Predicting Patient Reported Outcome Measures (PROMs) Using Routinely Collected Healthcare Data

Joseph Taggart[1], Dr Helen Parsons[2], Professor Martyn Plummer[1], Professor Rhiannon K. Owen[3], Professor Keith R. Abrams[1,2]

[1]Department of Statistics, University of Warwick; [2]Warwick Medical School, University of Warwick; [3]Population Data Science, Swansea University Medical School, University of Swansea

Background:

- Real-world data is increasingly used in Health Technology Assessment, to complement data from randomised controlled trials (RCTs).
- Trials and designed observational studies can determine which Patient Reported Outcome Measures (PROMs) are collected, routinely collected electronic health record (EHR) data often does not contain such measures.
- Hence, use of EHR data is limited due to the lack of quality of life related PROMs, used for NHS decision making by NICE and SMC.
- Trials are not inclusive; women, minority ethnic groups, the elderly, and individuals living with multiple long-term conditions are often under-represented.
- Decisions on treatments are informed by evidence that does not reflect the experiences of large proportions of the population, contributing to health inequalities.
- If PROMs were routinely available in EHR data, decision makers would have access to a broader, more diverse patient population.

Objective:

Predict the 4-dimensional quality of life PROM CASP-19 (Control, Autonomy, Self-Realisation, Pleasure) from routinely collected variables (sex, ethnicity, body mass index, comorbidities, Index of Multiple Deprivation) using waves 1-6 (2002-2012) of the English Longitudinal Study of Ageing (ELSA), with approximately 10,000 individuals and 31,000 observations.

Methods:

Multiple Bayesian mixed-effects models were fit, and their predictive accuracy assessed using MSE.

Model 1 consisted of just a linear mixed-effects model, with overall CASP-19 score as the longitudinal outcome.

$$y_i(t) = \mu_i(t) + \epsilon_i(t), \quad \epsilon_i(t) \sim \mathcal{N}(0, \sigma^2)$$

$$\mu_i(t) = \beta_0 + b_{0i} + (\beta_1 + b_{1i})t + \boldsymbol{\beta}^T \mathbf{X}_i(t)$$

with $b_{0i} \sim \mathcal{N}(0, \sigma_{b_0}^2)$ an individual level random intercept, $b_{1i} \sim \mathcal{N}(0, \sigma_{b_1}^2)$ an individual level random slope, and $\mathbf{X}_i(t)$ a vector of covariates for individual i.

Model 2 then included a survival sub-model for time to death, using the current value of the longitudinal outcome as a covariate in the survival model:

$$y_i(t) = \mu_i(t) + \epsilon_i(t), \quad \epsilon_i(t) \sim \mathcal{N}(0, \sigma^2)$$

$$\mu_i(t) = \beta_0 + b_{0i} + (\beta_1 + b_{1i})t + \boldsymbol{\beta}^T \mathbf{X}_i(t)$$

$$\lambda_i(t) = \lambda_0 \exp\left(\boldsymbol{\alpha}^T \mathbf{A}_i + \gamma \mu_i(t)\right)$$

where $\mathbf{A}_i$ denotes a vector of baseline covariates, γ the association parameter linking the longitudinal and survival processes via a current-value association, and λ_0 a constant baseline hazard.

Model 3 was a joint model with each CASP-19 subdimension modelled separately:

$$y_i(t) = \sum_{j=1}^{4} \mu_{i,j}(t) + \epsilon_i(t), \quad \epsilon_i(t) \sim \mathcal{N}(0, \sigma^2)$$

$$\mu_{ij}(t) = \beta_{0j} + b_{0ij} + \left(\beta_{1j} + b_{1ij}\right)t + \boldsymbol{\beta}_j^T \mathbf{X}_i(t)$$

$$\lambda_i(t) = \lambda_0 \exp\left(\boldsymbol{\alpha}^T \mathbf{A}_i + \textstyle\sum_{j=1}^{4} \gamma_j \mu_{ij}(\mathrm{t})\right)$$

Here, each subdimension is has its own association parameter γ_j.

Why Bayesian? Models were estimated using Markov Chain Monte Carlo (MCMC) using JAGS (Just Another Gibbs Sampler) to allow for full quantification of the uncertainty around parameter estimates and predictions.

Results:

Each model was evaluated using Monte Carlo cross-validation, where subsets of the dataset were randomly sampled to be trained on, whilst the remaining data was used for testing.

Model 1 was expectedly the worst performing model, with a MSE of approximately 140.

Model 2 has an MSE of 130, suggesting that accounting for mortality improves predictive accuracy.

Model 3 had an MSE of 87. This was a large improvement by modelling each of the CASP-19 subdimensions.

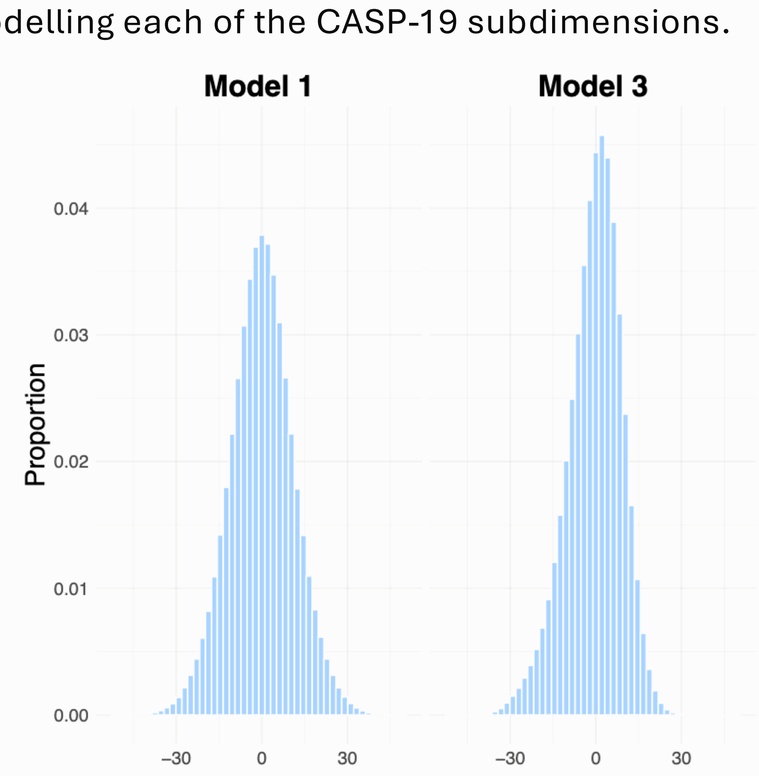

Figure 1: Differences between actual and predicted CASP-19 scores.

Quality of life PROMs can be reliably predicted from routinely collected healthcare data using Bayesian mixed-effects models. Increasing model complexity, by jointly modelling mortality and CASP-19 subdimensions, led to gains in predictive accuracy.

Future Work

- Exploration of Bayesian approximation methods, particularly Variational Inference, to jointly model all 19 individual CASP-19 questions, due to the high computational burden of MCMC.
- Future application to routinely collected population primary and secondary care data using SAIL Databank (NIHR Doctoral Award application).

Funded by NIHR | National Institute for Health and Care Research