Author’s response to reviews

Title: Integrating expert opinion with clinical trial data to extrapolate long-term survival: A case study of CAR-T therapy for children and young adults with relapsed or refractory acute lymphoblastic leukemia

Authors:
Shannon Cope (shannon.cope@precisionxtract.com)
Dieter Ayers (dieter.ayers@precisionxtract.com)
Jie Zhang (jie.zhang@novartis.com)
Katharine Batt (katharine.batt@gmail.com)
Jeroen Jansen (jeroen.jansen@precisionxtract.com)

Version: 2 Date: 22 May 2019

Author’s response to reviews:

Dear Editor,

Thank you very much for reviewing our manuscript, ‘Integrating expert opinion with clinical trial data to extrapolate long-term survival: A case study of CAR-T therapy for children and young adults with relapsed or refractory acute lymphoblastic leukemia’ (BMRM-D-17-00515R1). We also greatly appreciate the reviewers for their thoughtful comments and suggestions. We have revised the manuscript based on their feedback and included point-by-point responses to their comments below. We hope that you find our responses satisfactory and that the manuscript is now acceptable for publication in BMC Medical Research Methodology.

Sincerely,

Shannon Cope
Reviewer reports:

Paul Gaynon (Reviewer 1):

Comment: I am a clinician with some experience in clinical trials and trial design. I am uncertain of your audience.

Response: The audience for this paper is intended to be healthcare researchers involved in developing and/or applying methodologies for evidence synthesis evaluating the long-term efficacy of new interventions. The audience may include regulators or decision makers from health technology assessment agencies, pharmaceutical companies or trialists with only short-term data. Although clinicians are not our main target audience, physicians involved in this field or those involved in evaluating new interventions based on early data from clinical trials may also be interested in the process and results of this exercise.

Comment: With regard to Kymriah, only a fraction of eligible patients undergo pheresis, only a fraction of harvested patients have a useful product, and only a fraction of patient with a useful product are actually infused. You start your analysis from infusion. I believe that attrition runs about one third. Thus the subset of patients with a successful infusion differs from the subset of patients enrolled on trial in an uncertain manner.

Response: This is a valid concern, and it is necessary to ensure that the different periods of treatment are accounted for. Similarly it is necessary to ensure that patient drop out is accurately described in any analyses of patients receiving this type of treatment. The aim of the current study was to assess the long term survival of patients who had received treatment. The elicitation process focused exclusively on the subset of patients who had received a successful infusion. Since experts were not asked to generalize the results to broader population, potential drop out should not bias the current results. Nonetheless, this is an interesting point, which could lead to a validation exercise, where experts could be provided with data from the complete set of patients, and they could estimate the survival from that cohort. This could be compared to an estimate composed of estimates of the infusion rate and the survival of infused patients. We have now highlighted this point in the discussion to ensure that the interpretation of these results accounts for this factor.
Comment: If you plan to inform bedevilled clinicians such as myself, you will do well to provide further inform about the implications of your various cure models - Weibull and Gompertz.

This is a critical issue as we depend too strongly on the logrank statistic - a measure of time to event and not the eventual cure rate.

Current treatments in ALL have both delayed and decreased relapse.

Response: In our study we present a new approach to synthesize expert information with information from clinical trials. The specific form of the model is not the central point, which is why we have highlighted in the discussion that this method to integrate expert-elicited estimates can be applied to many different kinds of survival models.

Comment: I understand, I think, your careful refinement of your expert's thoughtful opinions but I see no empiric validation of their consensus.

Are your experts any more accurate than kindergartners or kangaroos?

You offer no validation of your quite reasonable method in other circumstances where such validation may exist.

Response: The criteria for selecting experts is outlined in the methodology section. Since this treatment is relatively new, there is not a large pool of potential experts from which to draw. The experts that were selected represented a group of practitioners with experience in this area. Therefore, the experts were well positioned to provide accurate estimates of survival. During the peer review process for this manuscript, additional follow-up data from ELIANA became available. As such, we have revised the manuscript to compare these results to those including the expert information. It appears that the expert estimates were accurate at 24 months. In future, it will be beneficial to further validate the expert estimates based on longer follow-up from this trial, which we have highlighted in the discussion.

Comment: Figure 3 is too busy - too many curves. Too hard to distinguish the Gompertz curve.

Response: We agree that the plot is busy, though that is of primary importance. However, we feel this plot is important to demonstrate the variability in extrapolations that result from the use of alternative models. This helps to illustrate the structural uncertainty, which provides the motivation for integrating expert information.
The manuscript entitled "Integrating expert opinion with clinical trial data to extrapolate long-term survival: A case study of CAR-T therapy for children and young adults with relapsed or refractory acute lymphoblastic leukemia" aims to assess the feasibility of formally integrating long-term survival estimates from experts with empirical clinical trial data to provide more credible extrapolated survival curves. To this effect, the authors apply a case study regarding relapsed or refractory B-cell pediatric and young adult acute lymphoblastic leukemia (r/r pALL) regarding long-term survival for CTL019 (chimeric antigen receptor T-cell) from which evidence from the phase II ELIANA trial is available up to 1.5 years. Seven pediatric oncologists and hematologists experienced with CAR-T therapies were recruited in a double-blinded manner. Survival rates and related uncertainty at 2, 3, 4, and 5 years were elicited from experts using a web-based application adapted from SHeffield ELicitation Framework. The experts' results were combined with the ELIANA trial data using time-to-event parametric models in a Bayesian framework that accounted for experts' uncertainty in their estimates, producing an overall distribution of survival over time. The authors found that extrapolated survival curves based on ELIANA trial data without expert information were uncertain, differing based on the model choice. Survival estimates between 2 to 5 years also varied between individual experts. However, incorporating their estimates improved the precision in the extrapolated survival curves. The authors conclude that by the proposed methodology expert opinion can be elicited and synthesized with observed survival data using a transparent and formal procedure. The manuscript and proposed approach is of interest and seeks to integrate empirical and experimental data.

Comment: However, further information is necessary regarding the selection of experts. In particular, selection and recruitment procedures ought to be described in further detail, and a justification of the number of experts included should be provided.

Response: The expert selection procedure is outlined in the methodology section. The text in this section has been modified to provide more detail, including a justification for the number of experts, which was based on the SHeffield ELicitation Framework recommendations. If necessary, we can also provide the Participant Information Statement used for recruitment as a supplemental file. Finally, we have added a point to the discussion to highlight that the exercise could include additional experts in future.
Comment: Further information regarding discrepancies and differences in expert opinions should be presented.

Response: Figure 4 visually illustrates all of the estimates elicited from the individual experts in order to be transparent regarding the results. Overall, the experts were fairly consistent, though additional text has been added to the results section to describe the observed differences. These differences were not substantial enough to affect the estimation of long-term survival, and the estimate of survival based on these elicitations was agreed upon in the consensus meeting.

Comment: Finally, the proposed methodology ought to be validated in the experimental ELIANA cohort following adequate follow-up, and further replicated in an independent trial.

Response: We agree with this assessment, and have revised the paper to include the results from ELIANA that were published during the peer review process for this manuscript. The discussion has been revised substantially to explain reasons for observed differences between our results and the longer follow-up from ELIANA. Additionally, we propose some possible improvements on the elicitation process that require additional research. We are in the process of replicating the study for another oncology trial and hope to publish those results as separate manuscript.