Joint longitudinal and time-to-event models for multilevel hierarchical data

Sam Brilleman[1,2], Michael J Crowther[3], Margarita Moreno-Betancur[2,4,5], Jacqueline Buros Novik[6], James Dunyak[7], Nidal Al-Huniti[7], Robert Fox[7], Rory Wolfe[1,2]

39th Conference of the International Society for Clinical Biostatistics (ISCB)

Melbourne, Australia

26-30th August 2018

[1] Monash University, Melbourne, Australia

[2] Victorian Centre for Biostatistics (ViCBiostat), Melbourne, Australia

[3] University of Leicester, Leicester, UK

[4] Murdoch Childrens Research Institute, Melbourne, Australia

[5] University of Melbourne, Melbourne, Australia

[6] Icahn School of Medicine at Mount Sinai, New York, NY, USA

[7] AstraZeneca, Waltham, MA, USA

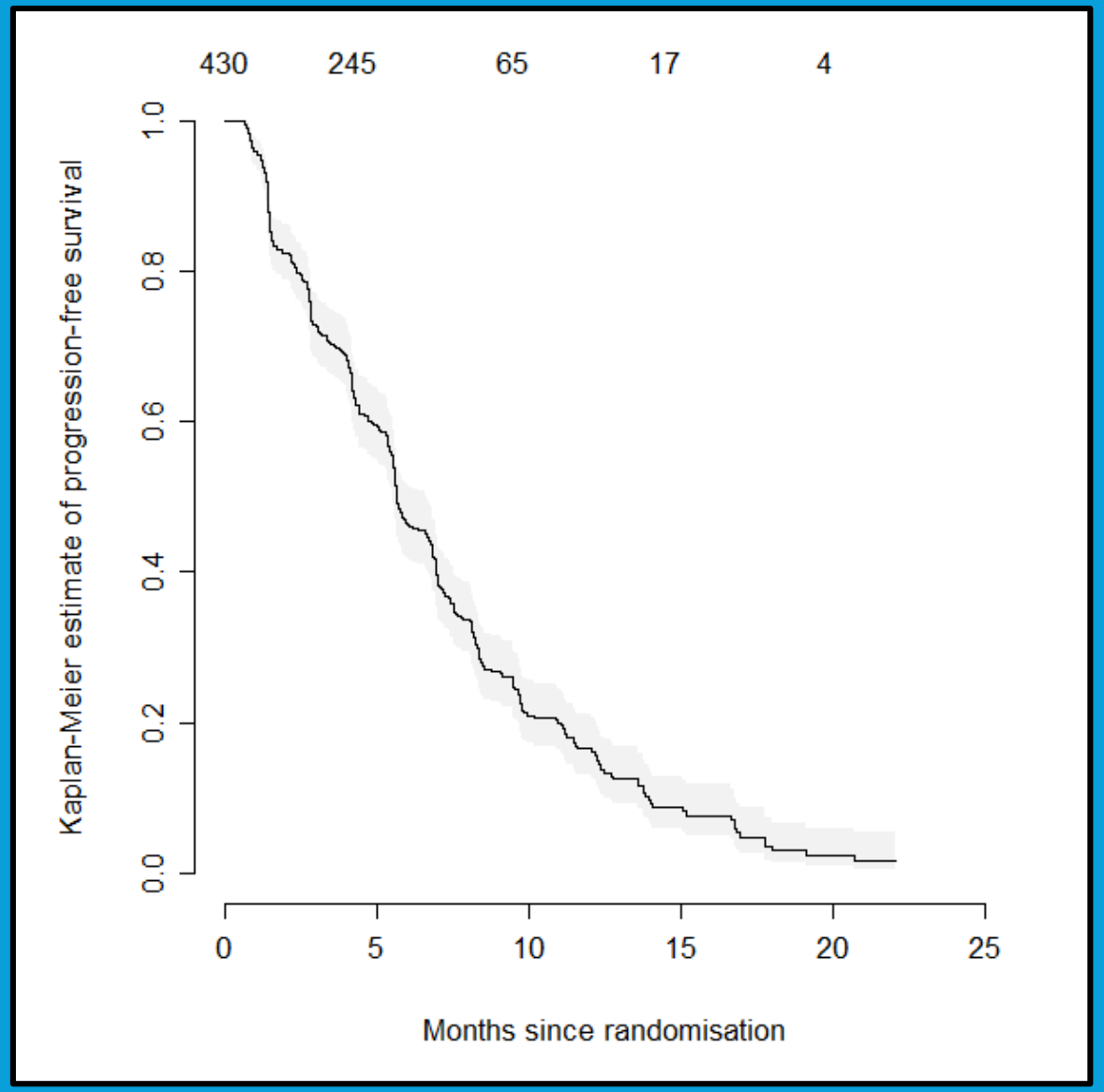
MONASH University ViCBiostat

Motivating application

- Data from the Iressa Pan-Asia Study (IPASS)
  - phase 3 trial of N = 1,217 untreated non-small cell lung cancer (NSCLC) patients in East Asia randomized to either (i) gefitinib or (ii) carboplatin + paclitaxel [1]
  - primary outcome was progression-free survival
  - main trial results suggested that an epidermal growth factor receptor (EGFR) mutation was associated with treatment response (i.e. treatment by subgroup interaction) [2]
- We performed a **secondary analysis** of data for the **N = 430 (35%) patients with known EGFR mutation status**
- We used a **joint modelling approach to explore how changes in tumor size are related to death or disease progression**

Outcome variables

- Time-to-event outcome:
  - progression-free survival

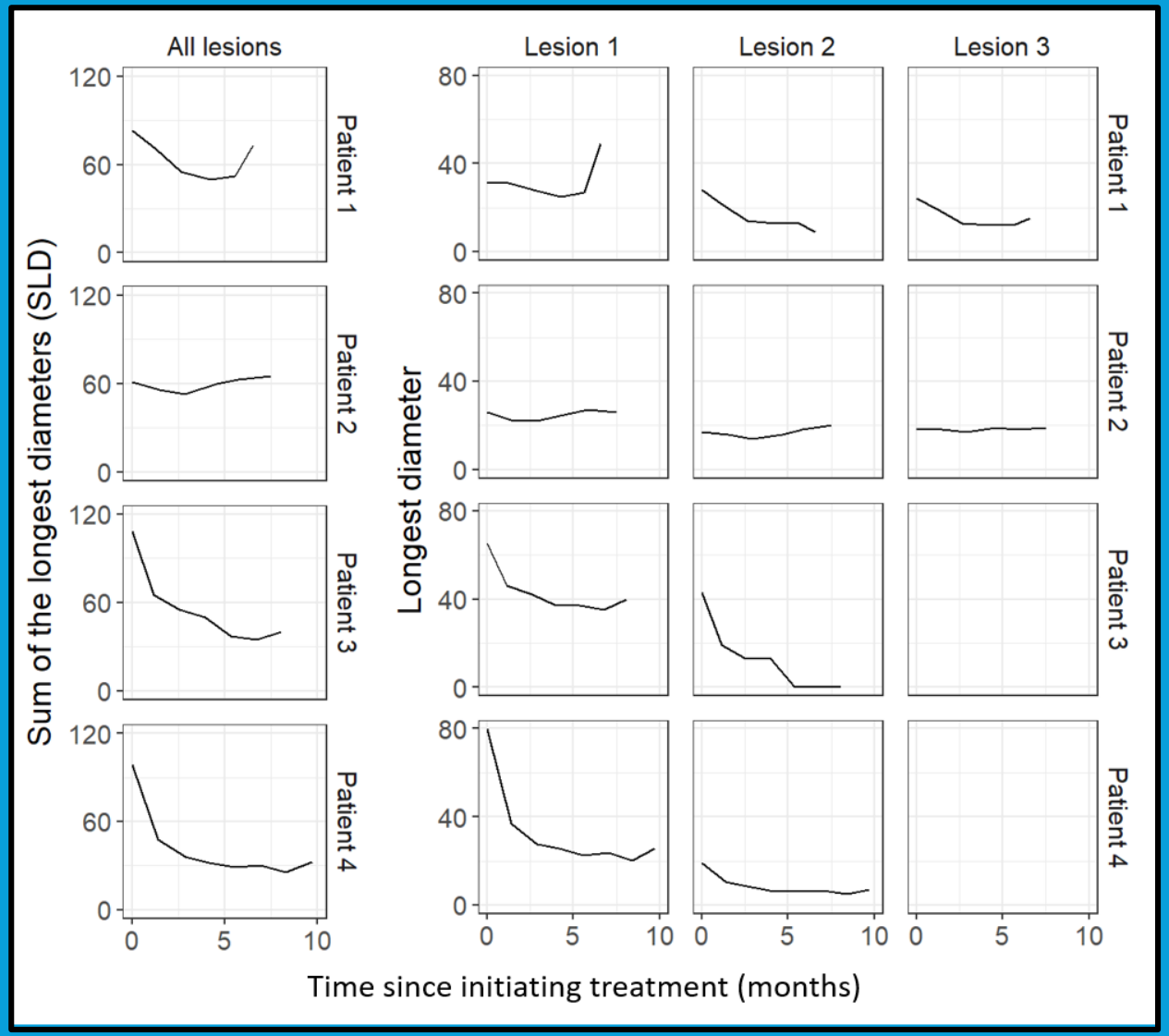


MONASH University VICBiostat

Outcome variables

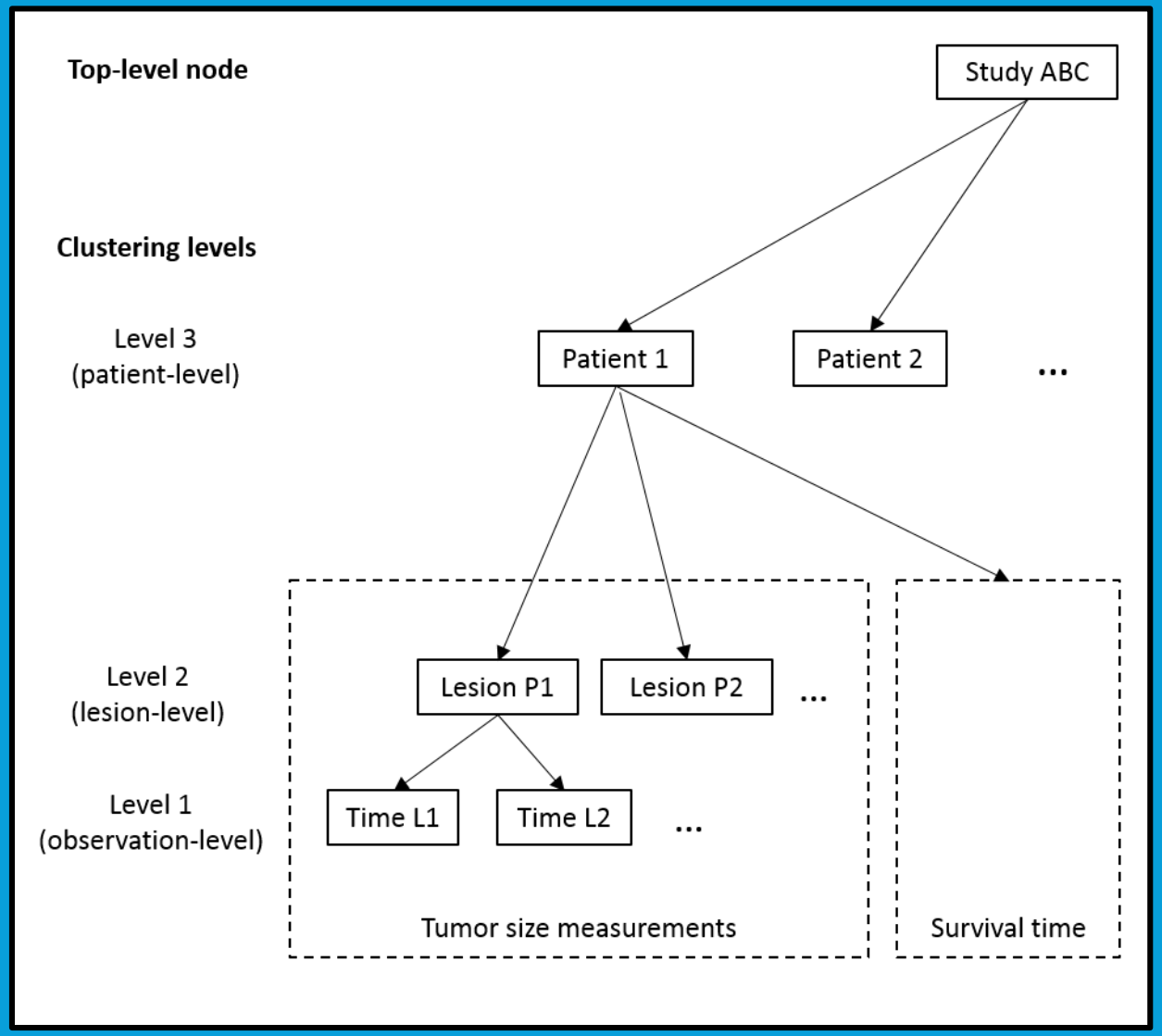
- Time-to-event outcome:
  - progression-free survival

- Longitudinal outcome:
  - tumor size, often captured through "**sum of the longest diameters**" (SLD) for target lesions defined at baseline
  - but can we do better?
  - why not model the (changes in the) longest diameter of the individual lesions rather than their sum?



Data structure

- Patients can have >1 tumor lesions
- The number of lesions might differ across patients
- There may not be any natural ordering for the lesions (i.e. they are exchangeable with respect to the correlation structure)
- Data contains a **three-level hierarchical structure** in which the longitudinal outcome (lesion diameter) is observed at:
  - time points < lesions < patients



Joint modelling

- Joint estimation of regression models which traditionally would have been estimated separately:
  - a **mixed effects model** for a longitudinal outcome (“longitudinal submodel”)
  - a **time-to-event model** for the time to an event of interest (“event submodel”)
  - the submodels are linked through **shared parameters**

Joint modelling

- Joint estimation of regression models which traditionally would have been estimated separately:
  - a **mixed effects model** for a longitudinal outcome (“longitudinal submodel”)
  - a **time-to-event model** for the time to an event of interest (“event submodel”)
  - the submodels are linked through **shared parameters**

- Most common shared parameter joint model has included one longitudinal outcome (a repeatedly measured “biomarker”) and one terminating event outcome

Joint modelling

- Joint estimation of regression models which traditionally would have been estimated separately:
  - a **mixed effects model** for a longitudinal outcome ("longitudinal submodel")
  - a **time-to-event model** for the time to an event of interest ("event submodel")
  - the submodels are linked through **shared parameters**

- Most common shared parameter joint model has included one longitudinal outcome (a repeatedly measured "biomarker") and one terminating event outcome

- However, **a vast number of extensions** have been proposed, for example:
  - competing risks, recurrent events, interval censored events, multiple longitudinal outcomes, …

Joint modelling

- Joint estimation of regression models which traditionally would have been estimated separately:
  - a **mixed effects model** for a longitudinal outcome ("longitudinal submodel")
  - a **time-to-event model** for the time to an event of interest ("event submodel")
  - the submodels are linked through **shared parameters**

- Most common shared parameter joint model has included one longitudinal outcome (a repeatedly measured "biomarker") and one terminating event outcome

- However, **a vast number of extensions** have been proposed, for example:
  - competing risks, recurrent events, interval censored events, multiple longitudinal outcomes, …

- But a **common aspect has been a two-level hierarchical data structure**:
  - longitudinal biomarker measurements are observed at time points (level 1) < patients (level 2)

A 3-level joint model

$y_{ijk}(t)$ is the observed diameter at time t for the
k^{th} time point ($k = 1, \dots, K_{ij}$)
clustered within the j^{th} lesion ($j = 1, \dots, J_i$)
clustered within the i^{th} patient ($i = 1, \dots, I$)
T_i is "true" event time, C_i is the censoring time
$T_i^* = \min(T_i, C_i)$ and $d_i = I(T_i \leq C_i)$

Longitudinal submodel

$$y_{ijk}(t) \sim N(\mu_{ijk}(t), \sigma_y^2)$$

$$\mu_{ijk}(t) = \boldsymbol{x'_{ijk}}(t)\boldsymbol{\beta} + \boldsymbol{z'_{ijk}}(t)\boldsymbol{b_i} + \boldsymbol{w'_{ijk}}(t)\boldsymbol{u_{ij}}$$

for fixed effect parameters $\boldsymbol{\beta}$, patient-specific parameters $\boldsymbol{b_i}$, and lesion-specific parameters $\boldsymbol{u_{ij}}$,

and assuming $\boldsymbol{b_i} \sim N(0, \boldsymbol{\Sigma}_b)$, $\boldsymbol{u_{ij}} \sim N(0, \boldsymbol{\Sigma}_u)$, $\text{Corr}(\boldsymbol{b_i}, \boldsymbol{u_{ij}}) = 0$

A 3-level joint model

$y_{ijk}(t)$ is the observed diameter at time t for the
k^{th} time point ($k = 1, \dots, K_{ij}$)
clustered within the j^{th} lesion ($j = 1, \dots, J_i$)
clustered within the i^{th} patient ($i = 1, \dots, I$)
T_i is "true" event time, C_i is the censoring time
$T_i^* = \min(T_i, C_i)$ and $d_i = I(T_i \leq C_i)$

Longitudinal submodel

$$y_{ijk}(t) \sim N(\mu_{ijk}(t), \sigma_y^2)$$

$$\mu_{ijk}(t) = \boldsymbol{x'_{ijk}}(t)\boldsymbol{\beta} + \boldsymbol{z'_{ijk}}(t)\boldsymbol{b_i} + \boldsymbol{w'_{ijk}}(t)\boldsymbol{u_{ij}}$$

for fixed effect parameters $\boldsymbol{\beta}$, patient-specific parameters $\boldsymbol{b_i}$, and lesion-specific parameters $\boldsymbol{u_{ij}}$,

and assuming $\boldsymbol{b_i} \sim N(0, \boldsymbol{\Sigma}_b)$, $\boldsymbol{u_{ij}} \sim N(0, \boldsymbol{\Sigma}_u)$, $\text{Corr}(\boldsymbol{b_i}, \boldsymbol{u_{ij}}) = 0$

Event submodel

$$h_i(t) = h_0(t) \exp\left(\boldsymbol{v'_i}(t)\boldsymbol{\gamma} + \sum_{q=1}^{Q} \alpha_q \, f_q(\boldsymbol{\beta}, \boldsymbol{b_i}, \boldsymbol{u_{ij}};\ j = 1, \dots, J_i) \right)$$

for fixed effect parameters $\boldsymbol{\gamma}$ and α_q $(q = 1, \dots, Q)$, and some set of functions $f_q(.)$ applied to the J_i

lesion-specific quantities (e.g. expected values or slopes) for the i^{th} patient at time t.

A 3-level joint model

$y_{ijk}(t)$ is the observed diameter at time t for the
k^{th} time point ($k = 1, \dots, K_{ij}$)
clustered within the j^{th} lesion ($j = 1, \dots, J_i$)
clustered within the i^{th} patient ($i = 1, \dots, I$)
T_i is "true" event time, C_i is the censoring time
$T_i^* = \min(T_i, C_i)$ and $d_i = I(T_i \leq C_i)$

Longitudinal submodel

$$y_{ijk}(t) \sim N(\mu_{ijk}(t), \sigma_y^2)$$

$$\mu_{ijk}(t) = \boldsymbol{x'_{ijk}}(t)\boldsymbol{\beta} + \boldsymbol{z'_{ijk}}(t)\boldsymbol{b_i} + \boldsymbol{w'_{ijk}}(t)\boldsymbol{u_{ij}}$$

for fixed effect parameters $\boldsymbol{\beta}$, patient-specific parameters $\boldsymbol{b_i}$, and lesion-specific parameters $\boldsymbol{u_{ij}}$,

and assuming $\boldsymbol{b_i} \sim N(0, \boldsymbol{\Sigma}_b)$, $\boldsymbol{u_{ij}} \sim N(0, \boldsymbol{\Sigma}_u)$, $\text{Corr}(\boldsymbol{b_i}, \boldsymbol{u_{ij}}) = 0$

Event submodel

$$h_i(t) = h_0(t) \exp\left(\boldsymbol{v'_i}(t)\boldsymbol{\gamma} + \sum_{q=1}^{Q} \alpha_q f_q(\boldsymbol{\beta}, \boldsymbol{b_i}, \boldsymbol{u_{ij}};\ j = 1, \dots, J_i) \right)$$

"association structure" for the joint model

for fixed effect parameters $\boldsymbol{\gamma}$ and α_q $(q = 1, \dots, Q)$, and some set of functions $f_q(.)$ applied to the J_i

lesion-specific quantities (e.g. expected values or slopes) for the i^{th} patient at time t.

Association structures

- The association structure for the joint model is determined by $f_q\left(\boldsymbol{\beta}, \boldsymbol{b_i}, \boldsymbol{u_{ij}};\ j = 1, \ldots, J_i\right)$, for $q = 1, \ldots, Q$

Association structures

- The association structure for the joint model is determined by $f_q(\boldsymbol{\beta}, \boldsymbol{b_i}, \boldsymbol{u_{ij}};\ j = 1, \ldots, J_i)$, for $q = 1, \ldots, Q$
- There are two aspects to consider:
    1. Need to define which **aspect of the longitudinal trajectory** we want to be associated with the (log) hazard of the event, for example, expected size of the lesion $\left(\mu_{ij}(t)\right)$ or rate of change in size of the lesion $\left(\frac{d\mu_{ij}(t)}{dt}\right)$

Association structures

- The association structure for the joint model is determined by $f_q(\boldsymbol{\beta}, \boldsymbol{b_i}, \boldsymbol{u_{ij}};\ j = 1, \ldots, J_i)$, for $q = 1, \ldots, Q$
- There are two aspects to consider:
    1. Need to define which **aspect of the longitudinal trajectory** we want to be associated with the (log) hazard of the event, for example, expected size of the lesion $\left(\mu_{ij}(t)\right)$ or rate of change in size of the lesion $\left(\frac{d\mu_{ij}(t)}{dt}\right)$
    2. Need to define the set of functions $f_q(.)$ that determine how we **combine information across lesions** clustered within a patient into some form of patient-level summary, for example, **sum, mean, max** or **min**

Association structures

- The association structure for the joint model is determined by $f_q(\boldsymbol{\beta}, \boldsymbol{b_i}, \boldsymbol{u_{ij}};\ j = 1, \ldots, J_i)$, for $q = 1, \ldots, Q$
- There are two aspects to consider:
  1. Need to define which **aspect of the longitudinal trajectory** we want to be associated with the (log) hazard of the event, for example, expected size of the lesion $\left(\mu_{ij}(t)\right)$ or rate of change in size of the lesion $\left(\frac{d\mu_{ij}(t)}{dt}\right)$
  2. Need to define the set of functions $f_q(.)$ that determine how we **combine information across lesions** clustered within a patient into some form of patient-level summary, for example, **sum, mean, max** or **min**

- For example, consider the following definitions for $f_q(\boldsymbol{\beta}, \boldsymbol{b_i}, \boldsymbol{u_{ij}};\ j = 1, \ldots, J_i)$

$$\sum_{j=1}^{J_i} \mu_{ij}(t) \longrightarrow \text{“total tumor burden” for patient } i \text{ at time } t$$

$$\max\left(\frac{d\mu_{ij}(t)}{dt}; j = 1, \ldots, J_i\right) \longrightarrow \text{fastest growing lesion for patient } i \text{ at time } t;$$
e.g. the one that escaped treatment and will drive disease progression?

Model specification

- Longitudinal submodel
  - Fixed effect covariates:
    - 3 category group variable (EGFR+; EGFR- with carboplatin plus paclitaxel; EGFR- with gefitinib)
    - Linear and quadratic terms for time (orthogonalised)
    - Interaction between group and the linear & quadratic terms
  - Random effect covariates:
    - Patient-level: random intercept
    - Lesion-level: random intercept, linear and quadratic terms for time

Model specification

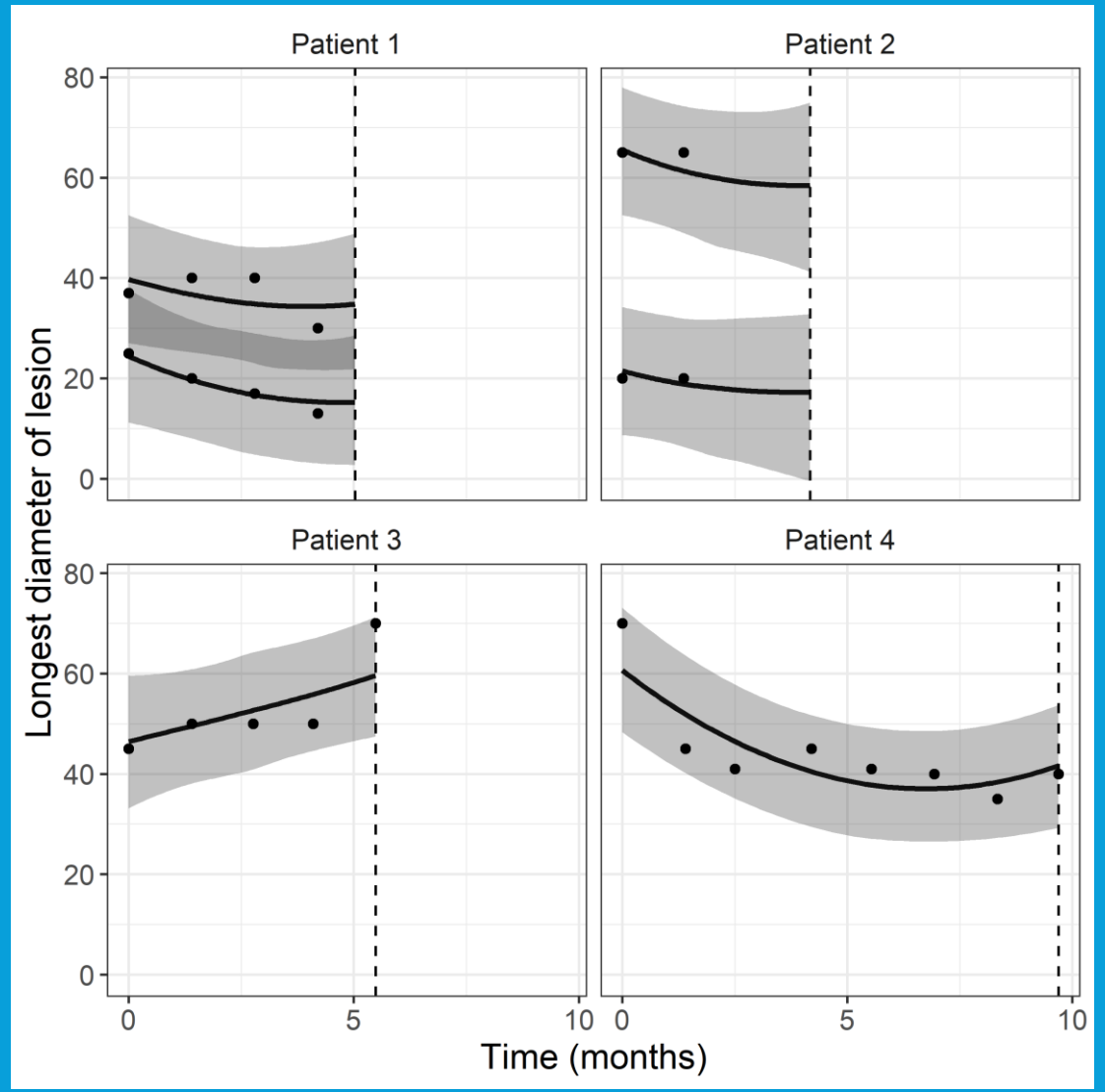
- Longitudinal submodel
  - Fixed effect covariates:
    - 3 category group variable (EGFR+; EGFR- with carboplatin plus paclitaxel; EGFR- with gefitinib)
    - Linear and quadratic terms for time (orthogonalised)
    - Interaction between group and the linear & quadratic terms
  - Random effect covariates:
    - Patient-level: random intercept
    - Lesion-level: random intercept, linear and quadratic terms for time
- Event submodel
  - B-splines used to model the log baseline hazard
  - Fixed effect covariates:
    - 3 category physical functioning measure (normal activity; restricted activity; in bed >50% of the time)
  - Association structure: sum, mean, min, or max of the lesion-specific values and/or slopes

Model estimation

- Estimated under a Bayesian approach, with prior distributions on all unknown parameters
- Implemented as part of the *stan_jm* modelling function in the **rstanarm** R package [3,4]
- The user can easily specify the hierarchical joint model using customary R formula syntax and data frames
- Various options for model fitting as well as post-estimation tools

https://cran.r-project.org/package=rstanarm

https://github.com/stan-dev/rstanarm



Model comparison

- We compared models with different association structures using a **time-dependent AUC measure** [5], adapted to the three-level hierarchical setting
- To calculate the AUC measure we used each patient's longitudinal **biomarker data up to 5 months**, and then **predicted their event status at 10 months**
- Overall predictive performance was poor, however:
  - the **smallest and slowest** growing lesion provided the worst predictive performance, and
  - the **largest and fastest** growing lesion provided the "best" predictive performance

Association structure	Time-dependent AUC
No biomarker data (i.e. no association structure)	0.50
Lesion-specific value	
Sum	0.62
Average	0.56
Maximum	0.61
Minimum	0.55
Lesion-specific value & slope	
Sum	0.65
Average	0.64
Maximum	0.66
Minimum	0.59

Abbreviations. AUC: area under the (receiver operating characteristic) curve.

Summary

- Joint modelling approaches have previously been limited to a two-level hierarchical data structure
- However, many clinical research settings present us with data that has additional levels of clustering
- Our proposed approach models the longitudinal measurements for lower-level clusters, and combines them into a patient-level summary that we assume is associated with the event rate
- From an inferential perspective, the method allows for association structures that would not have otherwise been possible
- From a model performance perspective, the method can potentially improve model fit since it provides greater flexibility, i.e. we can directly model the longitudinal trajectories for distinct lower-level units clustered within a patient
- The method has been implemented in general-purpose, freely-accessible, user-friendly software

Thank you


sam.brilleman@monash.edu


https://www.sambrilleman.com


@sambrilleman

References

[1] Mok TS et al. Gefitinib or Carboplatin–Paclitaxel in Pulmonary Adenocarcinoma. *New England Journal of Medicine.* 2009; 361: 947–957

[2] Fukuoka M et al. Biomarker Analyses and Final Overall Survival Results From a Phase III, Randomized, Open-Label, First-Line Study of Gefitinib Versus Carboplatin/Paclitaxel in Clinically Selected Patients With Advanced Non–Small-Cell Lung Cancer in Asia (IPASS). *Journal of Clinical Oncology.* 2011; 29: 2866–2874

[3] Stan Development Team. 2018. *rstanarm: Bayesian applied regression modeling via Stan*. R package version 2.17.4. http://mc-stan.org/rstanarm

[4] Brilleman SL et al. Joint longitudinal and time-to-event models via Stan. *In: Proceedings of StanCon 2018*. Pacific Grove, CA, USA. DOI: 10.5281/zenodo.1284334

[5] Rizopoulos D. Dynamic Predictions and Prospective Accuracy in Joint Models for Longitudinal and Time-to-Event Data. *Biometrics.* 2011; 67: 819–829.

MONASH University VICBiostat