Statistical analysis plan

Matthias Klimek, Sebastian Ludwig, Andreas Ziegler

July 8, 2025

Version 2.0

Summary

This individual participant data meta-analysis (IPDMA) of randomized controlled trials (RCTs) compares transcatheter aortic valve implantation (TAVI) and surgical aortic valve replacement (SAVR) in patients with severe symptomatic aortic stenosis at low to intermediate risk. Data from four investigator-initiated trials (DEDICATE, UK TAVI, NOTION, NOTION-2) will be analyzed. The IPD from two industry-sponsored trials are currently not available. An aggregate data meta-analysis including these trials will be added. The primary endpoint is stroke-free survival at 1 year. The primary analysis will be based on the one-step IPDMA using a random effect model, stratified by STS-PROM score, in the intention-to-treat population. Correlated frailties will be used to account for trial heterogeneity. Additional subgroup and interaction analyses for gender, chronic kidney disease, and pacemaker aim to inform personalized treatment decisions between TAVI and SAVR in low and intermediate risk patients.

1. Title

Overall and sex-specific individual participant data meta-analysis to compare transcatheter with surgical aortic valve replacement in severe symptomatic aortic stenosis in patients with low to intermediate risk

2. Aims

- To explore clinical characteristics and outcomes of patients with severe symptomatic aortic stenosis assigned to transcatheter aortic valve implantation (TAVI) or surgical aortic valve replacement (SAVR).
- To estimate sex-specific differences on selected time-to-event endpoints.
- To estimate the effect in patients with chronic kidney disease (CKD).
- To estimate the effect in patients receiving/not receiving a pacemaker.

3. Search strategy

We systematically searched the MEDLINE databases until May 19, 2025, for randomized controlled trials (RCTs) that included clinical endpoints comparing TAVI and SAVR in patients with severe, symptomatic aortic stenosis at low or intermediate surgical risk.

Search terms were

#1: “aortic stenosis” AND

#2: "transcatheter aortic valve implantation" OR "transcatheter aortic valve replacement" AND

#3: "surgical aortic valve replacement" AND

#4: "randomized"

The search revealed 8 eligible RCTs. DEDICATE-DZHK6 (ClinicalTrials.gov Identifier: NCT03112980; sponsor: University Medical Center Hamburg-Eppendorf, Hamburg, Germany), UK TAVI (ISRCTN Identifier: ISRCTN57819173; University of Leicester, Leicester, United Kingdom), NOTION (ClinicalTrials.gov Identifier: NCT01057173; Rigshospitalet, Copenhagen, Denmark) and NOTION-2 (ClinicalTrials.gov Identifier: NCT02825134; Rigshospitalet, Copenhagen, Denmark) are investigator-initiated trials (IIT). The sponsors of the other four trials are medical device companies: PARTNER 3 (ClinicalTrials.gov Identifier: NCT03003299; Edwards Lifesciences), Evolut Low Risk (ClinicalTrials.gov Identifier: NCT02701283; Medtronic Cardiovascular), PARTNER 2A (NCT01314313; Edwards Lifesciences), and SURTAVI (NCT01586910; Medtronic Cardiovascular).

4. Data used for analyses

The respective principal investigators of all listed trials were asked to provide individual participant data (IPD). IPD of the four IITs was obtained and analyzed:

- DEDICATE-DZHK6 Trial
- UK TAVI Trial
- NOTION Trial
- NOTION-2 Trial

Until today, we did not obtain IPD for the trials:

- Evolut Low Risk
- PARTNER 3
- SURTAVI
- PARTNER 2A

5. Inclusion criteria

- Patients enrolled in contemporary randomized controlled trials comparing TAVI versus SAVR for severe, symptomatic aortic stenosis.
- Only RCTs including patients at low or intermediate surgical risk will be incuded.

6. Analysis populations

The primary analysis of the IPD data will be based on the intention-to-treat (ITT) populations. Analyses will be followed by the as-treated (AT) population to evaluate the consistency of results.

7. Variables used for analyses

7.1. Outcomes: time-to-event

- Primary endpoint: All-cause death (ACD; death from any cause) or stroke (disabling or non-disabling stroke) at 1 year after randomization.
- Secondary endpoints were stroke, disabling stroke, all-cause death, and rehospitalization, all measured up to 1 year after randomization. The complete list of secondary endpoints is provided in the supplementary material.

7.2. Adjustment and subgroup variables

- STS-PROM score (continuous) and STS-PROM classes: <2%, ≥2%
- Age (continuous) and age classes: <70 years, 70-75 years, ≥75 years; Age (<75 years, ≥75 years)
- Gender (e.g., sex-specific models)
- BMI (<25 kg/m^2, 25-30 kg/m^2, ≥30 kg/m^2)
- NYHA stadium (≤2, >2)
- Coronary artery disease (yes/no)
- Previous myocardial infarction (yes/no)
- Previous stroke (yes/no)
- Cerebrovascular disease (yes/no)
- Peripheral vascular disease (yes/no)
- Chronic obstructive pulmonary disease (COPD) (yes/no)
- Left-ventricular ejection fraction (LVEF) (<55%, ≥55%)
- Aortic valve (AV) mean gradient (<40mmHg, ≥40mmHg)
- Moderate-to-severe chronic kidney disease (CKD) as determined by the estimated glomerular filtration rate (eGFR) (<60ml/min/1.73m^2/≥60ml/min/1.73m^2)
- New permanent implanted pacemaker (yes/no)

8. Statistical analyses

Data will be analyzed using both one-step and two-step IPD meta-analysis (IPDMA). Both random effect (RE) and common effect (CE) models will be estimated to investigate the consistency of results. All endpoints will be investigated at 1 year. Thus, administrative censoring will be used in case of a longer follow-up.

8.1. Inclusion criteria

Key endpoints are time-to-event variables. These data do not need to be imputed.

One-step MA using IPD

- One-step IPD meta-analysis (Burke et al., 2017) will be estimated using the Katsahian RE model (Katsahian and Boudreau, 2011). This model is an expansion of the Fine and Gray (1999) model, which allows for a frailty term.
- The following RE models will be estimated:
  - crude model (no adjustments),
  - main model: stratified by STS-PROM,
  - stratified by age groups,
  - stratified by STS-PROM and age groups,
- Competing risks: correlated frailties will be estimated from the Γ distribution using EM algorithm (Rueten-Budde et al., 2019).
- Fractional polynomials of degree 2 (Benner, 2005) will be used to identify the best transformations for the STS-PROM score and age.
- Estimation: overall subdistribution hazard ratios (sHR) will be estimated with 95% confidence intervals (CI) for the endpoints (Meddis et al., 2020).
- Additional analysis: cause-specific hazard ratios (csHR) will be estimated for competing risks in addition to sHRs (Putter et al., 2007, Schmitt et al., 2023).
- For each competing event to all-cause death, pairwise variances and frailty correlation coefficients will be estimated with 95% CI. Frailty plots will display differences between trials.
- Overall cumulated incidence function (CIFs) will be estimated by the pooled and interpolated CIF functions from the subgroup combinations of the STS-PROM and age classes.
- I^2 statistics will be estimated to investigate heterogeneity between the trials (Langan et al., 2019, Veroniki et al., 2016).
- All steps will be applied to gender-stratified analysis, i.e., separately for men and women.
- Analyses will also be done with an interaction term between procedure (TAVI/SAVR) and gender.

8.2. Two-step MA

Two-step meta-analyses will be performed based on IPD as well as on IPD added by aggregated published data from studies for which IPD are not available.

All analyses will be done for the overall population and with stratification by gender. The disadvantage of the two-step MA based on both IPD, and aggregated data is that interaction terms will not be available in those studies, for which IPD are not available.

8.2.1. Two-step MA based on IPD

- MA performed as described in Section 8.2 on each trial with IPD
- Correlated frailties will be estimated (as described in Section 8.2) to adjust for possible trial center effects. The statistical models for the two-step MA based on IPD will differ from the models for the one-step MA.
- RE models will be estimated to pool the sHRs and the csHRs by restricted maximum-likelihood estimator (REML) (Langan et al., 2019, Veroniki et al., 2016). The sHRs and csHRs will also be estimated by common effect (CE) models.

8.2.2. Two-step MA based on IPD and aggregated data

- Joint analyses will be performed for all RCTs (four trials with IPD and two trials with aggregated data)
- The same models will be used as those described above in Section 8.2
- If no HR estimates are available at 1 year after randomization, we will use the reported event rates per treatment group to form a relative risk (RR). The reported sample size at baseline for respective endpoint will be used to estimate the standard error.

8.3. Baseline characteristics

Analysis will be performed descriptively. P-values will be provided based on two-sided tests to illustrate differences between treatment groups. Fisher's exact test will be used to compare dichotomous variables. For continuous variables, the t-test for independent samples will be used for normally distributed data and the Mann-Whitney test for non-normally distributed data.

Summary statistics are presented:

- Overall
- Per treatment group
- Per trial
- Per trial and treatment group
- Per sex
- Per trial and sex

8.4. Procedural and periprocedural characteristics

Analysis will be performed descriptively. P-values will be provided based on two-sided tests to illustrate differences between treatment groups. Fisher's exact test will be used to compare dichotomous variables. For continuous variables, the t-test for independent samples will be used for normally distributed data and the Mann-Whitney test for non-normally distributed data.

Summary statistics are presented:

- Overall

- ☐ Per treatment group
- ☐ Per trial
- ☐ Per trial and treatment group
- ☐ Per sex
- ☐ Per trial and sex

References

Benner, A. 2005. Mfp: Multivariable fractional polynomials. *R News,* 5**,** 20–23.

Burke, D. L., Ensor, J. & Riley, R. D. 2017. Meta-analysis using individual participant data: one-stage and two-stage approaches, and why they may differ. *Stat Med,* 36**,** 855–875.

Fine, J. P. & Gray, R. J. 1999. A proportional hazards model for the subdistribution of a competing risk. *J Am Stat Assoc,* 94**,** 496–509.

Katsahian, S. & Boudreau, C. 2011. Estimating and testing for center effects in competing risks. *Stat Med,* 30**,** 1608–1617.

Langan, D., Higgins, J. P. T., Jackson, D., Bowden, J., Veroniki, A. A., Kontopantelis, E., Viechtbauer, W. & Simmonds, M. 2019. A comparison of heterogeneity variance estimators in simulated random-effects meta-analyses. *Res Synth Methods,* 10**,** 83–98.

Meddis, A., Latouche, A., Zhou, B., Michiels, S. & Fine, J. 2020. Meta-analysis of clinical trials with competing time-to-event endpoints. *Biom J,* 62**,** 712–723.

Putter, H., Fiocco, M. & Geskus, R. B. 2007. Tutorial in biostatistics: competing risks and multi-state models. *Stat Med,* 26**,** 2389–2430.

Rueten-Budde, A. J., Putter, H. & Fiocco, M. 2019. Investigating hospital heterogeneity with a competing risks frailty model. *Stat Med,* 38**,** 269–288.

Schmitt, S., Buchholz, A. & Ozga, A. K. 2023. Systematic comparison of approaches to analyze clustered competing risks data. *BMC Med Res Methodol,* 23**,** 86.

Veroniki, A. A., Jackson, D., Viechtbauer, W., Bender, R., Bowden, J., Knapp, G., Kuss, O., Higgins, J. P., Langan, D. & Salanti, G. 2016. Methods to estimate the between-study variance and its uncertainty in meta-analysis. *Res Synth Methods,* 7**,** 55–79.