The University Of Sheffield.

School Of Health And Related Research

DO FLEXIBLE SURVIVAL MODELS PROVIDE BETTER EXTRAPOLATIONS THAN STANDARD SURVIVAL MODELS? A SIMULATION STUDY.

Kearns B[1], Stevenson M[1], Triantafyllopoulos K[1], Manca A[2],
[1]The University of Sheffield [2]The University of York

Background

Health technology funding decisions often require accurate extrapolations of future survival. There is guidance on the use of standard survival models for extrapolation, and growing interest in the use of more flexible approaches. However, it is unclear if more flexible models provide more accurate extrapolations than standard models. **The aim of this study** was to assess the extrapolation performance of the more flexible models, compared against standard models, using a simulation study [1].

Simulation methods

Data-generating mechanism:

A mixture-Weibull model representing 2 sub-groups with short and long survival [2]. We varied length of follow-up (2, 3 and 4 years) and sample size (100, 300 and 600 patients), with 200 simulations per scenario. Figure 1 shows true values (black line) and simulations (grey lines).

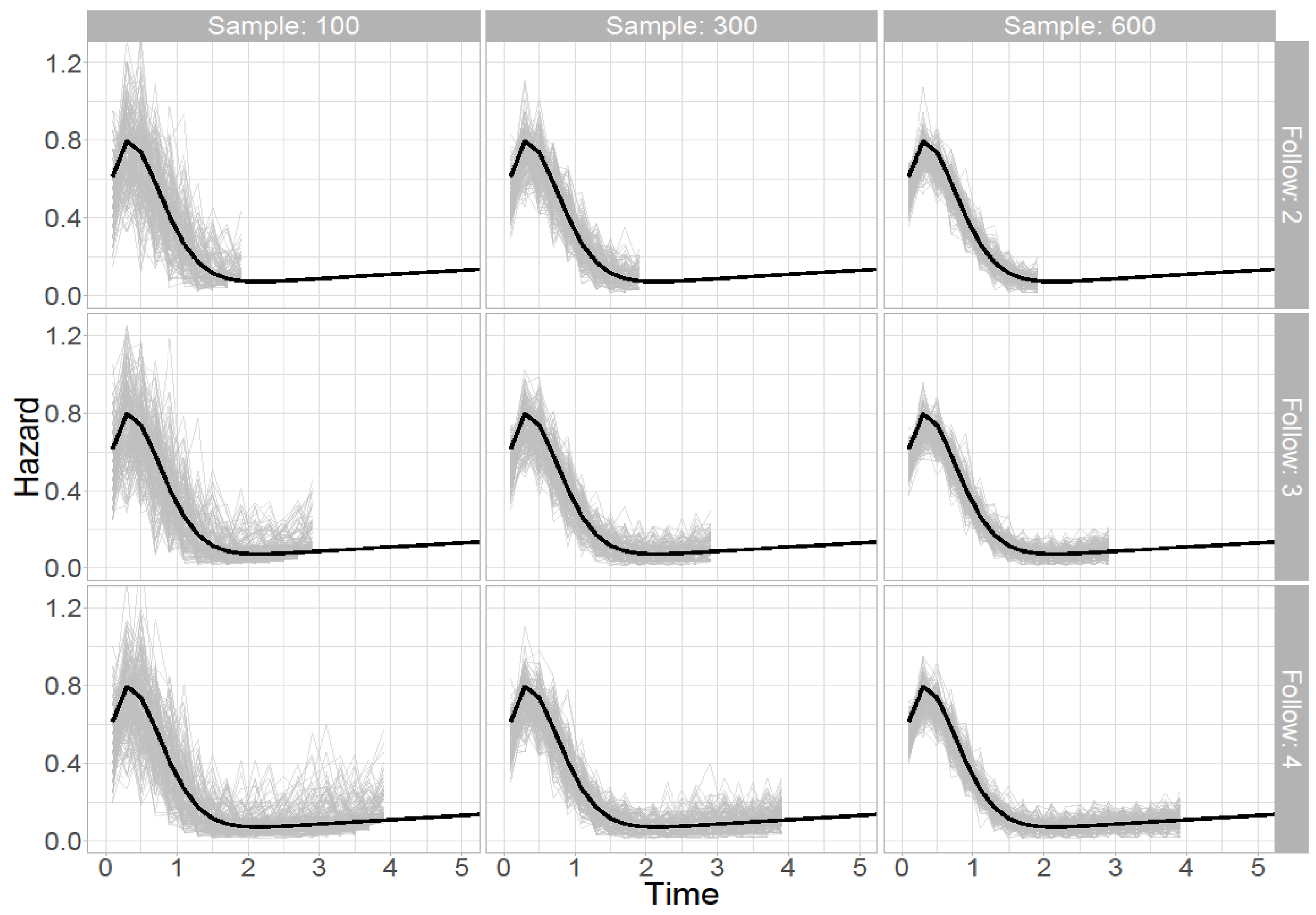
Models:

We considered four classes of survival model. For more details see Kearns *et al* (2019) [3].

1. Standard practice: *exponential, Weibull, gamma, loglogistic, lognormal, generalised gamma*.
2. Spline-based: Royston-Parmar (up to 5 knots) and generalised additive models.
3. Fractional polynomials: first-order FP(1) and second-order FP(2) models.
4. Dynamic survival models: local trend and damped trend. Model parameters vary over time; this variation is modelled as a time-series.

In general, the choice of extrapolating model was based on Akaike Information Criterion.

Figure 1: Truth and simulated data



Estimand and performance measure:

The log of extrapolated hazards was the estimand. Mean-squared error (MSE) was the performance measure, this penalises for bias and variability of estimates.

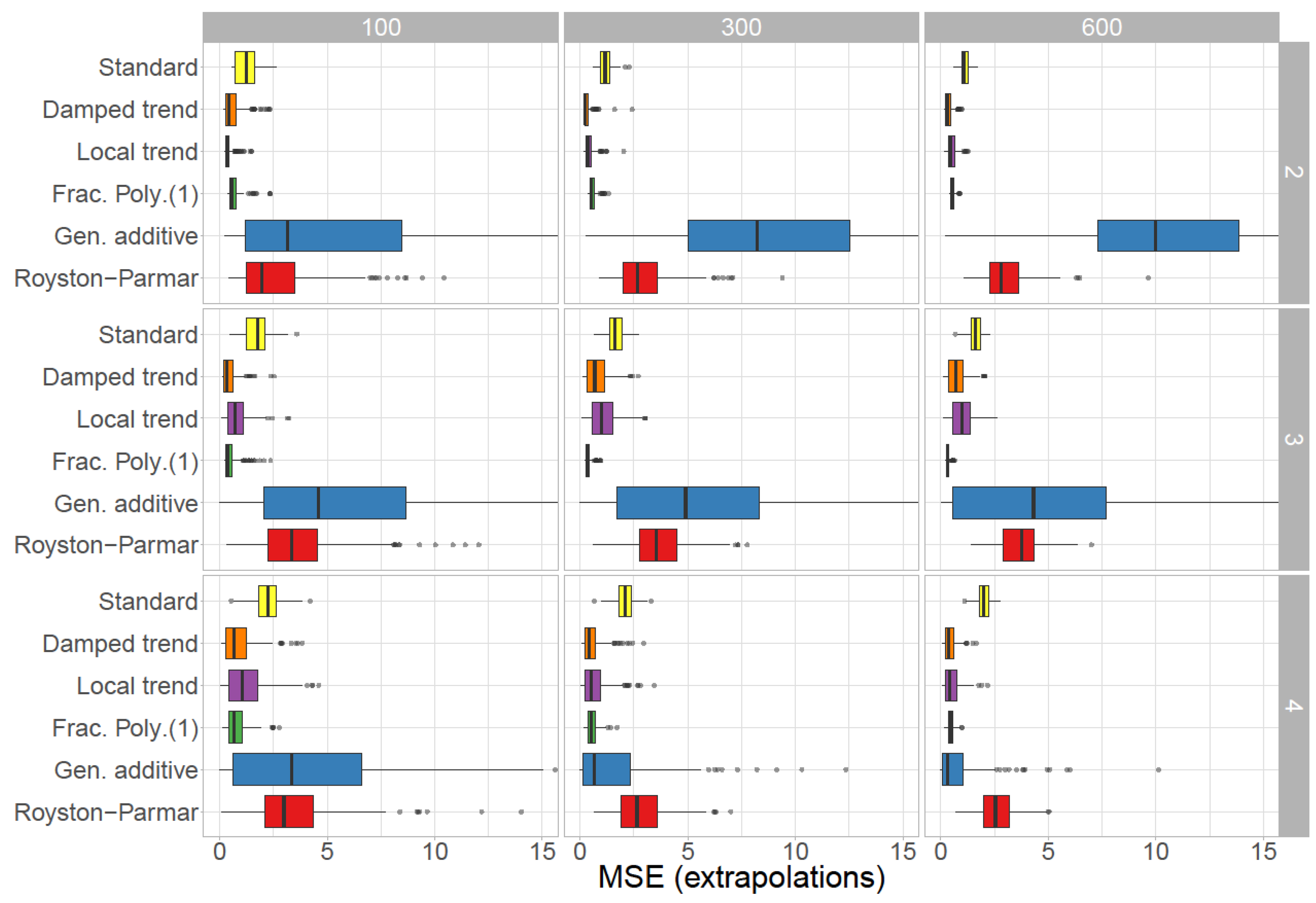
Results and Implications

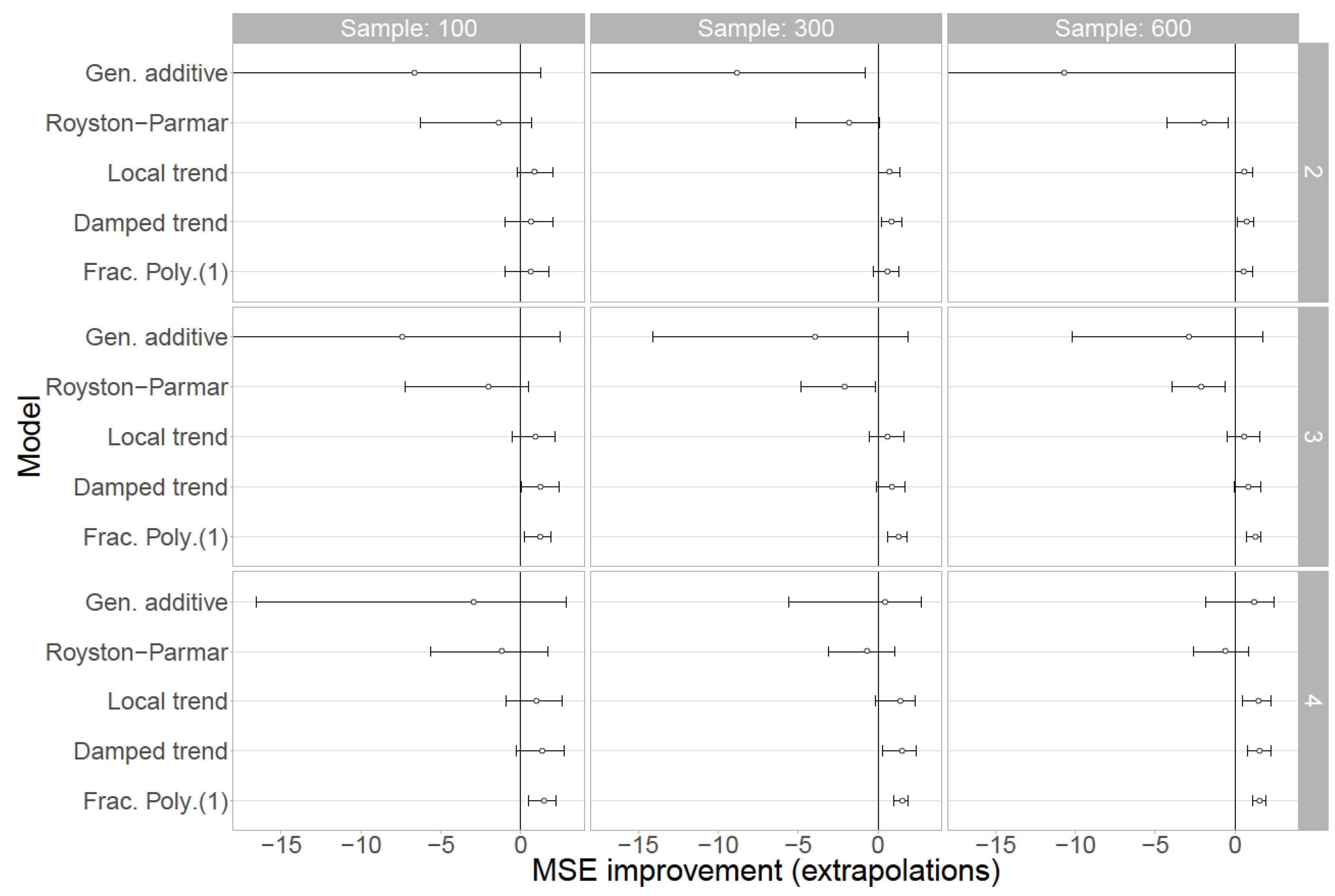
In general, the flexible models provided improved within-sample estimates of the observed data, with the exception of the FP(1) for which estimates had a higher MSE than standard practice. The extrapolation MSE is displayed in Figures 2 (absolute values) and 3 (improvement vs standard practice). Results were very poor for FP(2) models, so are not displayed. The two spline-based models generally had higher MSE, whilst dynamic models and FP(1) gave extrapolations with lower MSE.
In a scenario analysis, the best-extrapolating standard practice model was used as a benchmark. There were no significant differences in MSE for the dynamic models despite the advantage provided to standard methods by using more data. Future work could explore different models and data-generating mechanisms, incorporating external evidence, and compare performance using case-studies. See poster PCN444 for an example case-study.

Dynamic survival models had superior extrapolation performance to standard models. Performance was similar to if the best-extrapolating standard model was known in advance.

Figure 2 (left): Extrapolation MSE by model.

Figure 3 (right): Improvement in extrapolation MSE compared with standard practice.





References: [1] https://doi.org/10.1002/sim.8086 **[2] h**ttps://doi.org/10.1080/00949655.2013.845890 **[3]** https://doi.org/10.1177/0272989X19873661

Acknowledgements and funding: This research was funded by the NIHR Doctoral Research Fellowship (DRF-2016-09-119) 'Good practice guidance for the prediction of future outcomes in health technology assessment'. BK was supported in the preparation of this work by the HEOM Theme of the NIHR CLAHRC Yorkshire and Humber. www.clahrc-yh.nihr.ac.uk. The views expressed are those of the authors, and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

Presented at ISPOR Europe 2019 (PNS327)

b.kearns@sheffield.ac.uk www.scharrheds.blogspot.co.uk

@benjaminkearns2
@scharrheds

FUNDED BY NIHR | National Institute for Health Research