On features influencing surrogacy validations

Mario Ouwens, senior statistical science director AstraZeneca, PSI Webinar: Surrogacy

29 May 2024

Disclaimer

- The views and opinions expressed in the following PowerPoint slides are those of the individual presenter and should not be attributed to AstraZeneca.

Shift in presentation

- We have run simulations
- Based on simulations, conclusions were drawn
- Most conclusions are relatively straightforward, and could have been formulated without simulation
- Shift: Because of limited presentation time, the focus of this presentation will be on the conclusions from the simulation

What are criteria that need to be fulfilled for high quality surrogacy validation?

Adequate data:

1. Adequate estimand/adequate PICOS statement used for trial inclusion

Sufficient power:

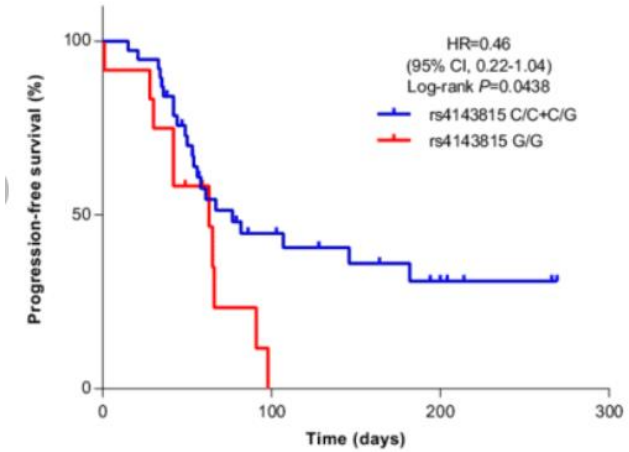
2. Sufficient events/follow-up time per treatment arm per trial
3. Sufficient number of studies included
4. Sufficient variation in treatment effect on surrogate

Correctly modelled relationship:

5. Linear relationship

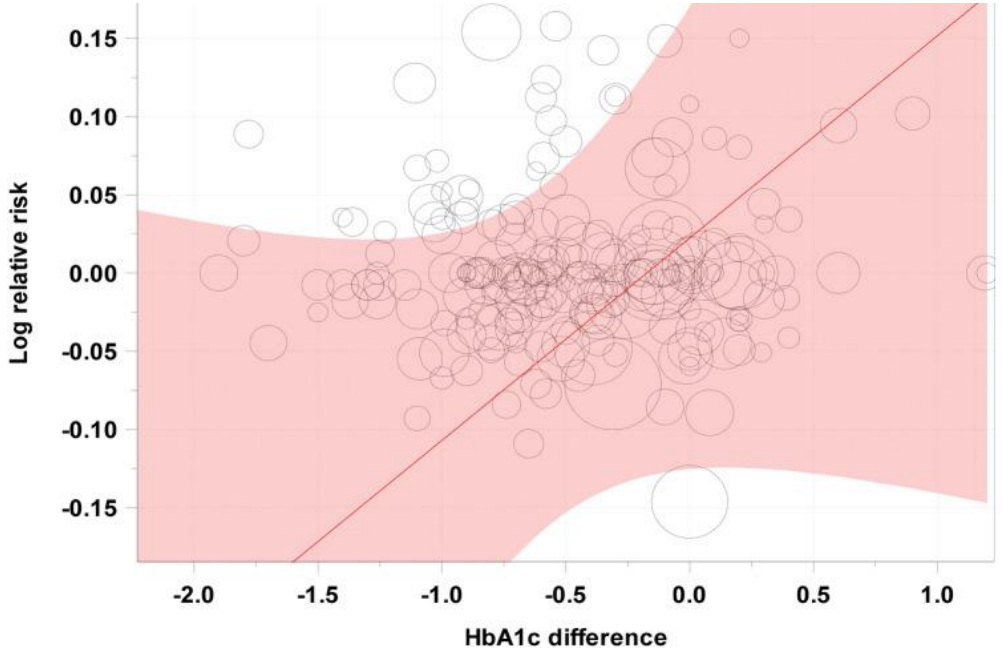
1. Adequate estimand/adequate PICOS statement (Population, Intervention, comparator, outcome study design) used for trial inclusion

- Biological/clinical plausibility
  - Population (Target population; Conditional versus marginal surrogacy)
  - Treatment mechanisms: Surrogacy for statins does not necessarily imply surrogacy for PCSK9s
  - Variable: Surrogacy for “transition to later CKD state” does not necessarily mean surrogacy for mortality
  - Summary statistic:
    Hazard Ratio may not be an adequate summary statistic
  - Approach to intercurrent events: Surrogacy in the past may have led to change of treatment after surrogacy outcome, so that surrogacy no longer hold



2. Sufficient events per treatment arm per trial - Diabetes

- Example publication: Baechle et al 2021:
  "Is HbA1c a valid surrogate for mortality in type 2 diabetes? Evidence from a meta-analysis of randomized trials"
  - 122,245 observations
  - 361 deaths (0.30% death)
  - 205 single trials
  - On average less than 1 death per treatment arm/trial

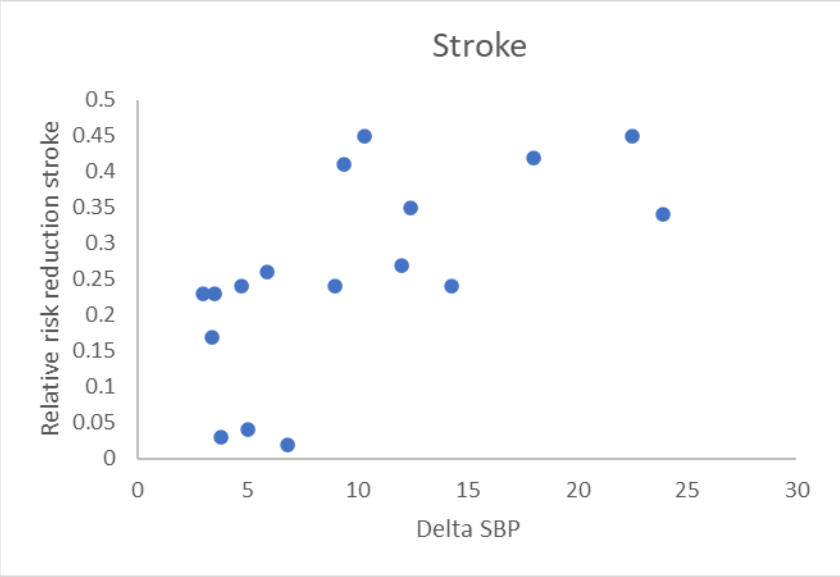


- Surrogacy evaluated using *observed* effect on mortality in trials. Is that making sense in this case?
- What requirements in general?

2. Sufficient percentage of patients with events per treatment arm per trial – Hypertension

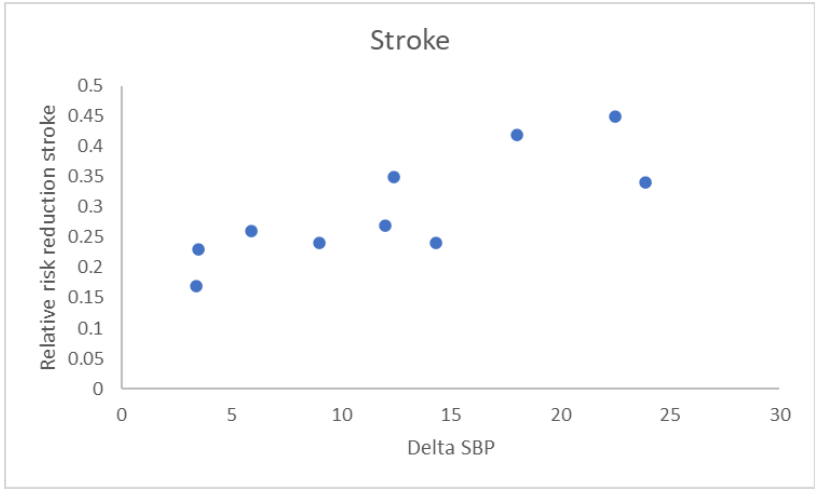
- Lassere et al: “Is blood pressure reduction a valid surrogate endpoint for stroke prevention? an analysis incorporating a systematic review of randomised controlled trials, a by-trial weighted errors-in-variables regression, the surrogate threshold effect (STE) and the biomarker-surrogacy (BioSurrogate) evaluation schema (BSES)

All data (Correlation 0.64; R squared 0.41)



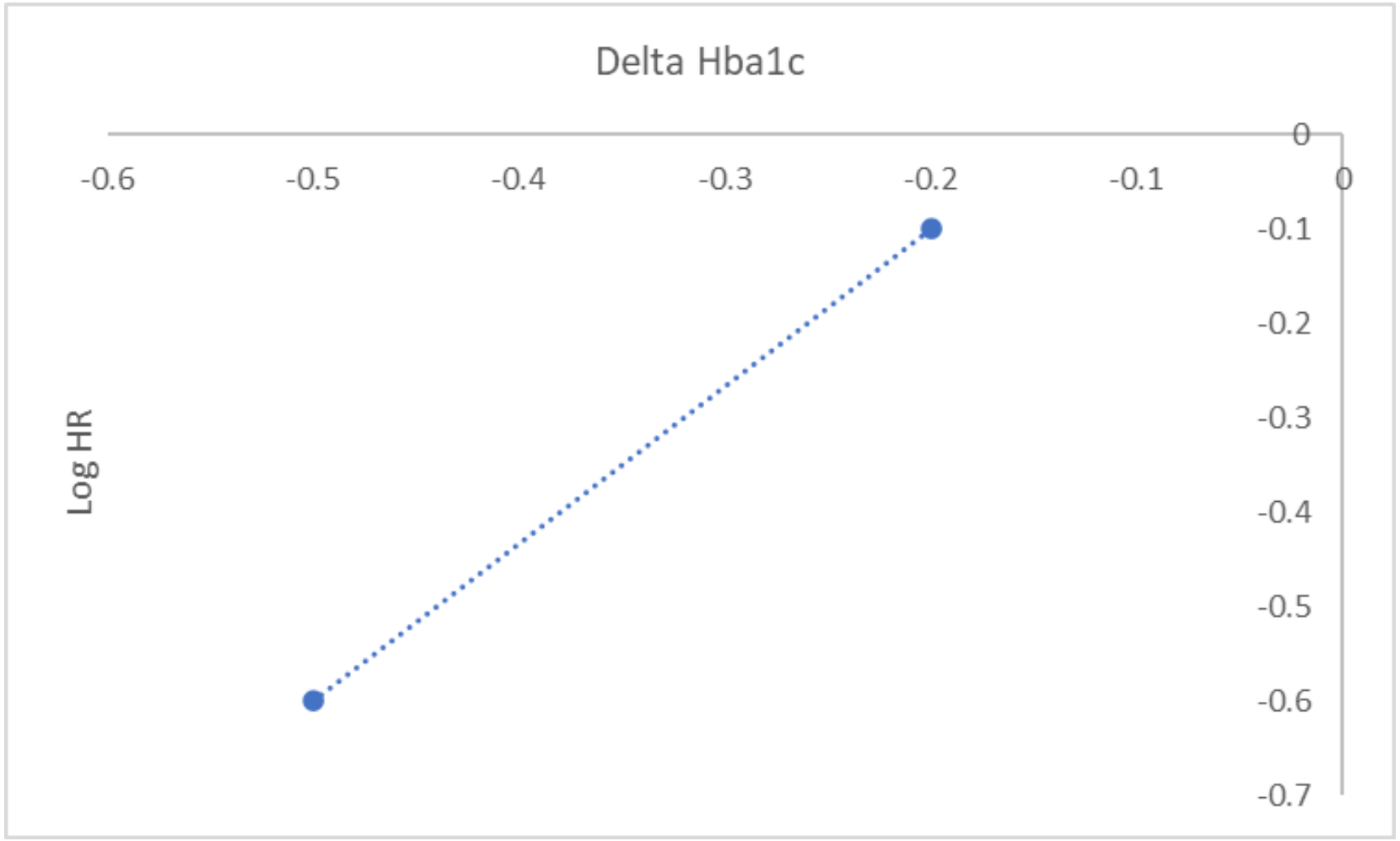
Solely data with >3% events on trial (Correlation 0.81; R squared 0.66)

(for illustrational purposes)



Note: % events influenced by trial duration as well as patient characteristics

3. Sufficient number of studies included



IQWiG criterion for oncology considers lower bound correlation

Uncertainty in the correlation is a function of the number of studies

4. Sufficient variation in treatment effect on surrogate

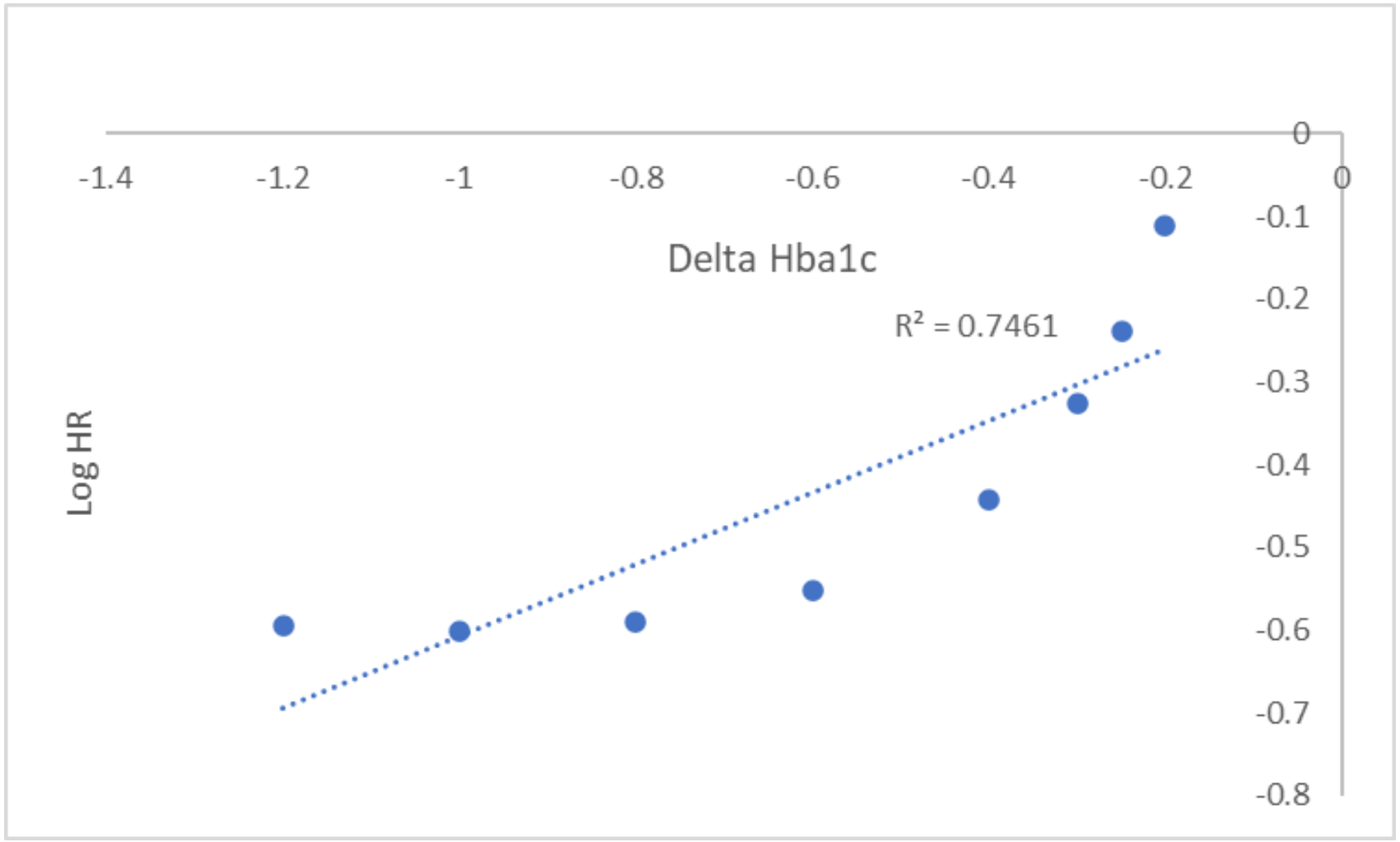
$$Y = b_0 + b_1 X + \varepsilon$$

$$b_1 = r \frac{s_y}{s_x}$$

$$r = b_1 \frac{s_x}{s_y}$$

- For observed values the s_y is a function of
  - true underlying survival probability and
  - random occurrence of death given underlying survival probability
- For true surrogacy:
  - Small variation in surrogate treatment effect implies
  - Small variation in true underlying HR implies
  - random chance dominates
- Studies thus need to vary sufficiently in treatment effect on surrogate treatment effect

5. Linear relationship



In general
linear trend assumed

Is linear trend the
correct summary?

What are criteria that need to be fulfilled for high quality surrogacy validation?

1. Adequate estimand/adequate PICOS statement used for trial inclusion
2. Sufficient events per treatment arm per trial
3. Sufficient number of studies included
4. Sufficient variation in treatment effect on surrogate
5. Linear relationship

Discussion and conclusion

- It is questionable whether we can require analytical surrogacy validations in low-mortality risk diseases like diabetes or that biological/clinical plausibility is all we should ask for

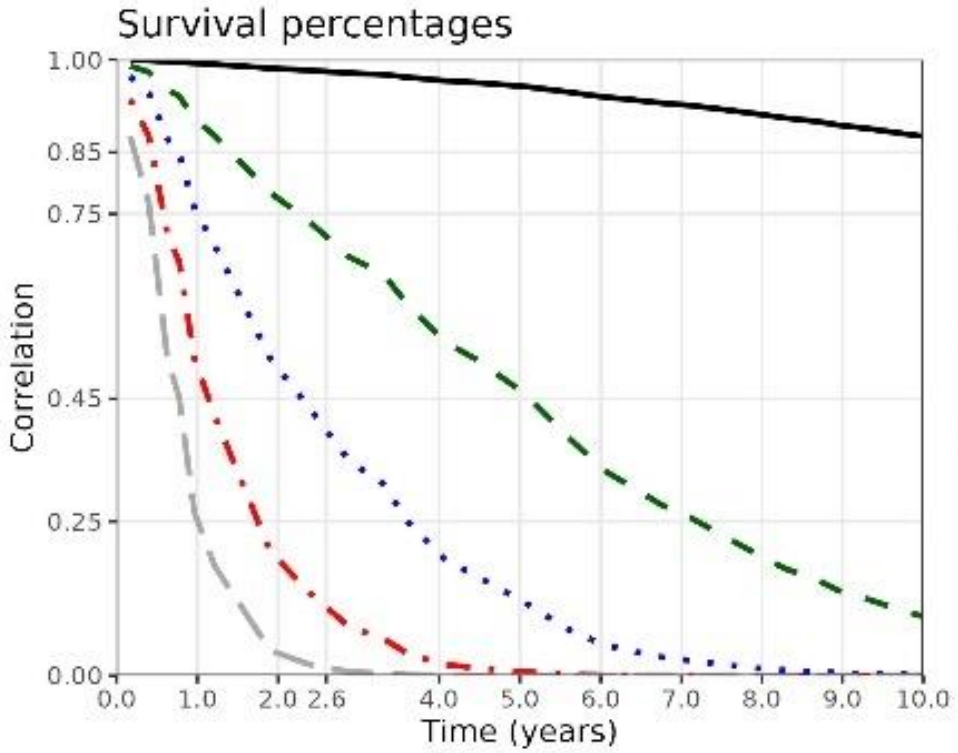
- Five criteria are provided that are important to assess when performing surrogacy analyses, but these are often not fulfilled. We need more data than is usually available, in a more representative population restricting the number of trials included, at a time where often novel subgroups/biomarkers are used even more restricting the number of trials that can be included

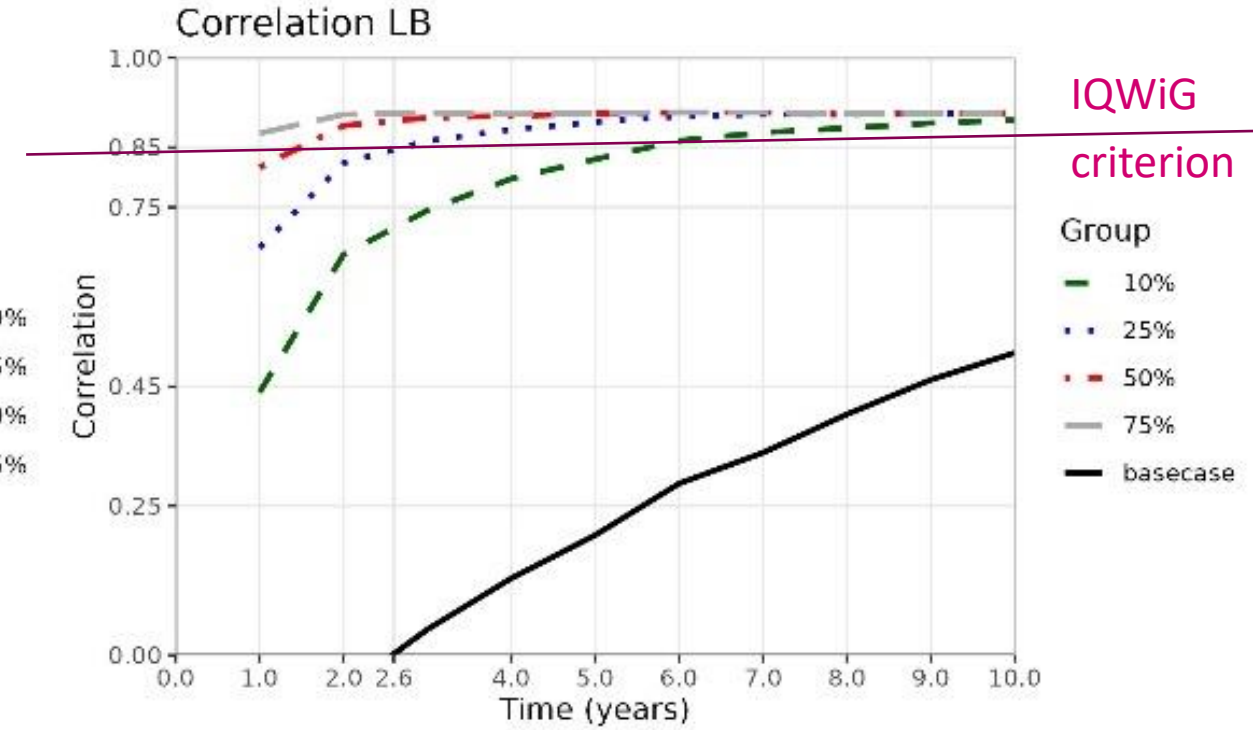
- Evaluating surrogacy while these criteria are not fulfilled may, however, lead to misleading results and conclusions and ultimately lead to not being able to use certain outcomes for licensing that could have caused (faster) access to treatments

- More guidance should be developed when we should stay at biological/clinical plausibility and just accept that analytical surrogacy validations would be of (very) low quality

Background slides

Simulation to assess duration to achieving IQWiG rule of LB 0.85





Assumptions: Perfect relationship, mean treatment effect HbA1c -0.43, standard deviation 0.17, HR 0.80 with model ln(HR) = beta*delta HbA1c, 19 trials, 7226 patients per trial, estimation via Cox model. Using weighted correlation, weighted for variance in estimated HR (assuming delta HbA1c known).