UMC Utrecht

Prediction of absolute treatment effect

Challenges and opportunities

UTI-IPD symposium; 7 may 2021

Thomas Debray

About me

Affiliations

- University Medical Center Utrecht
- University of Oxford
- University College London
- Cochrane

Research of statistical methods

- Evidence synthesis
- Real World Data

@TPA_Debray

Background

Estimates of relative risk are commonly used to

- Evaluate efficacy and safety of medical interventions
- Develop medical guidelines
- Decide upon treatment for individual patients

However, estimates of relative risk

- Do not indicate how individual outcomes are affected by treatment
- Cannot directly be used to personalize healthcare

Personalized decision making

What outcomes are most likely to happen for the specific individual in the presence and in the absence of the intervention?

- Both outcomes can be predicted using multivariable models
- The mere difference between these two absolute outcome predictions provides the absolute intervention effect for that specific individual
- The absolute intervention effect may differ between individuals if the relative treatment effect is constant

Personalized decision making

Absolute outcome risk without the intervention

Absolute outcome risk with the intervention

Absolute intervention effect

Prognosis research strategy (PROGRESS) 4: Stratified medicine research

OPEN ACCESS

Annals of Internal Medicine RESEARCH AND REPORTING METHODS

The Predictive Approaches to Treatment effect Heterogeneity (PATH) Statement

David M. Kent, MD, MS; Jessica K. Paulus, ScD; David van Klaveren, PhD; Ralph D'Agostino, PhD; Steve Goodman, MD, MHS, PhD; Rodney Hayward, MD; John P.A. Ioannidis, MD, DSc; Bray Patrick-Lake, MFS; Sally Morton, PhD; Michael Pencina, PhD; Gowri Raman, MBBS, MS; Joseph S. Ross, MD, MHS; Harry P. Selker, MD, MSPH; Ravi Varadhan, PhD; Andrew Vickers, PhD; John B. Wong, MD; and Ewout W. Steyerberg, PhD

Prediction of absolute treatment effect

Develop a multivariable (e.g. regression) model directly on RCT data with inclusion of

Prognostic variables

- Subject characteristics (e.g. age, gender)
- History and physical examination results (e.g. blood pressure)
- Imaging results
- (Bio)markers (e.g. coronary plaque)

Treatment variables

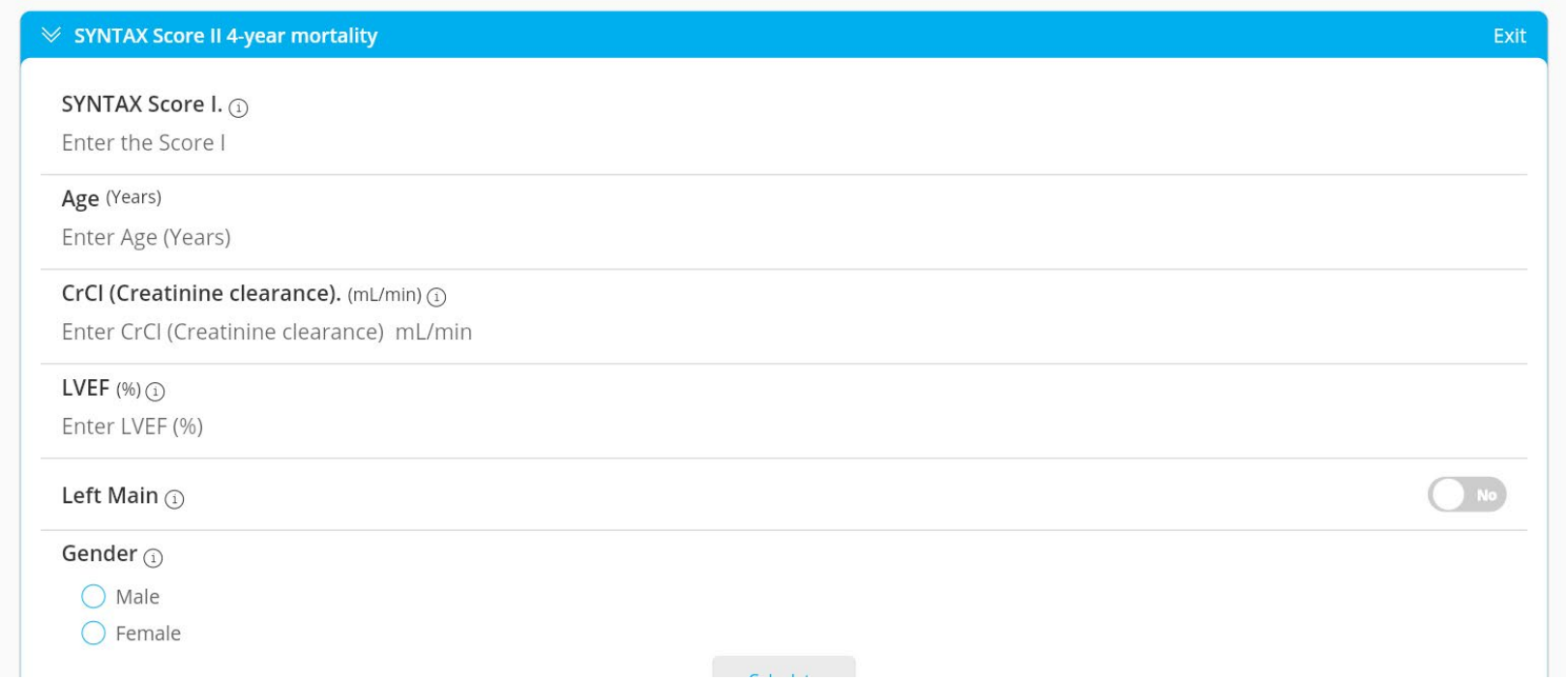
- With potential for effect modification

Example: The SYNTAX score II

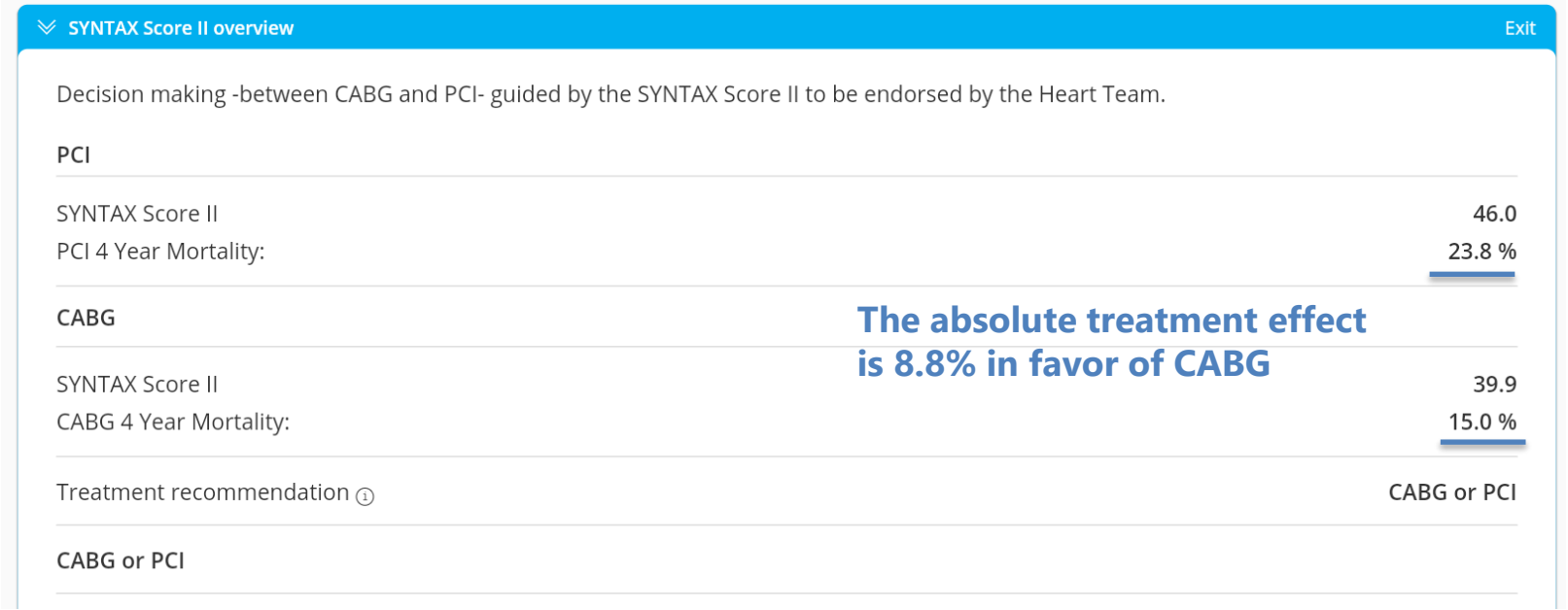
"The SYNTAX score II is a clinical tool that combines clinical variables with the anatomical SYNTAX score, providing expected 4-year mortality for both coronary artery bypass grafting (CABG) and percutaneous coronary intervention (PCI) — thus recommending either PCI only, CABG only or equipoise in treatment based on long-term mortality."

DOI: 10.21037/acs.2018.07.02 ; web calculator: https://calculator.syntaxscore2020.com/

Example: The SYNTAX score II



Example: The SYNTAX score II



The absolute treatment effect is 8.8% in favor of CABG

@TPA_Debray

Estimation of absolute treatment effect

- Accurate estimates of relative treatment effect remain key
  - An obvious choice to develop treatment effect models is to use patient-level data from randomized clinical trials (RCTs)
- Accurate estimates of prognostic effects are important
  - Prognostic effects can reliably be estimated in randomized data but also in non-randomized studies
- Accurate estimates of baseline risk are important
  - Due to strict eligibility criteria, RCT-based estimates of baseline risk may not generalize well to real-world populations

Estimation of absolute treatment effect

Concerns:

- Are (typical) RCTs large enough to develop accurate multivariable treatment effect models?

- Can treatment effect models provide accurate predictions when they are applied in “the real world” ?
  (where baseline risk, prognostic & treatment effects may differ)

A selection of modeling approaches

- Risk magnification
  - Prediction model where treatment status is included as a main effect
  - #parameters: *p (predictors) + 1 (treatment) + 1 (intercept)*
- Full modelling
  - Prediction model with one or more treatment-covariate interactions
  - #parameters: *p (predictors) + 1 (treatment) + 1 (intercept) + q (interactions)*
- Baseline risk modification
  - Prediction model that includes an interaction between treatment and the linear predictor
  - #parameters: *p (predictors) + 1 (treatment) + 1 (intercept) + 1 (interaction)*

Findings from simulation studies

A tutorial on individualized treatment effect prediction from randomized trials with a binary endpoint

J Hoogland[1], J IntHout[2], M Belias[2], MM Rovers[2], RD Riley[3], FE Harrell Jr[4], KGM Moons[1,5], TPA Debray[1,5], and JB Reitsma[1,5]

ELSEVIER

Journal of Clinical Epidemiology ■ (2015) ■

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

Estimates of absolute treatment benefit for individual patients required careful modeling of statistical interactions

David van Klaveren[a,*], Yvonne Vergouwe[a], Vasim Farooq[b,c], Patrick W. Serruys[b], Ewout W. Steyerberg[a]

Findings from simulation studies

- Risk magnification generally works well
- Shrinkage and selection critical even in large RCTs
- Inclusion of interaction terms should be driven by domain knowledge
- Modelling of interaction terms only beneficial in large trials

Are (typical) RCTs large enough to develop accurate multivariable treatment effect models? Yes, but only to some extent

Personalizing treatment in major depression

- Data available from a large multi-center trial *(N = 1544)*
- Comparison of regression-based and machine learning methods
- Internal-external validation to evaluate accuracy
- Treatment effect models based on Support Vector Machines improved treatment recommendations for a minority of participants
  (as compared to regression-based methods & group average approach)

Research paper

Can personalized treatment prediction improve the outcomes, compared with the group average approach, in a randomized trial? Developing and validating a multivariable prediction model in a pragmatic megatrial of acute treatment for major depression

Toshi A Furukawa[a,*], Thomas P A Debray[b,*], Tatsuo Akechi[c,*], Mitsuhiko Yamada[d,*], Tadashi Kato[e,*], Michael Seo[f,*], Orestis Efthimiou[g,*]

Generalizability

Treatment effect models developed in RCT may have limited validity and/or applicability

- Danger of overfitting
- Baseline risk often differs between RCTs and routine care settings
- Lack of long-term outcomes
- Potential for efficacy-effectiveness gap

Meta-analysis & integration of real-world data appears desirable

Meta-analysis of IPD

- Increase power to include treatment-covariate interactions
- Facilitate inclusion of nonlinear relationships

TUTORIAL IN BIOSTATISTICS | Open Access

Individual participant data meta-analysis to examine interactions between treatment effect and participant-level covariates: Statistical recommendations for conduct and planning

Richard D. Riley, Thomas P.A. Debray, David Fisher, Miriam Hattle, Nadine Marlin, Jeroen Hoogland, Francois Gueyffier, Jan A. Staessen, Jiguang Wang, Karel G.M. Moons, Johannes B. Reitsma, Joie Ensor

Potential advantages of RWD

- Increase power to estimate model parameters
- Improve estimation of prognostic effects & baseline risk
- Adjust for different treatment modalities (~estimands)
- Evaluate accuracy of absolute risk predictions

Future directions

- Critical appraisal of RWD
- Methodology
  - Causal inference in RWD
  - Cross-design synthesis
  - Meta-analysis of IPD
  - Missing data
- Guidance & software

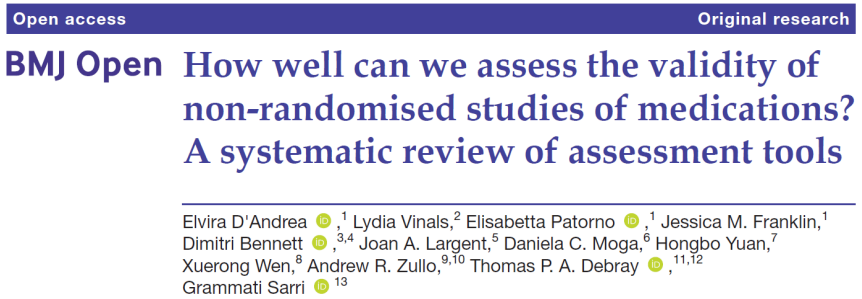
General medicine

OPEN ACCESS

Evidence synthesis

Framework for the synthesis of non-randomised studies and randomised controlled trials: a guidance on conducting a systematic review and meta-analysis for healthcare decision making

Grammati Sarri [1], Elisabetta Patorno [2], Hongbo Yuan,[3] Jianfei (Jeff) Guo,[4] Dimitri Bennett [5], Xuerong Wen,[6] Andrew R Zullo [7], Joan Largent,[8] Mary Panaccio,[9] Mugdha Gokhale,[10] Daniela Claudia Moga,[11] M Sanni Ali,[12,13,14] Thomas P A Debray [15,16]



Open access Original research

BMJ Open

How well can we assess the validity of non-randomised studies of medications? A systematic review of assessment tools

Elvira D'Andrea [1], Lydia Vinals,[2] Elisabetta Patorno [1], Jessica M. Franklin,[1] Dimitri Bennett [3,4] Joan A. Largent,[5] Daniela C. Moga,[6] Hongbo Yuan,[7] Xuerong Wen,[8] Andrew R. Zullo,[9,10] Thomas P. A. Debray [11,12] Grammati Sarri [13]