

We describe a Bayesian multinomial meta-analysis model that can take any number of pairs of sensitivity and specificity from each study and explicitly estimates the accuracy of the test across all thresholds. Our model does not require the analyst to pre-specify the distributional form of underlying test results. We assume only that some unspecified Box-Cox transformation of test results in the diseased and disease-free populations has a logistic distribution: the transformation parameter is estimated from the data. In addition to credible intervals for the pooled sensitivity and specificity across all available thresholds, we produce prediction intervals, allowing for between-study heterogeneity in all parameters.

We will demonstrate the model using a data set provided by Antonia Zapf, quantifying the sensitivity and specificity of a Neutrophil Gelatinase-Associated Lipocalin (NGAL) for the diagnosis of renal damage.

Meta-analysis of diagnostic test accuracy studies with multiple cutoffs: The R package **diagmeta**.

G Rücker (1), G Schwarzer (1), S Steinhauser (2)

(1) Institute of Medical Biometry and Statistics, Faculty of Medicine and Medical Center - University of Freiburg, Germany
(2) Institute of Medical Statistics and Computational Biology, University of Cologne, Germany

The R package **diagmeta** implements and extends our model for meta-analysis of diagnostic test accuracy (DTA) studies allowing for multiple cutoffs (Steinhauser 2016). This parametric model assumes that the values of the underlying biomarker follow two correlated distributions for individuals with/without the target condition. Data can be entered either as study label, cutoff, TP (true positive), TN (true negative), FP (false positive), FN (false negative), or as individual participant data (study label, individual's measurement, status). Users can choose between several mixed linear models and specify the type of distribution (logistic or normal), and the weighting method for studies (e.g., inverse variance weighting). For determination of an optimal cutoff, weights for sensitivity and specificity can be specified.

diagmeta provides basic information such as the number of studies and cutoffs, the empirical distribution of cutoffs, the optimal cutoff, sensitivity and specificity at this cutoff, and the area under the summary ROC curve. For given cutoffs, pairs of sensitivity and specificity with confidence intervals can be tabulated. If a prevalence is specified, predicted values are calculated. In addition, a flexible plot function is provided to produce cumulative distribution plots, density plots, Youden index curves, study-specific ROC curves, the summary ROC curve, and the summary operating point, optionally with a corresponding confidence region.

We exemplify the use of **diagmeta** by applying it to a data set provided by Antonia Zapf.

(Steinhauser 2016) Steinhauser S et al.: Modelling multiple thresholds in meta-analysis of diagnostic test accuracy studies. BMC Med Res Methodol. 2016;16(1):97.doi:10.1186/s12874-016-0196-1.

Bivariate time-to-event models for the meta-analysis of diagnostic tests accounting for multiple thresholds

Annika Hoyer[1], Oliver Kuss[1]
[1]German Diabetes Center, Leibniz Center for Diabetes Research at Heinrich Heine University Düsseldorf, Institute for Biometrics and Epidemiology

Diagnostic tests are often evaluated at several thresholds in diagnostic accuracy studies, aiming to determine optimal thresholds for practice. When doing meta-analyses, researchers frequently ignore that single studies report on full receiver operating characteristic (ROC) curves and select only one single threshold per study. Although methods have been proposed that can use the full information, these still have disadvantages. We propose an alternative parametric approach for the meta-analysis of full ROC curves that is based on bivariate time-to-event models for interval-censored data using Weibull, log-normal and log-logistic distributions [1].
The model is illustrated by the example of kidney injury, based on a data set provided by Antonia Zapf.

[1] Hoyer A, Hirt S, Kuss O. Meta-analysis of full ROC curves using bivariate time-to-event models for interval-censored data. Research Synthesis Methods 2018;9(1):62-72