Joint modeling and Bayesian bivariate meta-analysis for surrogate endpoint validation in interstitial lung disease

h_da HOCHSCHULE DARMSTADT UNIVERSITY OF APPLIED SCIENCES
fbi FACHBEREICH INFORMATIK

h_da HOCHSCHULE DARMSTADT UNIVERSITY OF APPLIED SCIENCES
fbmn FACHBEREICH MATHEMATIK UND NATURWISSENSCHAFTEN

Nikolai Spuck
Darmstadt University of Applied Sciences

Boehringer Ingelheim

Objective

Clinical trials that investigate treatments for interstitial lung disease regularly apply forced vital capacity (FVC), a measure for lung capacity, as primary endpoint. Besides being an important biomarker, FVC is also understood as a surrogate for all-cause mortality. Due to the rarity of the disease and the long time-period required to observe a number of deaths that is sufficient to draw meaningful conclusions from statistical analyses, mortality is impractical as a primary endpoint in clinical trials investigating new treatments. Although the surrogacy of FVC for mortality is biologically plausible and FVC is widely accepted as a surrogate for mortality, statistical evidence that supports the surrogacy assumption is still lacking.
The objective of this work was to apply state-of-the-art methodology for surrogate endpoint evaluation to interstitial lung disease trial data in order to explore the validity of FVC as an appropriate surrogate for mortality. In particular, two types of statistical modeling approaches were considered: joint models for longitudinal and time-to-event data and Bayesian bivariate meta-analysis.

Surrogacy levels

To assess the extent of evidence on a surrogacy relationship, Taylor and Elston (2009) defined three surrogacy levels:

- **Level 3**: evidence based on biological plausibility of the relationship
- **Level 2**: evidence of a general association between the surrogate and the clinical endpoint
- **Level 1**: evidence showing that the treatment effect on the surrogate endpoint is correlated with the treatment effect on the clinical endpoint across many randomized controlled trials

FVC fulfills the criteria for surrogacy level 3. Level 2 may be considered established by the recent work of Paterniti et al. (2017). Level 1 is not yet established as the relationship between the treatment effects on FVC and the treatment effects on mortality has not been analyzed statistically yet.

Joint models for longitudinal and time-to-event data

Joint models allow the combined analysis of longitudinal and time-to-event endpoints in a single model and the consideration of the relationship between them. The models were fitted to interstitial lung disease trial data in order to investigate in the association between FVC and mortality and potentially confirm surrogacy level 2. A joint model consists of two submodels: the longitudinal and the time-to-event submodel.
Considering a sample of n independent subjects, the longitudinal submodel is given by

$$\begin{aligned} y_i(t) &= m_i(t) + \varepsilon_i(t), \\ m_i(t) &= x_i^T(t)\beta + z_i^T(t)b_i, \\ b_i &\sim \mathcal{N}(0, D),\ \varepsilon_i(t) \sim \mathcal{N}(0, \sigma^2) \end{aligned}$$

for subjects $i = 1, ...n$, where $m_i(t)$ is the true value of the longitudinal marker of subject i at time t, the $y_i(t)$ are the observed values of the marker, the $x_i(t)$ are the design vectors of the fixed effects β, $z_i(t)$ are the design vectors of the random effects b_i and $\varepsilon_i(t)$ are the time-dependent measurement error terms. D denotes the covariance matrix of the random effects and σ^2 is the variance of the error terms.
The standard time-to-event submodel is given by

$$h_i(t|\ \mathcal{M}_i(t), w_i) = h_0(t)\exp\{\gamma^T w_i + \alpha m_i(t)\}$$

for $i = 1, ...n$, where $h_i(t)$ is the instantaneous risk or hazard for an event for subject i at time t, $\mathcal{M}_i(t)$ is the history of $m_i(.)$ up to t, w_i are the baseline covariate values of subject i with coefficients γ, α is the association parameter and $h_0(.)$ is the baseline hazard function.

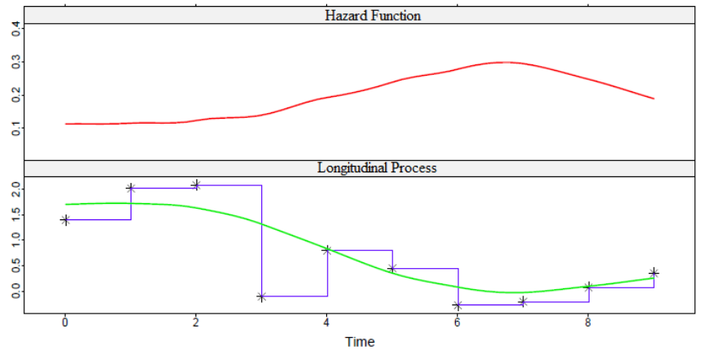

Figure 1: Intuitive representation of handling a time-varying covariate in an extended Cox model and in a joint model for longitudinal and time-to-event data

Figure 1 illustrates the concept of jointly modeling longitudinal data and time-to-event data in form of a hazard function. The blue graph represents the approach to handling longitudinal covariates in an extended Cox model, where the last observed values of are carried forward until a new value is observed. The green line corresponds to the way longitudinal outcomes are dealt with by joint models. Here, a function for the true level of the longitudinal outcome is fitted based on the observed values. The joint modeling approach is likely to result in more realistic values of the longitudinal marker then the last observation carried forward approach.
FVC level was modeled linearly in longitudinal submodel and risk of death was modeled in the time-to-event submodel. The models were fitted on a pooled data set that contain data from six interstitial lung disease trials investigating the effect of the drug nintedanib. Since the association parameter α then quantifies the effect of FVC level on risk of death, it is the value of interest for evaluating surrogacy of level 2.

Bivariate meta-analysis models

Bujkiewicz et al. (2019) proposed different bivariate meta-analytic models with random effects to estimate the correlation between the treatment effects on FVC and thereby investigate level 1 surrogacy.
The models consist of a within-study and a between-study model. The within-study model gives the distribution of the treatment effect estimates on both endpoints given the true effect estimates as a bivariate normal distribution. In the between-study model, the distribution of true study-specific effects is given. They are also assumed to be normally distributed and deviate from one another due to heterogeneity of the studies (e.g. differences in study design). For surrogate endpoint evaluation, the most essential part of the between-study model is given by

$$\theta_{2i}|\ \theta_{1i} \sim \mathcal{N}(\lambda_0 + \lambda_1\theta_{1i}, \psi_2^2)$$

for $i = 1, ..., k$, where k is the number of included studies, θ_{2i} is the true treatment effect on the endpoint of ultimate interest (in this case mortality) in study i, and θ_{1i} is the true treatment effect on the potential surrogate endpoint (in this case FVC) in study i. Then λ_0 is the expected effect on the endpoint of ultimate interest if there is no effect on the potential surrogate, λ_1 quantifies how an effect on the surrogate changes the expected effect on the endpoint of ultimate interest and ψ_2^2 is the conditional variance of the effect on the endpoint of ultimate interest given the effect on the surrogate endpoint. Hence, the parameters λ_0, λ_1 and ψ_2 are the most important when investigating level 1 surrogacy. In particular, $\lambda_1 \neq 0$ is a necessary condition for level 1 surrogacy.
Treatment effect estimates from six nintedanib and two pirfenidone trials were used. The FVC effect estimates were given by mean differences in percentage of the predicted FVC value between the treatment and the placebo groups and mortality effect estimates were given as log hazard ratios. The models were fitted in a Bayesian framework with the probabilistic programming language Stan. Using the same study data, treatment effect estimates based on different models were determined and applied for fitting the meta-analytic models to assess the sensitivity of the models. In addition, different approaches for handling the within-study correlations (i.e. the correlation of the effect estimates given the true study-specific effects) were applied, including different prior distributions and estimates based on a bootstrap procedure.

References

S. Bujkiewicz, F. Achana, T. Papanikos, R. D. Riley, and K. R. Abrams. *NICE DSU Technical Support Document 20: Multivariate meta-analysis of summary data for combining treatment effects on correlated outcomes and evaluating surrogate endpoints.* National Institute for Health and Care Excellence, 2019.

M. O. Paterniti, Y. Bi, D. Rekic, Y. Wang, B. A. Karmini-Shah, and B. A. Chowdhurry. Acute Exacerbation and Decline in Forced Vital Capacity Are Associated with Increased Mortality in Idiopathic Pulmonary Fibrosis. *Ann Am Thorac Soc*, 14(9):1395–1402, 2017.

R. S. Taylor and J. Elston. The use of surrogate outcomes in model-based cost-effectiveness analyses: A survey of UK Health Technology Assessment reports. *Health Technol Assess*, 13(8):1–50, Jan 2009