NICE National Institute for Health and Care Excellence

University of BRISTOL Guidelines Technical Support Unit

MSR61

Lessons learned from network meta-analysis of survival data with fractional polynomials

Dietz J[1], Downing B[2], Tebbs H[1], Claxton L[1], Welton NJ[2]

[1] National Institute for Health and Care Excellence

[2] Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

Background

Fractional polynomials (FP) have been proposed for network meta-analysis (NMA) of survival data without relying on the proportional hazards assumption (1). FP NMAs require the analyst to make a number of choices. First, one must choose what time intervals to use. FP NMAs require individual patient data (IPD) to be aggregated into the number of events and number of people at risk for each of a series of time intervals. However, there is limited guidance on the number and width of time intervals to use when aggregating IPD. Second, one must choose what power to use. First order fractional polynomials can be described with the following formula: $\gamma = \beta_0 + \beta_1 t^p$, where power, p is chosen from: -2, -1, -0.5, 0, 0.5, 1, 2, 3 (1). This research aims to make recommendations for fitting FP models based on lessons learned from fitting a large number of FP models for the recent NICE melanoma guideline update (2).

What we did and why

Using a network of 10 studies on 9 treatments for advanced melanoma, we aggregated reconstructed Kaplan-Meier data (3) for Progression Free Survival (PFS) and Overall Survival (OS) in 9 different ways: 3 or 8 intervals, every 1, 2, 3, 4, 5, or 6 months, or every 1 month for 20 months followed by every 2 months.

We then used the above aggregate data to fit first-order FP models with powers from the set -2, −1, −0.5, 0, 0.5, 1, 2, 3, estimated in WinBUGS. We compared models using:

- the Deviance Information Criteria (DIC) to compare between models with the same data inputs,
- posterior mean deviance per data point to compare model fit with different data inputs, and
- visual inspection.

Table 1: Convergence for select OS models (✓ indicates convergence, × indicates model did not converge)

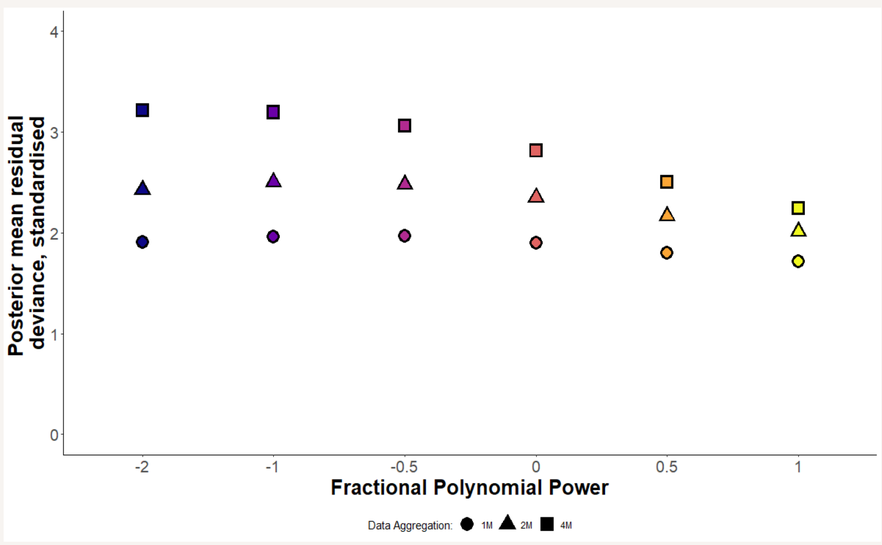
Power	Time interval aggregated on		
	4 monthly	2 monthly	1 monthly
-2	✓	✓	✓
-1	✓	✓	✓
-0.5	✓	✓	✓
0	×	✓	✓
0.5	×	✓	✓
1	×	×	×
2	×	×	×
3	×	×	×

Outcomes and impact

We were able to fit most FP models, however, convergence was an issue for FP models with data aggregated over wider time intervals. For example, where data were aggregated in 3 intervals of 0-6 months, 6-12 months and 12-120 months, the models converged for powers (-2, -0.5, 0.5) for PFS and for powers (-2, -1, -0.5) for OS. Convergence was better when the data were aggregated in 1-month intervals, with models with powers (-2, -1, -0.5, 0, 0.5) converging for both PFS and OS.

Convergence was also more problematic for FP models with higher powers, specifically powers 1, 2 and 3. This finding was true for both PFS and OS.

Figure 1 (Below): Posterior mean residual deviance, standardised by the size of the dataset for select OS models. Plotted for different data aggregation intervals.



We found that models aggregated on finer intervals have smaller deviance per data point (standardised deviance, Figure 1) indicating better average model fit. Furthermore, NMA estimates gave a better visual fit to KM data (Figures 2 & 3) when aggregating the data using the finest time interval (every 1 month). However, the degree to which fit improves varies depending on whether or not one is looking at PFS or OS and depending on what treatment one is looking at. Run-time was increased with a finer aggregation; however, this was not prohibitively long. We found that the DIC and visual inspection were in agreement as to which power is the best fitting model.

Figure 3 (Right): Predicted OS curves for Encorafenib + Binimetinib for select FP Models with different time intervals

What we learnt

We recommend when fitting first-order FP models:

- That data is aggregated in fine time intervals
- It is essential to check convergence. Further sense checks of model fit statistics and visual inspection of the survival predictions should be made to check the model is estimable. In our example, although the Gelman-Rubin-Brooks plots indicated certain powers converged (often higher powers, specifically 3), the DIC values for these models were incredibly large and the plotted survival predictions were implausible indicating numerical issues in the model estimation.
- DIC together with visual inspection can be used to select models for inference and economic models.

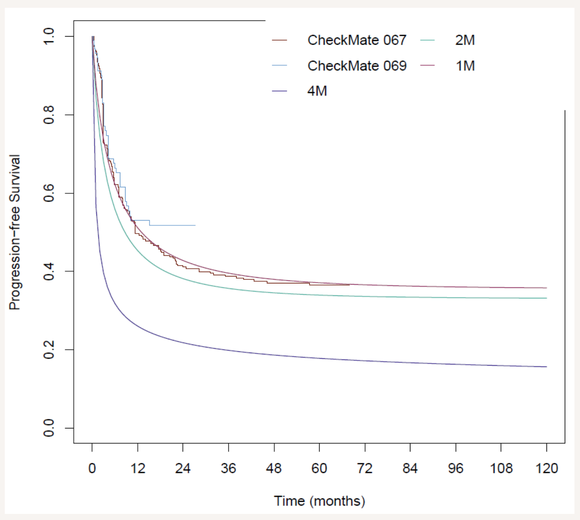
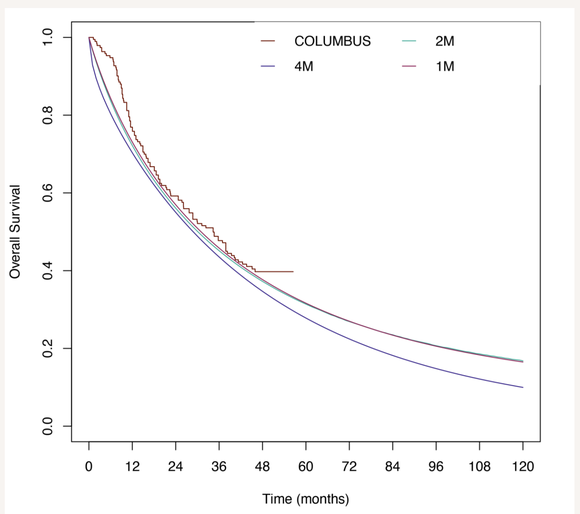


Figure 2 (Above): Predicted PFS curves for Nivolumab + Ipilimumab for select FP Models with different time intervals



References

1. Jansen JP. Network meta-analysis of survival data with fractional polynomials. BMC medical research methodology. 2011 Dec;11(1):1-4.
2. National Institute for Health and Care Excellence (2022) Melanoma: assessment and management, NG14
3. Guyot P, Ades AE, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. BMC medical research methodology. 2012 Dec;12(1):1-3.

Jeremy Dietz
Senior Technical Analyst - Economics
National Institute for Health and Care Excellence
Email: Jeremy.Dietz@nice.org.uk
Telephone: 020 7045 2146

nice.org.uk