Use of generalised fluctuation test to optimise piecewise fitting in immuno-oncology modelling

Murvin Jootun[1], David C. Bode[1], Ion Agirrezabal[1]
[1]Covance Market Access, London, United Kingdom

Introduction

- In oncology health technology assessments (HTAs), survival outcomes have traditionally been extrapolated beyond the duration of clinical trials using parametric survival models. However, it is now common for more flexible regression models to be used, with support from the HTA agency, the National Institute for Health and Care Excellence (NICE).[1,2]
- Piecewise modelling is suitable when standard parametric models produce implausible long-term extrapolation due to poor fit of clinical trial data, a challenge commonly observed in immuno-oncology therapy analyses.[1,2]
- Piecewise models split the duration of follow-up into mutually exclusive intervals and fit separate parametric models for each interval.[3,4] The time points at which parametric models change are known as the inflection points or breakpoints, where the hazard function is assumed to vary.[3]
- The choice of cut-off points (where separate models would be fitted before and after) can have a significant impact on PFS and OS extrapolation estimates; but there are no clear guidelines describing how these should be identified, and published economic models rarely report the rationale behind the choices made.

Objective

- The objective of the analysis was to identify the evolution of hazard across survival curves, and to identify optimal time points where changes in hazard occurs. This will facilitate the selection of breakpoints for piecewise fitting and selection of knots for spline models.

Methods

Study selection/ data extraction

- A targeted literature review was conducted to identify recently published data on IO therapies in any indication. The selection criteria included randomised clinical trials (RCTs) with long-term OS follow-up data (>30 months).
- The OS Kaplan-Meier (KM) curves available in the identified publications were digitised and individual patient-level data were reconstructed as described by Guyot *et al.*[5]

Statistical analysis

Overview

- Time series often have nonstationary properties: changing levels or variable values, alterations in variances and autocorrelations, or a combination for some if not all those attributes.[4]
- The gold standard methodology for the analysis of structural changes in univariate time series data relies on the functional central limit theorems), which are based on the least square method.[6]

The model

- A linear signal-plus-noise model for a univariate quantity y_i is described by:

$$y_i = \mu_i + e_i;\ i = 1,...,T$$

Where μ_i is the deterministic signal and e_i is the noise with $E[e_i]=0$ and $Var[e_i]=\sigma^2$ (i.e. e_i is assumed to be independent and normally distributed).

- The null hypothesis H_o: $\mu_i = \mu_o$ assumes no changes in signal (or structure) and considers the natural estimates of μ_o to be the following recursive estimates:

$$\hat{\mu}_k = k^{-1} \sum_{i=1}^{k} yi,\ k=1,...,T$$

with the residuals $\hat{\mu}$ given as:

$$\tilde{e}_i = y_i - \hat{\mu}_{i-1},\ i=2,...,T$$

- The main approach for detecting structural changes is to study the fluctuations of partial (or cumulative) sums (CUSUMs) of these recursive residuals. Test statistics are used to reject the null hypothesis of parameter stability if their fluctuations are excessive.
- This methodology is extended to detect the breakpoints at different time points in the data.

Analyses

- Survival estimates from KM curves were transformed into log-cumulative hazard and ordinary least squares (OLS) regression were fitted.
- Fluctuation tests and significant testing:
  - The generalised fluctuation tests analysed the fluctuation in the residuals resulting from the OLS model.
  - Three types of empirical fluctuation processes were tested (Rec-CUSUM, OLS-CUSUM and OLS moving sum [MOSUM] processes), each of which has its own unique underlying assumptions on the behaviour and movement of the residuals.
  - Significant testing at an alpha level of 0.05 was applied to reject the null hypothesis.
- Identification of breakpoints:
  - The optimal number of breakpoints was determined via the Bayesian Information Criterion (BIC) and residual sum of squares (RSS).
- All analyses were conducted using R version 3.5.3.

Results

- Ten KM curves with a minimum follow-up period of 30.6 months were analysed. The indications covered were melanoma, non-small-cell lung cancer (NSCLC) and renal cell carcinoma (RCC). A summary of the studies and findings is presented in Table 1.
- Ordinary least square regression provided good fit to the data for all studies ($R^2 > 0.95$).
- The fluctuation (OLS-CUSUM) and the significance tests showed evidence of more than two breakpoints in the data, indicating different hazards may exist at different time points. The BIC and RSS favoured four breakpoints.
- Table 1 displays the time of the four breakpoints, showing that the timing of breakpoints is comparable across the studies with the average time (standard error) for the first, second, third and fourth breakpoints occurring at 3.15 (0.31), 6.63 (0.65), 11.84 (0.90), 20.96 (0.86) months, respectively.
- The first breakpoint identified in our analysis was consistent with the medical literature where patients are known to experience a delay in clinical benefit at the beginning treatment phase. Also, when the fourth breakpoint was compared to the reported KM curves across the studies, it was observed that at those particular time points patients experience durable survival benefit.[1]
- The survival of patients at breakpoints 1 and 2 was comparable while the proportion of patients surviving at breakpoints 3 and 4 differs across studies, most likely due to different treatment effects of IO therapies as demonstrated by Hodi et al., 2018.[7]
- Rec-CUSUM and OLS-MOSUM fluctuation processes were also tested and they were in accordance with the results from the OLS-CUSUM fluctuation test (results not shown).

Table 1: Summary of clinical studies and results of analysis

Study	Indication	Intervention arm	Follow-up period (months)	Optimal number of BPs	Time and Survival		
					BPs	Time (months)	Survival probability
Antonia 2018[8]	Stage III, unresectable NSCLC	Durvalumab	40.8	4	BP1: BP2: BP3: BP4:	5.78 11.51 18.02 26.46	0.93 0.83 0.74 0.64
Ascierto 2018[9]	Unresectable or metastatic melanoma	Ipilimumab	46.7	4	BP1: BP2: BP3: BP4:	2.11 3.97 7.37 17.01	0.91 0.80 0.65 0.47
Hodi 2018[7]	Advanced melanoma	Ipilimumab	55.5	4	BP1: BP2: BP3: BP4:	3.77 6.99 12.6 22.63	0.88 0.79 0.64 0.46
Hodi 2018[7]	Advanced melanoma	Nivolumab + ipilimumab	57	4	BP1: BP2: BP3: BP4:	2.79 5.32 10.27 20.20	0.94 0.87 0.77 0.65
Hodi 2018[7]	Advanced melanoma	Nivolumab	54	4	BP1: BP2: BP3: BP4:	2.86 5.53 11.44 21.10	0.94 0.85 0.75 0.61
Mok 2019[10]	Locally advanced or metastatic NSCLC	Pembrolizumab	37.2	4	BP1: BP2: BP3: BP4:	2.66 6.42 11.76 21.11	0.85 0.71 0.58 0.44
Rini 2019[11]	Previously untreated metastatic RCC	Atezolizumab + bevacizumab	38	4	BP1: BP2: BP3: BP4:	3.25 8.27 14.74 23.99	0.93 0.85 0.75 0.63
Schadendorf 2015[12]	Unresectable or metastatic Melanoma	Ipilimumab	119	4	BP1: BP2: BP3: BP4:	2.43 5.24 9.27 17.69	0.87 0.72 0.54 0.34
Socinski 2018[13]	Metastatic non-squamous NSCLC	Atezolizumab	30.6	4	BP1: BP2: BP3: BP4:	3.22 7.92 12.84 18.64	0.91 0.78 0.65 0.52
Vokes 2018[14]	Previously treated advanced NSCLC	Nivolumab	53	4	BP1: BP2: BP3: BP4:	2.59 5.11 10.08 20.81	0.85 0.70 0.52 0.31

Key: BPs, breakpoints; NSCLC, non-small-cell lung cancer; RCC, renal cell carcinoma.

Figure 1: Time of breakpoints observed across IO therapies

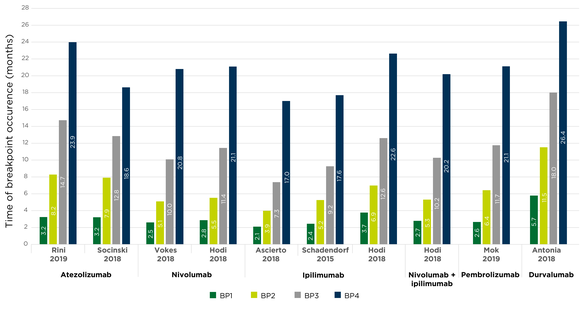
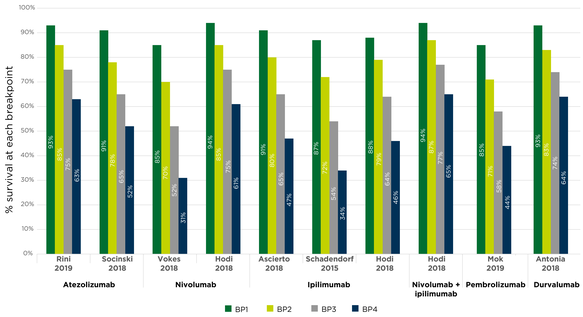


Figure 2: Survival observed at breakpoints across IO therapies



Conclusions

- The use of fluctuation tests captured relevant breakpoints, indicating that hazard changes over time, which is in line with the medical and health economics literature for IO therapies.[15]
- The results suggested a pattern may exist in survival estimates of patients receiving IO therapies: the optimal number of breakpoints was four across RCTs (approximately at 3, 7, 12 and 21 months after treatment initiation).
- This methodology provides a valuable framework for the identification of inflection points, as cut-off point selection is often done arbitrarily despite its potentially significant impact on long-term survival projections.
- However, the lack of patient-level data with associated covariates limited our ability to analyse whether patterns in breakpoints vary according to patients' characteristics.
- Further studies are needed to validate these results clinically and understand whether these patterns are somehow correlated with the biological evolution of the disease.

References

1) Deniz, B. et al. PloS one 2018 13, e0203406. **2)** Latimer, N. R. Medical decision making : an international journal of the Society for Medical Decision Making 2013 33, 743-754. **3)** Austin, P. C. International statistical review/ Revue internationale de statistique 2017 85, 185-203. **4)** Kleiber, C. Complexity and Synergetics 2017, 275–286. **5)** Guyot, P. et al. BMC medical research methodology 2012 12, 9. **6)** Bai, J. & Perron, P. Econometrica 1998 66, 47-78. **7)** Hodi, F. S. et al. The Lancet. Oncology 2018 19, 1480-1492. **8)** Antonia, S. J. et al. The New England journal of medicine 2017 377, 1919-1929. **9)** Ascierto, P. A. et al. The Lancet. Oncology 2017 18, 611-622. **10)** Mok, T. S. K. et al. Lancet (London, England) 2019 393, 1819-1830. **11)** Rini, B. I. et al. Lancet (London, England) 2019 393, 2404-2415. **12)** Schadendorf, D. et al. 2015 33, 1889-1894. **13)** Socinski, M. A. et al. The New England journal of medicine 2018 378, 2288-2301. **14)** Vokes, E. E. et al. Annals of oncology : official journal of the European Society for Medical Oncology 2018 29, 959-965. **15)** Ouwens, M. et al. PharmacoEconomics 2019 37, 1129-1138.