Methodology for analysis of joint longitudinal and time-to-event data from multiple studies

Maria Sudell, Dr Ruwanthi Kolamunnage-Dona,
Dr Catrin Tudur Smith

mesudell@liverpool.ac.uk

Definitions

Joint models for longitudinal and time-to-event data

- Methods to simultaneously model potentially related longitudinal and time-to-event data
- Often used to account for study dropout in longitudinal studies, to include a time varying covariate measured with error in time-to-event studies, or to investigate association between longitudinal and time-to-event outcomes
- Longitudinal sub-model and time-to-event sub-model linked through an association structure
- Many association structures exist – each with different interpretations. Here only random effects only proportional association considered

Meta-analysis

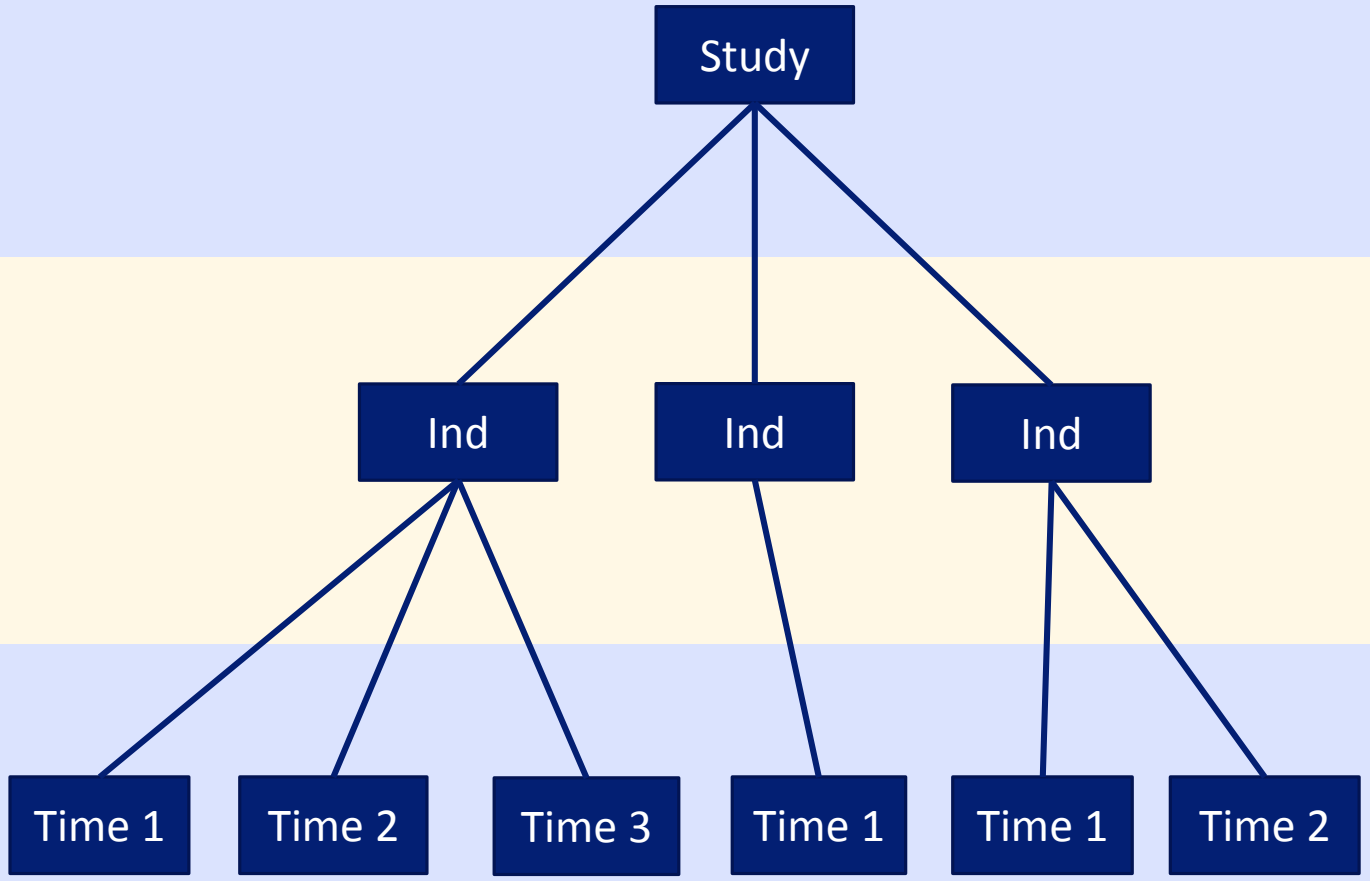
- Systematic pooling of results from multiple studies
- Allows increased precision, identification of effect sizes too small to be identified in single studies, and allows questions additional to those originally posed in the data to be answered
- Gold standard – Individual Participant/Patient Data (IPD) meta-analyses, where data for each individual recorded in studies identified in the meta-analysis is available.

The multi – study case, meta-analysis

Study k for k in $1 \ldots K$

Individual i for i in $1 \ldots n_k$

Longitudinal measurements at time points j for j in $1 \ldots m_{ki}$



Considerations with multi study joint data

- **Data from different studies likely to display heterogeneity, which needs to be accounted for**
  - Fixed interaction terms
  - Stratified baseline hazard in survival sub-model
  - Study level random effects

- **Large datasets** – issues fitting on individual computers

- **Estimation of standard errors**
  - Approximate standard error procedures – inaccurate if model includes study level random effects
  - Bootstrapping – time consuming

- **Estimation of random effects**
  - Difficulties with adaptive Gaussian quadrature – multi-level random effects potentially different specifications at each level
  - Pseudo – adaptive procedure used – calculate position and spread of random effects at each level based on the initial longitudinal fit.

Real Dataset – subset of INDANA dataset

- IPD from multiple studies investigating the effect of no treatment versus any treatment for hypertensive patients

- Longitudinal data measured at 6 months, then annually thereafter to maximum of 7 years. Measurement patterns varied between studies

- Using subset with data for longitudinal outcome systolic blood pressure (SBP) and time to death, data available from 6 studies. Data subset used as available computer didn't have sufficient memory for the full dataset. Proportions of individuals from each study, and proportions events/censored within each study kept same as full dataset

- Evidence of a changepoint in the data at 6 month, so time and treatment allowed different coefficients before and after changepoint

joineRmeta

- Package developed as extension to joineR package
- Currently one stage model permits
  - Random effects only proportional sharing structure
  - One continuous longitudinal outcome and one possibly censored time-to-event outcome
  - One changepoint permitted, with ability to specify variables to take different coefficients before and after. Changepoint is optional
  - Individual and study level random effects permitted, capped at 3 per level
  - Baseline hazard can be common across studies or stratified by study

- Package also contains (but not demonstrated here) :
  - Functions to easily plot multi-study joint data
  - Multi-study joint data simulation function
  - Function for second stage of two stage MA to pool joint model fits

Model group definitions – Group 0 – Naïve model

Longitudinal sub-model

$$Y_{ki} = \beta_{10} + \beta_{11} time + \beta_{12} treat + b_{0ki}^{(2)} + b_{1ki}^{(2)} time + \varepsilon_{ki}$$

Time-to-event sub-model

$$\lambda_{ki}(t) = \lambda_0(t) \exp(\beta_{21} treat + W_2)$$

Association Structure

$$W_2 \propto W_1$$

$$W_2 = \alpha^{(2)} \left(b_{0ki}^{(2)} + b_{1ki}^{(2)} time \right)$$

Key – model components accounting for between study heterogeneity

- Longitudinal fixed effect coefficients
- Time-to-event fixed effect coefficients
- Study level random effects
- Baseline hazard stratified by study

Model group definitions – Group 1 – Fixed effects

Longitudinal sub-model

$$Y_{ki} = \beta_{10} + \beta_{11} time + \beta_{12} treat + \beta_{13} study + \beta_{14} treat * study + b_{0ki}^{(2)} + b_{1ki}^{(2)} time + \varepsilon_{ki}$$

Time-to-event sub-model

$$\lambda_{ki}(t) = \lambda_0(t) \exp(\beta_{21} treat + \beta_{22} study + \beta_{23} treat * study + W_2)$$

Association Structure

$$W_2 \propto W_1$$

$$W_2 = \alpha^{(2)} \left(b_{0ki}^{(2)} + b_{1ki}^{(2)} time \right)$$

Key – model components accounting for between study heterogeneity

- Longitudinal fixed effect coefficients
- Time-to-event fixed effect coefficients
- Study level random effects
- Baseline hazard stratified by study

Model group definitions – Group 1 – Fixed effects

Longitudinal sub-model

$$Y_{ki} = \beta_{10} + \beta_{11} time + \beta_{12} treat + \beta_{13} study + \beta_{14} treat * study + b_{0ki}^{(2)} + b_{1ki}^{(2)} time + \varepsilon_{ki}$$

Time-to-event sub-model

$$\lambda_{ki}(t) = \lambda_0(t) \exp(\beta_{21} treat + \beta_{22} study + \beta_{23} treat * study + W_2)$$

Association Structure

$$W_2 \propto W_1$$

$$W_2 = \alpha^{(2)} \left(b_{0ki}^{(2)} + b_{1ki}^{(2)} time \right)$$

Key – model components accounting for between study heterogeneity

- Longitudinal fixed effect coefficients
- Time-to-event fixed effect coefficients
- Study level random effects
- Baseline hazard stratified by study

Model group definitions – Group 2 – Fixed effects and one study level random effect

Longitudinal sub-model

$$Y_{ki} = \beta_{10} + \beta_{11} time + \beta_{12} treat + \beta_{13} study + b_{0ki}^{(2)} + b_{1ki}^{(2)} time + b_{1k}^{(3)} treat + \varepsilon_{ki}$$

Time-to-event sub-model

$$\lambda_{ki}(t) = \lambda_0(t) \exp(\beta_{21} treat + \beta_{22} study + W_2)$$

Association Structure

$$W_2 \propto W_1$$

$$W_2 = \alpha^{(2)} \left(b_{0ki}^{(2)} + b_{1ki}^{(2)} time \right) + \alpha^{(3)} \left(b_{1k}^{(3)} treat \right)$$

Key – model components accounting for between study heterogeneity

- Longitudinal fixed effect coefficients
- Time-to-event fixed effect coefficients
- Study level random effects
- Baseline hazard stratified by study

Model group definitions – Group 2 – Fixed effects and one study level random effect

Longitudinal sub-model

$$Y_{ki} = \beta_{10} + \beta_{11} time + \beta_{12} treat + \beta_{13} study + b_{0ki}^{(2)} + b_{1ki}^{(2)} time + b_{1k}^{(3)} treat + \varepsilon_{ki}$$

Time-to-event sub-model

$$\lambda_{ki}(t) = \lambda_0(t) \exp(\beta_{21} treat + \beta_{22} study + W_2)$$

Association Structure

$$W_2 \propto W_1$$

$$W_2 = \alpha^{(2)} \left(b_{0ki}^{(2)} + b_{1ki}^{(2)} time \right) + \alpha^{(3)} \left(b_{1k}^{(3)} treat \right)$$

Key – model components accounting for between study heterogeneity

- Longitudinal fixed effect coefficients
- Time-to-event fixed effect coefficients
- Study level random effects
- Baseline hazard stratified by study

Model group definitions – Group 3 – Fixed effects and two study level random effects

Longitudinal sub-model

$$Y_{ki} = \beta_{10} + \beta_{11} time + \beta_{12} treat + b_{0ki}^{(2)} + b_{1ki}^{(2)} time + b_{0k}^{(3)} + b_{1k}^{(3)} treat + \varepsilon_{ki}$$

Time-to-event sub-model

$$\lambda_{ki}(t) = \lambda_0(t) \exp(\beta_{21} treat + W_2)$$

Association Structure

$$W_2 \propto W_1$$

$$W_2 = \alpha^{(2)} \left(b_{0ki}^{(2)} + b_{1ki}^{(2)} time \right) + \alpha^{(3)} \left(b_{0k}^{(3)} + b_{1k}^{(3)} treat \right)$$

Key – model components accounting for between study heterogeneity

- Longitudinal fixed effect coefficients
- Time-to-event fixed effect coefficients
- Study level random effects
- Baseline hazard stratified by study

Model group definitions – Group 3 – Fixed effects and two study level random effects

Longitudinal sub-model

$$Y_{ki} = \beta_{10} + \beta_{11} time + \beta_{12} treat + b_{0ki}^{(2)} + b_{1ki}^{(2)} time + b_{0k}^{(3)} + b_{1k}^{(3)} treat + \varepsilon_{ki}$$

Time-to-event sub-model

$$\lambda_{ki}(t) = \lambda_0(t) \exp(\beta_{21} treat + W_2)$$

Association Structure

$$W_2 \propto W_1$$

$$W_2 = \alpha^{(2)} \left(b_{0ki}^{(2)} + b_{1ki}^{(2)} time \right) + \alpha^{(3)} \left(b_{0k}^{(3)} + b_{1k}^{(3)} treat \right)$$

Key – model components accounting for between study heterogeneity

- Longitudinal fixed effect coefficients
- Time-to-event fixed effect coefficients
- Study level random effects
- Baseline hazard stratified by study

Model group definitions – Group 4 – Fixed effects, stratified baseline hazard

Longitudinal sub-model

$$Y_{ki} = \beta_{10} + \beta_{11} time + \beta_{12} treat + \beta_{13} study + \beta_{14} treat * study + b_{0ki}^{(2)} + b_{1ki}^{(2)} time + \varepsilon_{ki}$$

Time-to-event sub-model

$$\lambda_{ki}(t) = \lambda_{0k}(t) \exp(\beta_{21} treat + W_2)$$

Association Structure

$$W_2 \propto W_1$$

$$W_2 = \alpha^{(2)} \left(b_{0ki}^{(2)} + b_{1ki}^{(2)} time \right)$$

Key – model components accounting for between study heterogeneity

- Longitudinal fixed effect coefficients
- Time-to-event fixed effect coefficients
- Study level random effects
- Baseline hazard stratified by study

Model group definitions – Group 4 – Fixed effects, stratified baseline hazard

Longitudinal sub-model

$$Y_{ki} = \beta_{10} + \beta_{11} time + \beta_{12} treat + \beta_{13} study + \beta_{14} treat * study + b_{0ki}^{(2)} + b_{1ki}^{(2)} time + \varepsilon_{ki}$$

Time-to-event sub-model

$$\lambda_{ki}(t) = \lambda_{0k}(t) \exp(\beta_{21} treat + W_2)$$

Association Structure

$$W_2 \propto W_1$$

$$W_2 = \alpha^{(2)} \left(b_{0ki}^{(2)} + b_{1ki}^{(2)} time \right)$$

Key – model components accounting for between study heterogeneity

- Longitudinal fixed effect coefficients
- Time-to-event fixed effect coefficients
- Study level random effects
- Baseline hazard stratified by study

Model group definitions – Group 5 – Fixed effects, stratified baseline hazard, one study level random effect

Longitudinal sub-model

$$Y_{ki} = \beta_{10} + \beta_{11} time + \beta_{12} treat + \beta_{13} study + b_{0ki}^{(2)} + b_{1ki}^{(2)} time + b_{1k}^{(3)} treat + \varepsilon_{ki}$$

Time-to-event sub-model

$$\lambda_{ki}(t) = \lambda_{0k}(t) \exp(\beta_{21} treat + W_2)$$

Association Structure

$$W_2 \propto W_1$$

$$W_2 = \alpha^{(2)} \left(b_{0ki}^{(2)} + b_{1ki}^{(2)} time \right) + \alpha^{(3)} \left(b_{1k}^{(3)} treat \right)$$

Key – model components accounting for between study heterogeneity

- Longitudinal fixed effect coefficients
- Time-to-event fixed effect coefficients
- Study level random effects
- Baseline hazard stratified by study

Model group definitions – Group 5 – Fixed effects, stratified baseline hazard, one study level random effect

Longitudinal sub-model

$$Y_{ki} = \beta_{10} + \beta_{11} time + \beta_{12} treat + \beta_{13} study + b_{0ki}^{(2)} + b_{1ki}^{(2)} time + b_{1k}^{(3)} treat + \varepsilon_{ki}$$

Time-to-event sub-model

$$\lambda_{ki}(t) = \lambda_{0k}(t) \exp(\beta_{21} treat + W_2)$$

Association Structure

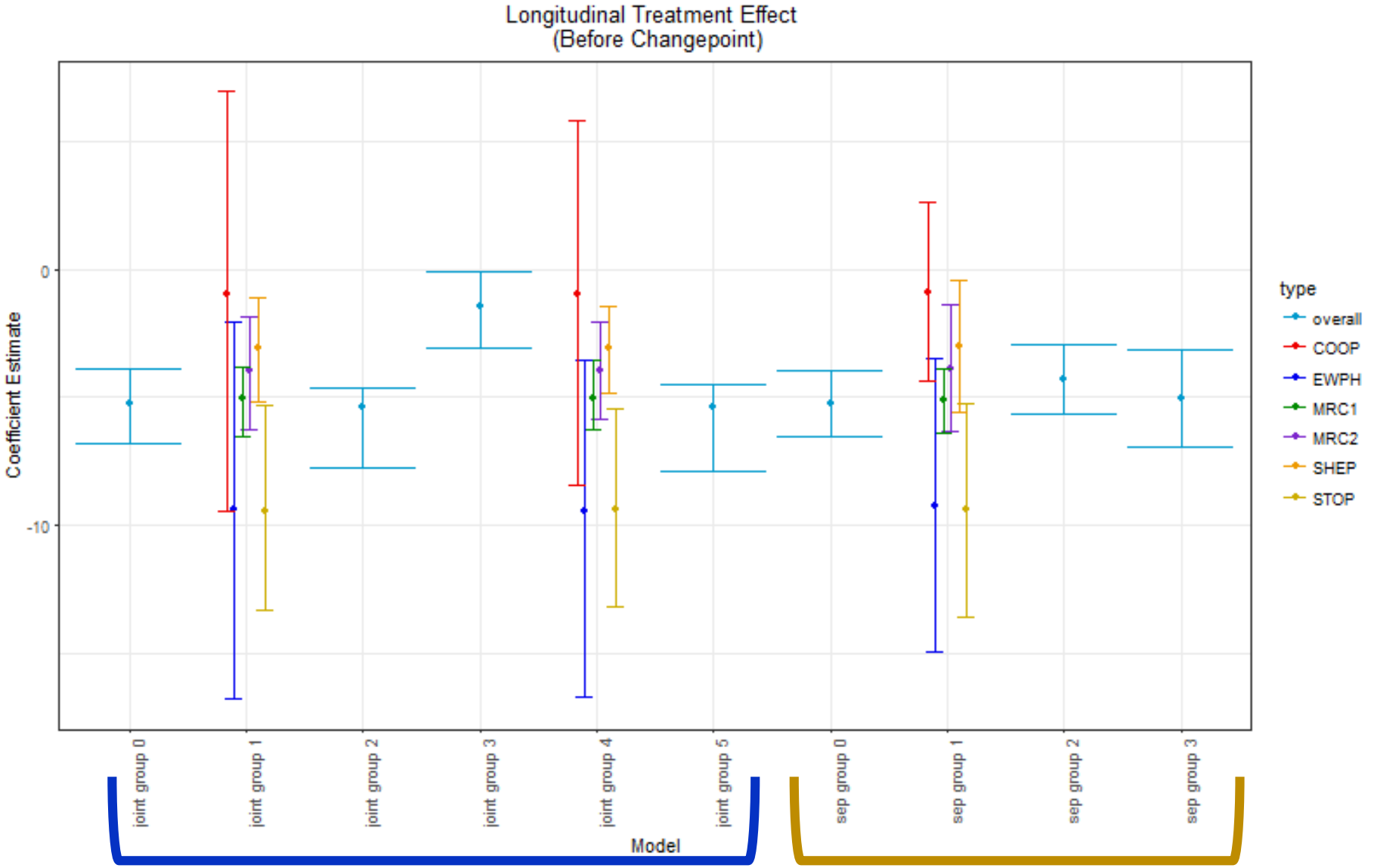
$$W_2 \propto W_1$$

$$W_2 = \alpha^{(2)} \left(b_{0ki}^{(2)} + b_{1ki}^{(2)} time \right) + \alpha^{(3)} \left(b_{1k}^{(3)} treat \right)$$

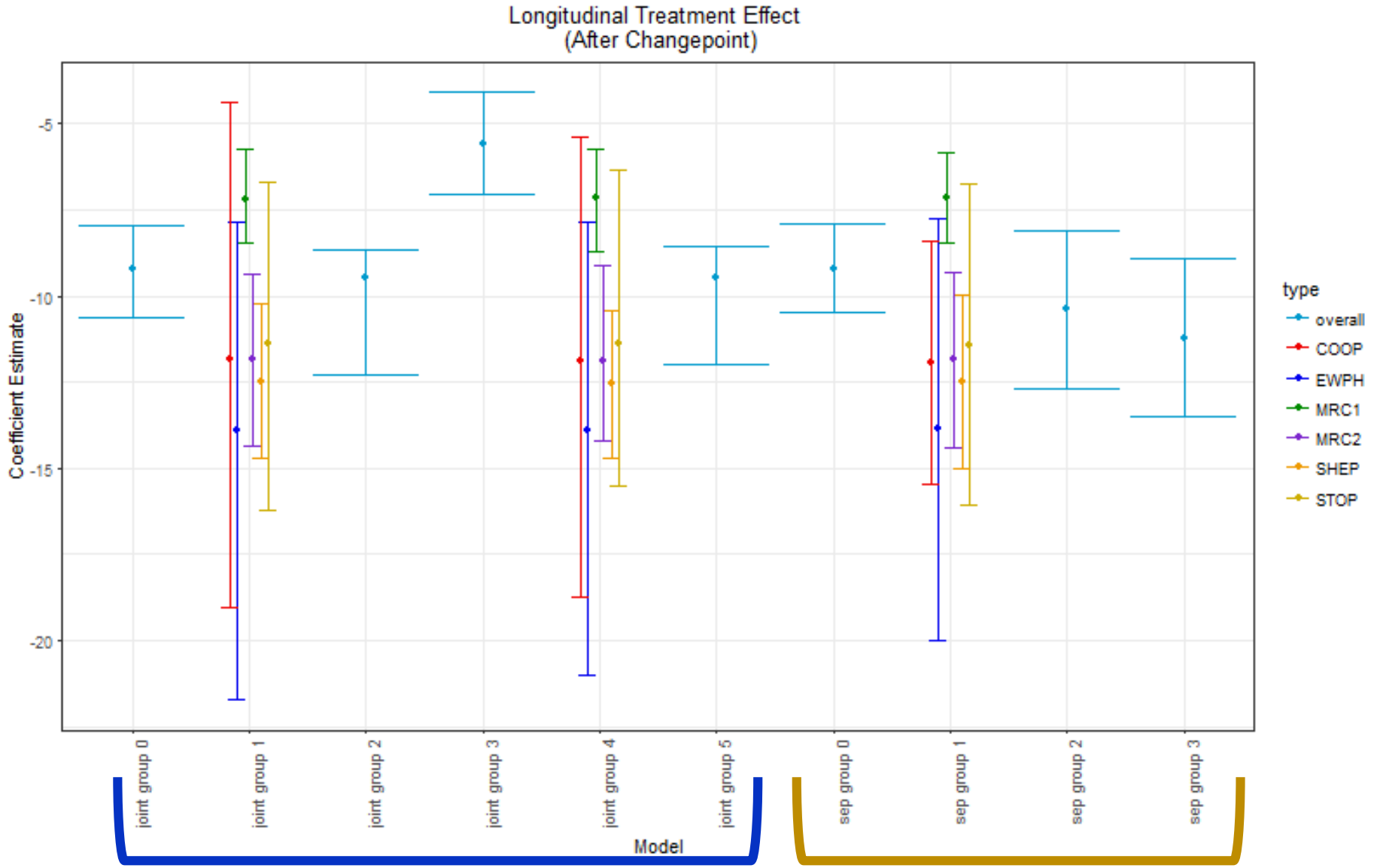
Key – model components accounting for between study heterogeneity

- Longitudinal fixed effect coefficients
- Time-to-event fixed effect coefficients
- Study level random effects
- Baseline hazard stratified by study

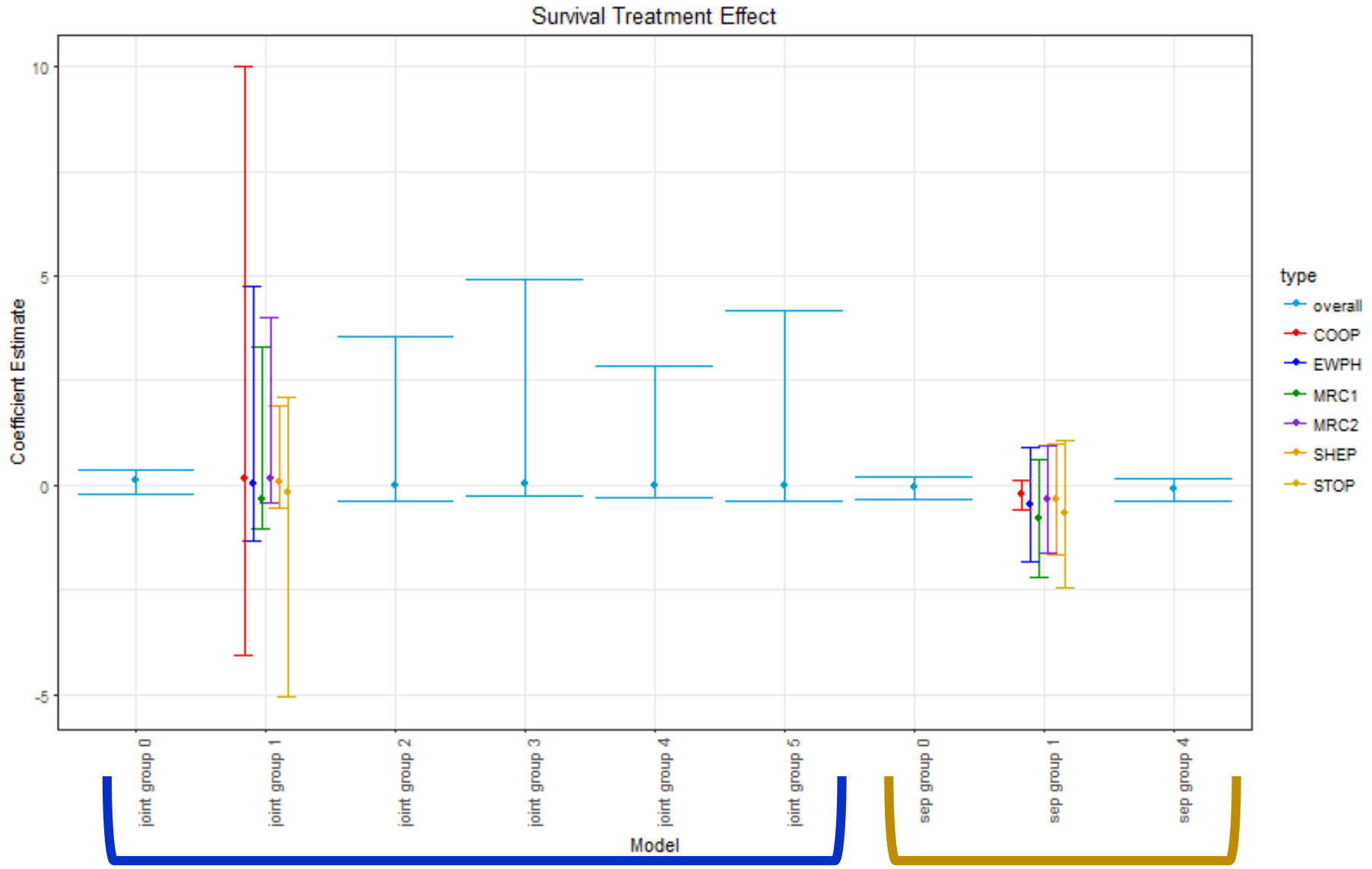
Results of one stage model fit – longitudinal treatment effect



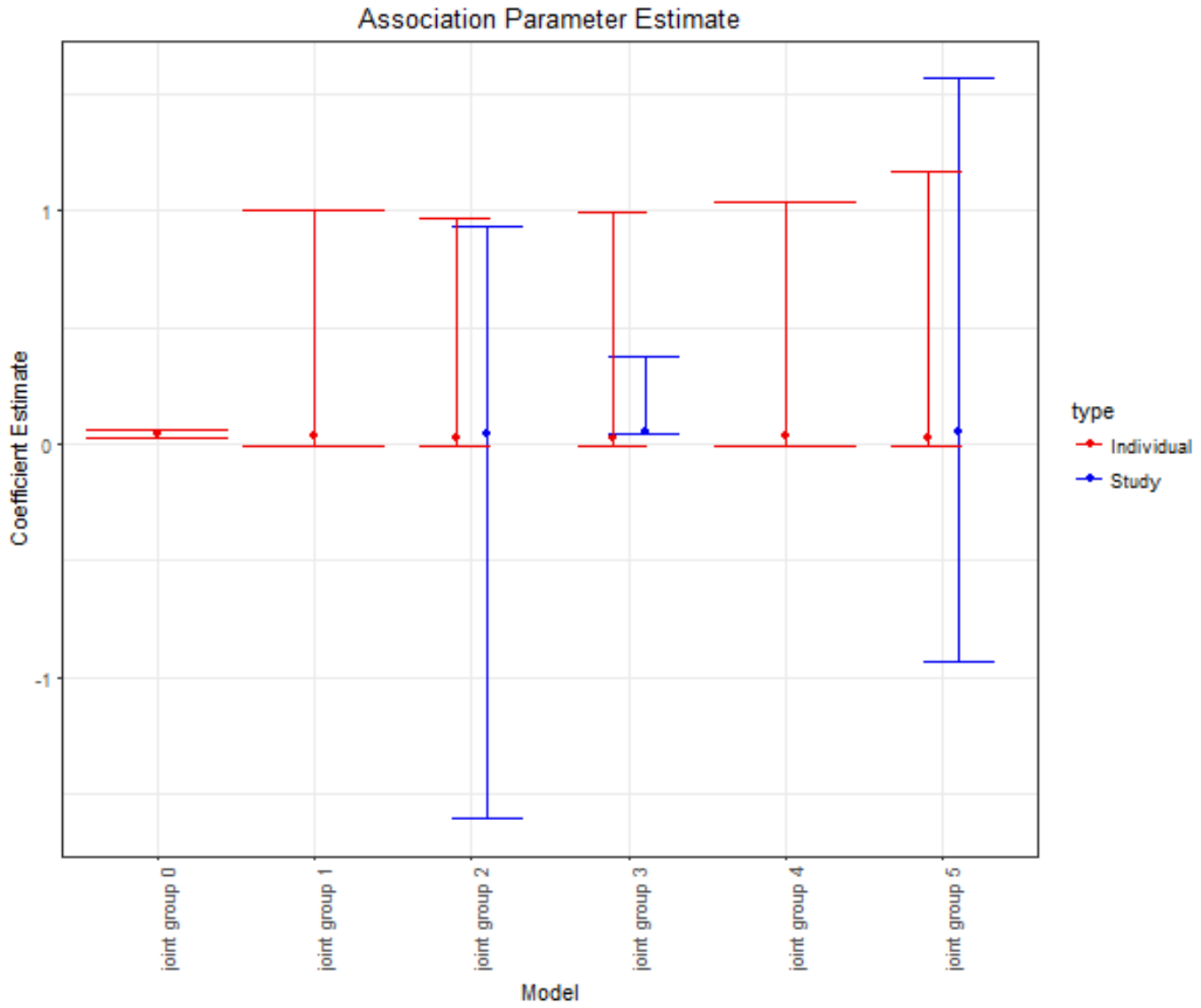
Results of one stage model fit – longitudinal treatment effect



Results of one stage model fit – time-to-event treatment effect



Results of one stage model fit – association parameter(s)



Simulation work

- Range of simulations being conducted to asses a variety of scenarios (ongoing)
  - Increasing number of studies
  - Increasing number of longitudinal measurements
  - Different levels of association at individual / study level
  - Simulated data contains no changepoint

- Initial observations
  - Issues with function with large datasets (due to sizes of some matrices used in survival sub-model)
  - One stage function appears to perform well for scenarios so far tested, but much more to be covered

Further work

- Continuation of simulation work
  - Varying number of studies, size of data, level of association
  - Model groups described here are being fitted to each scenario and examined to determine what model group preferable under a range of situations

- Expansion of joineRmeta package
  - Additional sharing structures (current value, first derivative,…)
  - Multiple longitudinal outcomes (e.g. modelling both systolic and diastolic blood pressure with time to death)
  - Improvement of efficiency (faster running, solving of memory issues with large datasets)

Conclusions

- Range of methods available to model multi-study joint data.
  - Methods available in R package, available on GitHub soon
  - Methods currently restricted to single longitudinal outcome, one type of association structure…
  - Further work needed to examine the association parameter estimation when changepoint included in the model.
  - Choice of type of model to use (fixed interaction terms, study level random effects) should be made based on data structure (e.g. number of included studies) and investigation aims (e.g. is it important to have study specific estimates of treatment effect?)

References

General

- Wulfsohn, M.S. and A.A. Tsiatis, *A Joint Model for Survival and Longitudinal Data Measured with Error*. 1997, International Biometric Society. p. 330.
- Henderson, R., P. Diggle, and A. Dobson, *Joint modelling of longitudinal measurements and event time data.* Biostatistics (Oxford, England), 2000. **1**(4): p. 465-480
- Hsieh, F., Y.-K. Tseng, and J.-L. Wang, *Joint Modeling of Survival and Longitudinal Data: Likelihood Approach Revisited.* Biometrics, 2006. **62**(4): p. 1037-1043
- Dempster, A.P., N.M. Laird, and D.B. Rubin, *Maximum Likelihood from Incomplete Data via the EM Algorithm*. 1977, Royal Statistical Society. p. 1.
- Rizopoulos, D., *Fast fitting of joint models for longitudinal and event time data using a pseudo-adaptive Gaussian quadrature rule.* Computational Statistics & Data Analysis, 2012. **56**(3): p. 491-501.
- R Core Team, *R: A Language and Environment for Statistical Computing*. 2015, R Foundation for Statistical Computing: Vienna, Austria.
- Philipson, P., et al., *joineR: Joint modelling of repeated measurements and time-to-event data*. 2012, R package version 1.0-3
- Lawrence Gould, A., et al., *Joint modeling of survival and longitudinal non-survival data: current methods and issues. Report of the DIA Bayesian joint modeling working group.* Stat Med, 2015. **34**(14): p. 2181-95.

INDANA Dataset

- Gueyffier, F., et al., *INDANA: a meta-analysis on individual patient data in hypertension. Protocol and preliminary results.* Therapie, 1995. **50**: p. 353-362.
- Amery, A., et al., *Mortality and morbidity results from the European Working Party on High Blood Pressure in the Elderly Trial.* Lancet, 1985. **i**: p. 1349-1354.
- Coope, J. and T.S. Warrender, *Randomised trial of treatment of hypertension in elderly patients in primary care.* BMJ, 1986. **293**: p. 1145-1151.
- Dalhöf, B., et al., *Morbidity and mortality in the Swedish Trial in Old Patients with Hypertension (STOP - Hypertension).* Lancet, 1991. **338**: p. 1281-1285.
- SHEP Cooperative Research Group, *Prevention of stroke by antihypertensive drug treatment in older persons with isolated systolic hypertension: final results of the Systolic Hypertension in the Elderly Program (SHEP).* JAMA, 1991. **265**: p. 3255-3264.
- Mitchell Perry Jr, H., et al., *Morbidity and mortality in the Systolic Hypertension in the Elderly Program (SHEP) Pilot Study.* Stroke, 1989. **20**: p. 4-13.
- Medical Research Council Working Party, *MRC Trial of treatment of mild hypertension: principal results.* BMJ, 1985. **291**: p. 97-104.
- MRC Working Party, *Medical Research Council trial of treatment of hypertension in older aldults: principal results.* BMJ, 1992. **304**: p. 405-412.

Thank you for listening
Any Questions?