Leveraging response-based survival observations from a historical trial to extrapolate from immature data in a study of relapsed/refractory multiple myeloma

MSR99

View all Parexel's posters at ISPOR Europe 2024

Daniel J. Sharpe[1], Ashley E. Tate[2], Kateryna Chepynoga[3], Tuli De[4], Jackie Vanderpuye-Orgle[4]

[1]Parexel, London, UK
[2]Parexel, Amsterdam, Netherlands
[3]Parexel, Copenhagen, Denmark
[4]Parexel, Durham, North Carolina, USA

Background

- Hematologic response has substantial impact on progression-free survival (PFS) in patients with relapsed/refractory multiple myeloma (RRMM). Patients who achieve complete response or better (≥CR) exhibit markedly improved survival outcomes vs patients whose response is less than complete (<CR)[1,2]
  - RRMM is a highly heterogeneous disease, with numerous key prognostic factors including cytogenetic risk and prior treatment history, and there are many indicated combination regimens with no single standard of care[3]
  - RRMM is a chronic disease, and long-term outcomes are poor[4], precluding the use of mixture cure models to analyze underlying survival heterogeneity and perform extrapolations
- We explored using Bayesian parametric mixture models[5] (B-PMMs), informed by mature data from a historical study of a related therapy that were stratified by hematological response (≥CR vs <CR), to extrapolate PFS outcomes in the experimental arm from a reconstructed early data cut of the CASTOR study[6]
  - the initial data cut of the phase III CASTOR study, investigating daratumumab, bortezomib, and dexamethasone (DVd) in RRMM, had 7.4 months minimum follow-up
    - these data were still highly immature, with PFS events observed for around 25% of patients
  - PFS projections from the B-PMM were compared to observations from a later data cut of CASTOR with 40.0 months median follow-up[9]
  - reconstructed PFS data for carfilzomib, lenalidomide, and dexamethasone (KRd) from the phase III ASPIRE study, with 32.3 months median follow-up[7,8], were used as prior information in the B-PMM to aid response-based survival extrapolations in CASTOR, after adjusting for between-trial confounding effects using the control arms of the two studies
    - while ASPIRE and CASTOR are well-matched on some key baseline characteristics (e.g., age and number of prior lines of therapy), there are differences in the distribution of other relevant biomarkers, such as staging

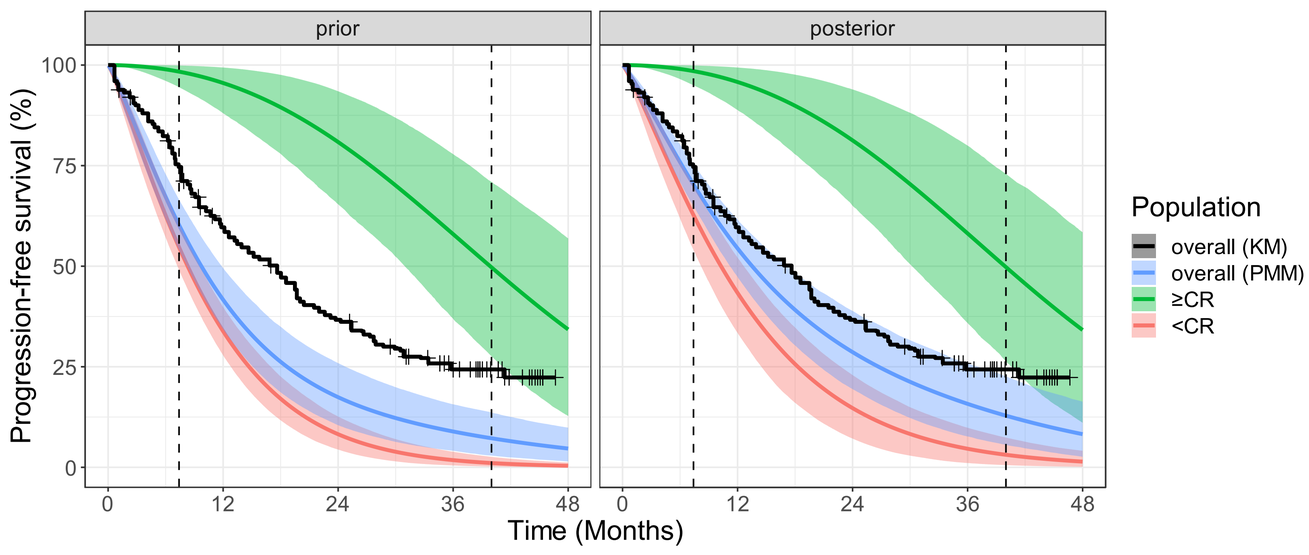
Methods

- Mature PFS data for the KRd arm from ASPIRE were used to derive prior Weibull distributions for latent ≥CR and <CR survival functions for the DVd arm of CASTOR
  - K and V are both proteasome inhibitors, while R and D are both immunotherapies, and thus there is a relation between KRd and DVd combinations that qualifies KRd as a reasonable, albeit non-ideal, source of *a priori* information on outcomes with DVd
  - however, KRd and DVd have different mechanisms of action, Kd is superior to Vd[10], and the efficacy of constituent therapies is liable to depend on synergistic effects[11]
- To broadly account for between-trial differences, the prior distribution for the latent <CR survival function was adjusted by a hazard ratio calculated from the respective control arms (namely, Rd and Vd; both of which are a targeted therapy plus dexamethasone combination) of the two trials (estimate: 0.36 ASPIRE vs CASTOR)
  - residual differences in patient populations across the two studies after this crude adjustment are essentially tolerated to a degree expressed by the variances of the prior distributions
- The fraction of patients classed as ≥CR at 6 months in the KRd arm of ASPIRE informed a prior beta distribution for the proportion of ≥CR patients in the B-PMM for the DVd arm of CASTOR
  - the variance of this prior was raised relative to the statistical uncertainty in ASPIRE (i.e., shape parameters were made to sum to one-tenth of the number of patients) to reflect confounding
- Effectively, the core assumptions of the B-PMM are as follows:
  1) ≥CR patients in both trials exhibit a somewhat similar survival pattern
  2) beyond the follow-up period of CASTOR, <CR patients in CASTOR have similar hazards to <CR patients in ASPIRE, after adjustment of the latter to reflect additional hazard in CASTOR
  3) the proportion of patients latently classed as ≥CR is somewhat similar between the two trials
- Model estimates were reported as posterior means [with 95% credible intervals]

Results and Discussion

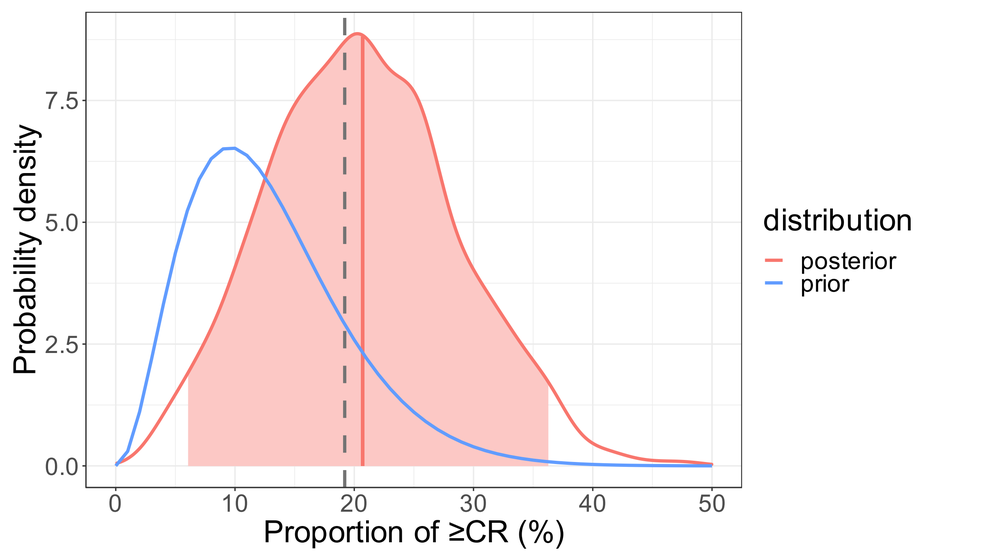
- PFS extrapolations for the overall DVd population, from the initial data cut of CASTOR, were consistent with later observations, albeit estimates were conservative (Fig. 1). e.g., 36-month PFS: 15.8% [7.7-25.8%] B-PMM vs 24% observed[9]
  - uncertainty levels in the survival extrapolations for the different populations were modest and realistic, reflecting the "true" uncertainty inclusive of prior knowledge from ASPIRE
- The estimate for the proportion of ≥CR patients (20.7% [6.2-36.5%]) in the DVd arm of CASTOR closely agreed with the CR rate reported in the initial data cut (19.2%), even though the prior expectation was conservative (12.8% [3.0-21.8%]) (Fig. 2)
  - this result demonstrates the capability of the B-PMM to deviate from the prior expectation and recover accurate estimates when the current evidence supports such behavior
- Patients treated with DVd in CASTOR that were latently classified as <CR in the B-PMM were predicted to have very high probability of progression or death on longer timescales (e.g., 36-month PFS: 4.6% [1.1-9.8%])
  - this prediction appears clinically reasonable, since selected subgroups in the DVd arm with poor prognosis at baseline had observed 36-month PFS <10%[9]
- Patients treated with DVd in CASTOR that were latently classified as ≥CR exhibited greatly improved survival, but with remaining moderate hazards, in accordance with observations from the KRd arm in ASPIRE (e.g., 36-month PFS: 58.1% [35.9-80.0%])
  - this prediction appears reasonable but potentially conservative upon comparison with PFS rates among the 14% of patients who were assessed as MRD-negative, an alternative criteria for strong response[3], in the DVd arm of CASTOR, for whom 36-month PFS was approx. 72%[9]
- There is some immortal time bias in the definition of the ≥CR population, since the corresponding prior survival function is based on observations at a 6-month landmark
  - this aspect can be rationalized as hematological responses being strongly correlated with favorable baseline characteristics that result in negligible hazards at early times
- While reasonable survival projections were transparently obtained from the B-PMMs, greater accuracy could potentially have been achieved by modeling key baseline subgroups, including subgroup-specific prior information

Figure 1: Prior and posterior B-PMM survival functions for the overall population and latent ≥CR and <CR subpopulations in the DVd arm of CASTOR, compared to Kaplan-Meier (KM) estimates from more mature trial data.



Vertical lines indicate median follow-up duration in the initial (to which the B-PMMs were fitted) and later (for which the trial observations are shown) data cuts of CASTOR. Shaded areas indicate 95% credible intervals. Crosses indicate censoring.

Figure 2: Prior and posterior estimates for the proportion of ≥CR patients in the DVd arm of CASTOR.



Vertical dashed line indicates the observed CR rate from the initial data cut. Vertical solid line indicates the posterior mean estimate. Shaded areas indicate 95% credible intervals.

Conclusions

- B-PMMs can reliably extrapolate complex survival trends borne by response-based survival heterogeneity in RRMM, even from the earliest data cuts when data remain highly immature
  - the models require suitable historical trial data that are stratified by hematological response and feature extended follow-up
- Despite between-trial confounding manifesting as less favorable outcomes in CASTOR vs ASPIRE, the historical PFS data for the experimental arm of ASPIRE proved useful in aiding reliable extrapolation of outcomes in the experimental arm of CASTOR, after a straightforward adjustment of selected priors
  - this study supports the use of data for a related, and not necessarily identical, experimental therapy from a historical study as prior information in an appropriately designed B-PMM or other Bayesian survival model

REFERENCES

[1] Daniele P, et al. *Future Oncol* 2023; 19(6): 463-471.
[2] Sonneveld P and Broijl A. *Haematologica*. 2016; 101(4): 396-406.
[3] Bhatt P, Kloock C, and Comenzo R. *Curr. Oncol*. 2023; 30(2): 2322-2347.
[4] Bazarbachi AH, et al. *Leukemia* 2019; 33: 2343-2357.
[5] Sharpe DJ. *Value Health* 2023; 26(12): S414-S415.
[6] Palumbo A, et al. *N Engl J Med* 2016; 375: 754-766.
[7] Stewart AK, et al. *N Engl. J. Med*. 2015; 372(2): 142-152.
[8] Dimopoulos M, et al. *J. Hematol Oncol.* 2018; 11: 49.
[9] Mateos M.-V., et al. *Clin. Lymphoma Myeloma Leuk.* 2019; 20(8): 509-518.
[10] Dimopoulos M, et al. *Lancet Oncol*. 2016; 17(1): 27-38.
[11] Yong KL, et al. *Haematologica.* 2021; 106(10): 2694-2706.

Contact: Daniel.Sharpe@parexel.com

www.parexel.com