Creating historical controls using data from a previous line of treatment – two non-standard approaches

Anthony James Hatswell, William George Sullivan

Corresponding author:

Anthony Hatswell

Department of Statistical Science
University College London
Gower Street
London
WC1E 6BT

07759 243157
ahatswell@deltahat.co.uk

Abstract

Where medical interventions are licensed based on only uncontrolled study data (for example a single-arm trial), a common approach for estimating the incremental benefit is to compare the treatment to a 'historical control'; data collected from patients who did not receive the intervention. We illustrate with motivating examples two methods for the creation of historical controls where disease progression and overall survival are typically the key clinically meaningful endpoints. The first method utilises information routinely collected in a clinical trial programme: patients' time to disease progression on their previous line of treatment against which outcomes can be compared. The second uses published clinical outcomes for the prior line of treatment which can be extrapolated to estimate outcomes at the next line. As examples we use two pharmaceuticals licensed on the basis of uncontrolled clinical studies - idelalisib for double-refractory follicular lymphoma and ofatumumab for double-refractory chronic lymphocytic leukemia. Although subject to limitations that should be considered on a case-by-case basis, the methods may be appropriate when trying to quantify the clinical benefit of treatment based on limited and uncontrolled trial data. As a result, the methods can be used to inform health technology adoption decisions.

1. Introduction

Novel medical interventions, especially pharmaceuticals, are usually studied in randomised controlled trials of the intervention against the most appropriate comparator. Where this is not the case, indirect evidence must be used to estimate the relative effectiveness of treatments. Methods for estimating comparative efficacy include Bucher indirect comparisons (1) and network meta-analysis (2), such methods however require a network of evidence with a common comparator. Where such a comparator is not available, comparisons with trials in similar populations that did not receive the intervention have been suggested, termed 'historical controls' (3).

Whilst historical controls have limitations regarding their applicability, they allow comparisons to be made where otherwise they could not be performed. Recent work in the area has involved the development of matching methods to reduce the bias in comparisons due to difference in patient populations (4,5), and the creation of controls from 'big data' where there exists no published information on an untreated cohort of patients (6). It is in this latter area where our work sits; the creation of controls where there are no published examples available.

The objective of this work was to present methods through which estimates of comparative effectiveness could be made in the absence of randomised controlled trial or historical comparator data from an external data source in a similar patient group and examined line of treatment. Our interest is primarily in the point estimates and uncertainty surrounding survival, though the methods would allow the estimation of other estimands (such as hazard ratios or p-values).

We highlight two practical approaches to the estimation of historical data to be used as controls in the estimation of comparative efficacy. The first approach utilises data taken from the pivotal regulatory trial patients' previous line of treatment and is particularly useful in conditions where patients experience multiple lines of treatment with disease management as the main objective, and when current care is well represented by available care for prior treatment. The second approach uses published literature on the previous line of treatment, in particular post-progression survival

(PPS) data, to estimate expected survival at the current treatment line. This approach is particularly useful when the type of current care is well represented by what patients would have received upon last disease progression i.e. where no new intervention has become available, as is often the case in end-stage cancers.

Motivating examples are presented for both methods to demonstrate the approaches and to evaluate the merit of the results produced. A companion R file is also made available to illustrate how the approaches can be applied using simulated data.

2. Historical control as outcomes from within-study previous treatment

2.1. General approach

In many cases treatments are used to manage a disease, with well-managed disease likely to lead to reduced risk of mortality, improved quality of life, or ideally both. Typically however, patients will eventually relapse and may benefit from further treatment – often based on a different mechanism of action. Such diseases include autoimmune diseases such as rheumatoid arthritis, human immunodeficiency virus, and as well as various cancers.

Where this pattern of treatments is followed through treatment lines, comparison with data for the pivotal trial patients' previous line of treatment may be informative. Assuming such treatments have been selected as they are believed to be efficacious, should the novel treatment provide disease management for at least as long, this may be a strong indicator of efficacy. In diseases where relapses are typically shorter with each disease progression (perhaps owing to accumulated toxicity from consecutive treatments, or development of more treatment-resistant disease) achieving an equal disease management period may even be a higher bar for indicating efficacy.

To estimate the benefit of the new treatment, data on the pivotal trial patients' previous treatment can be harnessed, if available. The difference between the projected disease management period with the novel intervention and the disease management period on the previous line of treatment

(used as a historical control) can then inform estimates of the benefit of the treatment. We illustrate this with reference to a motivating example in Sections 2.2 and 2.3.

2.2. Motivating example

Follicular lymphoma (FL) is a cancer of the lymphatic system where white blood cells multiply unchecked and collect in the lymph nodes; if untreated, patients are at risk of infection, abnormal blood counts, and have an elevated risk of death (7). Although the disease is treatable, patients will eventually become refractory to treatment, with each line of therapy typically achieving a shorter remission than the last (8).

Idelalisib (Zydelig®, Gilead) was licensed by the European Medicines Agency (EMA) and Food and Drug Administration (FDA) in 2014 for the treatment of double-refractory FL on the basis of Study 101-09 (9). The study enrolled 125 patients, 72 of whom had FL and on whom the license was based. These FL patients had an overall response rate of 54% and demonstrated median progression-free survival (PFS) of 11 months (95% Confidence Interval, CI: 8.1 – 13.6). It did not contain a control arm, and no external historical data was available on untreated patients to estimate comparative effectiveness.

2.3. Application of the approach to the motivating example

In this Section, we describe implementation of the general approach to estimate the absolute and relative survival associated with idelalisib, using Study 101-09 data.

To estimate the counterfactual outcomes, i.e. what would have happened without the availability of idelalisib, disease-progression data on the previous line of treatment were used, as collected in Study 101-09 (9). These data were then compared to time-to-progression (TTP) data for idelalisib in Study 101-09 (10). TTP is used in preference to PFS as a better definition of what was collected at the prior line – any patients who died pre-progression would never have entered Study 101-09 to receive idelalisib. Two assumptions were implicit in this approach – firstly that the risk of death

when pre-progression is the same across treatment lines (and negligible). Although a strong assumption in some applications, this was stated to be the case for this indication by the manufacturer (10). A second assumption was that the benefit of treatment is in delaying disease progression, with PPS equal between idelalisib and standard care – an assumption deemed appropriate in the context of FL where disease management is the primary endpoint of interest.

As patient level data (PLD) are not generally available to independent researchers, data were synthesized from publications (a company performing the analysis would be able to draw directly on the trial data). Data on FL patients treated with idelalisib in study 10-09 are presented in Salles et al. (11), including PFS for treatment with idelalisib and 'PFS' for the prior line of treatment. These data were digitised using the method of Guyot et al. (12). The algorithm requires Kaplan-Meier plots to be digitised using software, with reported censoring and numbers at risk used to then estimate the patient level data on which the plot was based. Where reported, numbers at risk are also used to space censoring and event times when these are unknown. The method has shown no systematic error and a high degree of reproducibility (12).

As data were reported for idelalisib PFS by response rate, it was possible to digitize these Kaplan-Meier curves individually giving greater accuracy for composite survival curve (which otherwise is limited by the resolution of published data). The data for prior line PFS presented is in reality better described as prior-line TTP (all events were progressions – patients who died pre-progression would not go on to the next line of treatment), idelalisib TTP was then calculated from the published PFS curves for idelalisib – this meant attempting to identify any PFS events that were deaths (and not progression), and using them as censors in the recreated Kaplan-Meier data. To do this, the overall survival curve from Salles et al. was digitized, and event dates were compared. In total, we estimated 5 survival events to have occurred prior to disease progression, though it should be noted it is not possible to be certain that these were pre-progression deaths, and not deaths that occurred in post-progression patients around the same time as a different patient progressed.

Parametric curves were then fitted to the data, using a Weibull model which presented the best visual fit and statistical fit to the observed data by Akaike information criterion (AIC) and Bayesian Information Criterion (BIC) statistics. This model was also selected by the manufacturer in their Scottish Medicines Consortium (SMC) submission (10). Summary statistics for the digitized survival data closely approximated the summary statistics reported by the manufacturer, with the recreated prior line of therapy having a median of 5.4 months TTP (compared to a reported 5.1 months), and recreated idelalisib data a median TTP of 11.1 months, close to of the reported 11.0 months (13). The digitized data and survival curve fits are shown in Figure 1 for both idelalisib treated patients and the prior line of treatment.
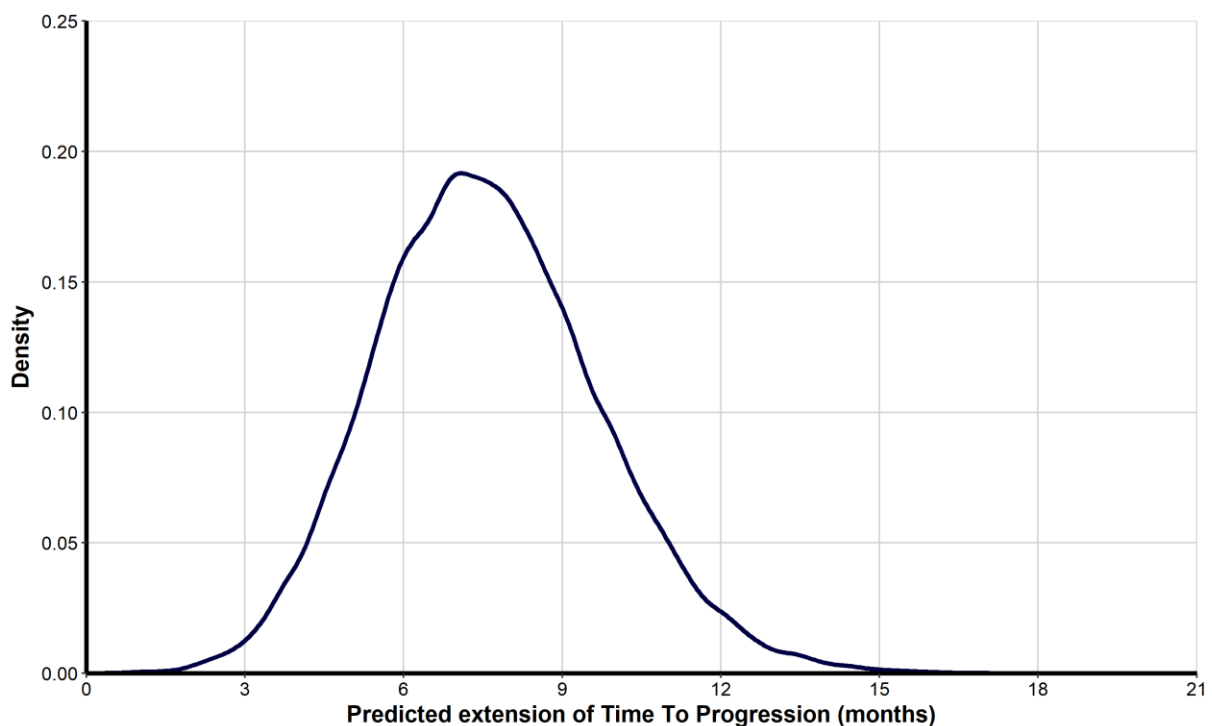
Figure 1: Estimated time to progression for idelalisib in FL compared to prior treatment time to progression including fitted Weibull curves
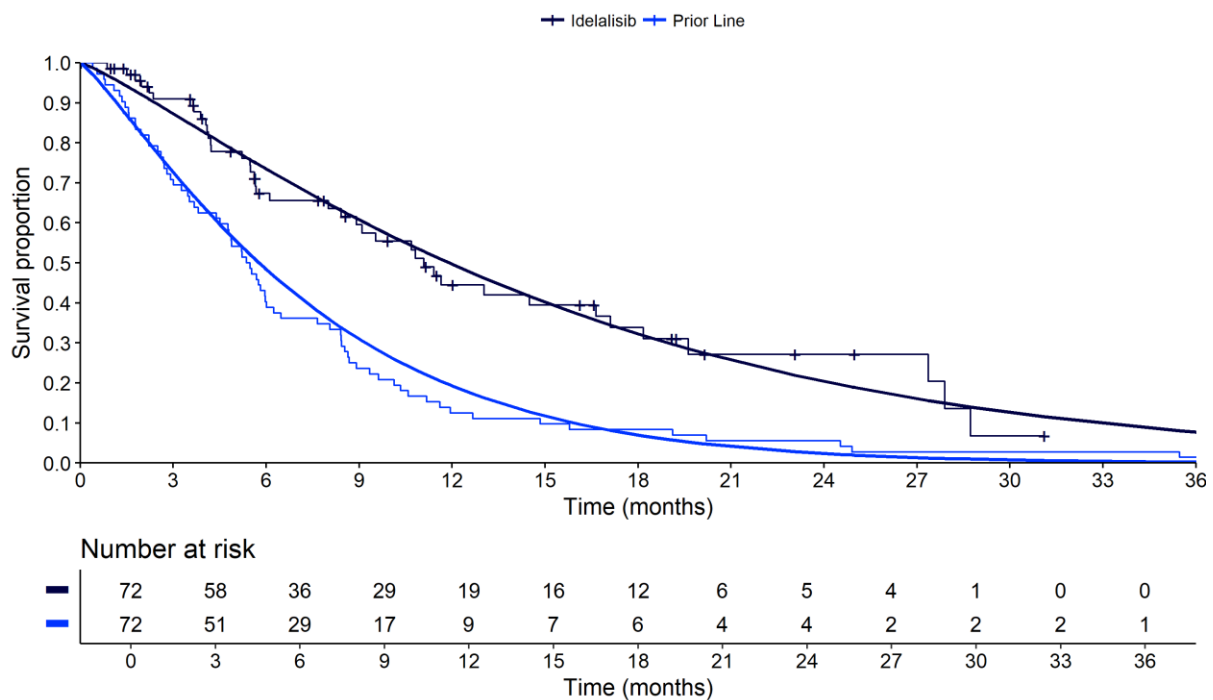
The area between the fitted survival curves gives an indication of the likely time to progression gain of idelalisib over the prior line of treatment. To estimate this difference between the extrapolated curves, 100,000 samples were taken from the fitted curve parameters, accounting for both the variance in parameter estimates, and covariance between parameter estimates. These were then

used to estimate the mean gain in TTP (estimated at a mean of 7.5 months), and shown as a density

plot in Figure 2. The approach of sampling from parametric survival curves is standard in

probabilistic sensitivity analysis (14–16), and allows for the uncertainty in estimates to be

represented – similar to bootstrapping, however sampling from the fitted distributions as opposed

to the raw data. The results of the analysis are then shown in a density plot of the predicted mean

survival gain i.e. difference between the area under the two sampled curves, which we would argue

is more informative than a simple confidence interval or interquartile range, though these can easily

be calculated or extracted.

Figure 2: Density plot of estimated time to progression benefit of idelalisib compared to outcomes seen with the prior line of

therapy

Overall, the results in Figures 1 and 2 suggest idelalisib offers an increase in time to progression. This

is likely a conservative estimate of the benefit of treatment, as the analysis assumes that

retreatment with drugs that have failed would yield the same time to disease progression as at the

previous line; in their SMC submission Gilead assumed time to progression at the next line of

treatment would be 10% shorter for retreatment (10).

2.4. Assumptions and limitations associated with the approach

Where Section 2.3 illustrated the applicability of the approach, this Section considers its viability and limitations in practice. The viability of this approach depends on the data available to the analyst. If a pharmaceutical company wishes to compare to a historical control of this type, they will have access to the PLD from their regulatory trial programme. In this case, comparison to previous treatment outcomes could be a useful and viable option with advantages over comparisons to aggregate data from trials not under their ownership (such as published historical controls). Documentation of clinical outcomes for prior therapy would be key, with the level of detail collected and robustness of data collection both vital. Key variables of interest may include the setting previous treatment was provided in, the treatments provided, and the dosage received - not all of which may be routinely available for prior therapy. If PLD are not available to the analyst, the creation of pseudo-PLD can be used, as shown in our illustrative example, though this does add a layer of uncertainty to the results.

Where viable, the merits of the approach described should be assessed on a case-by-case basis with consideration of the implicit assumptions involved. These include that the main aim of treatment is disease management, and that the TTP gains would translate to equivalent overall survival gains by delaying disease progression. Should prior therapy influence the subsequent line survival (either positively or negatively), the inference drawn from this approach could be misleading (though a similar approach with different PPS assumptions could be tested). Another implicit assumption in the approach presented is that survival risk is treatment independent i.e. patients have equal risk of death regardless of the treatment managing their disease. This assumption could be violated if different treatments have different life-threatening adverse event profiles, but could be relaxed easily within a similar framework.

3. Historical control as post-progression survival at the prior line of therapy from published evidence

3.1. General approach

When a treatment or intervention to extend survival is studied without a control arm (either active or placebo), often external uncontrolled datasets are used to inform the counterfactual outcomes; to estimate how long patients would have survived if they had received standard care rather than the intervention. Whilst there are issues with such comparisons (17–19), they often represent the most appropriate use of the best available clinical effectiveness evidence, however such historical controls are not always available.

Separate to the general approach in Section 2, where treatments are delivered in a defined pathway we propose that estimates can be made of the counterfactual survival time by estimating the time between disease progression and overall survival (OS) in external datasets of the previous line of treatment. The PPS for these patients may accurately reflect the period of time when a patient would have been eligible for the novel intervention (had it been available).

As clinical trials rarely follow all patients to PFS and OS completion (known as administrative censoring), we propose to estimate the difference between the two times (PFS and OS), using parametric curve fits to observed data (20). Using fitted survival curves, the time-to-event curve for PPS can be estimated - including estimating the unobserved portion.

3.2. Motivating example

Ofatumumab (Arzerra®, Novartis) was first licensed on the basis of the uncontrolled Hx-CD20-406 study in patients with chronic lymphocytic leukaemia (CLL) refractory to both fludarabine and alemtuzumab (21). In this study ofatumumab demonstrated an overall response rate (defined as the percentage of complete responders and partial responders) of 57% and progression free survival of 5.7 months, however there was no way to estimate the survival gain against standard of care. No historical control for patients receiving standard care was available in this patient population, making estimation of comparative effectiveness difficult.

Due to the lack of historical data at the same treatment line to use as an external control, the company performed a comparison against non-responders to estimate the survival gain of the product for use with health technology appraisal agencies for reimbursement submissions (22) – no comparative estimates were provided in regulatory submissions.

However, to be eligible for ofatumumab, patients must have been refractory to the previous line of treatment, alemtuzumab (MabCampath®, Genzyme). The pivotal study for alemtuzumab (the CAM211 study) enrolled 93 patients who had failed fludarabine therapy (23). Alemtuzumab was demonstrated to have a median disease-free survival (DFS) of 4.0 (95% CI 3.2 – 4.7) months, and a median OS of 15.9 (95% CI 11.8 – Not reached) months. These data, though from a prior line of therapy, have merit for understanding of the survival benefit associated with ofatumumab in CLL.

3.3. Application of the approach to the motivating example

In this Section, we describe implementation of the general approach to estimate the survival gain associated with ofatumumab for CLL using uncontrolled Hx-CD20-406 study data and comparator data from the CAM211 study.

In order to estimate the survival post disease progression on alemtuzumab treated patients in CAM211, PLD were recreated by digitizing the published survival curves then applying the method of Guyot et al. (12). Using the recreated PLD we fitted parametric curves using the process of Latimer to select the best fitting curves (20). The approach of Latimer gives an algorithm for selecting the preferred curve fits using the cumulative hazard function to select the appropriate set of parametric distributions, and then the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), clinical plausibility and visual inspection to select the most appropriate curve for the data.

Using the Latimer approach to curve fitting, we selected log-normal curve as the best fitting according to visual inspection as well as AIC and BIC (not shown) for both DFS and OS. The resulting digitized data and best fitting survival curves are shown in Figure 3 – the difference between the

curves being the time alive after treatment failure, and the time in which a patient would be eligible for ofatumumab.

Figure 3: Recreated patient level data for CAM211 study including log-normal curve fits
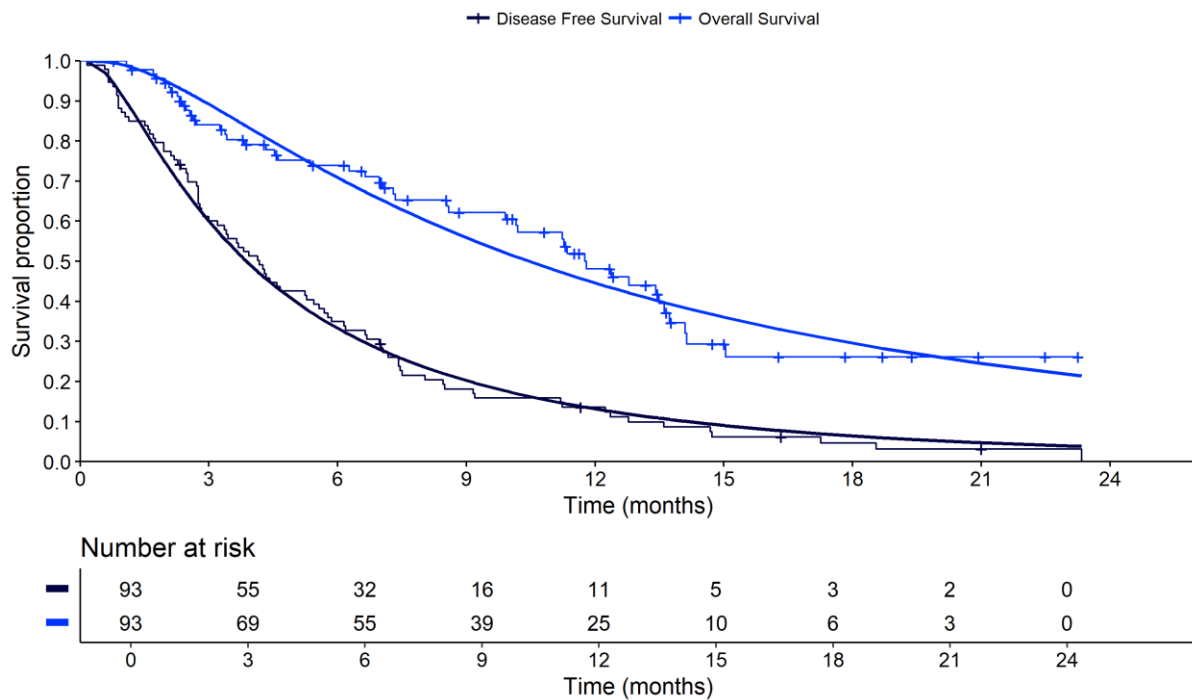

Using best fitting curves, 100,000 samples were taken and used to calculate the difference between DFS and OS (i.e. time spent alive after disease progression, PPS – when patients would be eligible for the next line of treatment) for the patients treated at this prior line. Across the samples taken the time alive after disease progression was projected to have a median of 12.3 months and mean of 13.0 months, though with a long tail. The same process was followed for the ofatumumab survival curve, where a Weibull curve was found to be the best fitting to the OS KM data, with estimated median survival of 14.6 months, and mean survival of 14.8 months (replicating published reanalysis of company data (24)). The digitised survival data for ofatumumab, fitted Weibull curve to the ofatumumab data, and time alive after disease progression (historical survival) are shown in Figure 4.

Figure 4: Digitized survival data and Weibull curve fit to Ofatumumab overall survival data, plotted against survival post disease progression from the CAM211 study
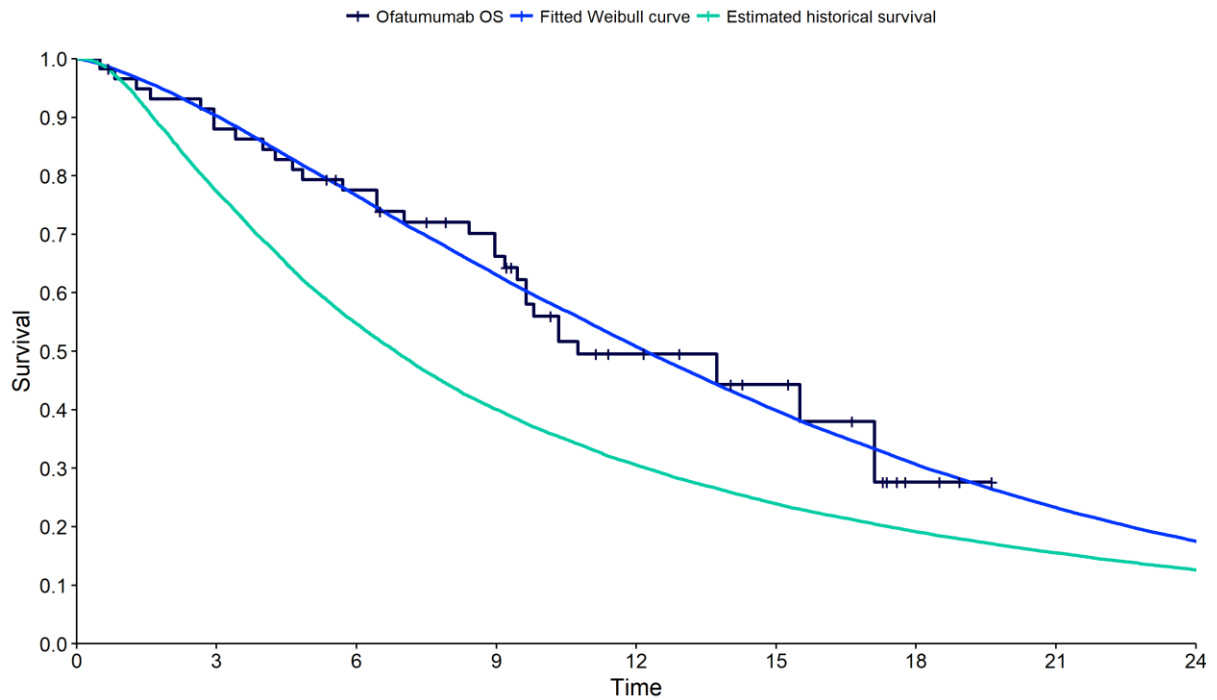
To estimate the survival gain of ofatumumab, the estimated time alive after disease progression was subtracted from the projected ofatumumab survival time for each of the 100,000 samples. As a part of the ofatumumab clinical study however, an exclusion criterion was that patients were not to have received any CLL treatment in the previous 4 weeks (21), with patients having had their disease progress a mean of 3.6 months prior to beginning ofatumumab treatment (25). Two further estimates are therefore presented in the density plot of the estimated survival gain for ofatumumab – firstly if 28 days (exclusion criteria for the trial), or 3.6 months (mean observed time) of time had elapsed before patients received treatment (Figure 5).

Figure 5: Density plot of estimated survival benefit of ofatumumab using assumptions of immediate treatment with ofatumumab, 28-day delay, and 56-day delay
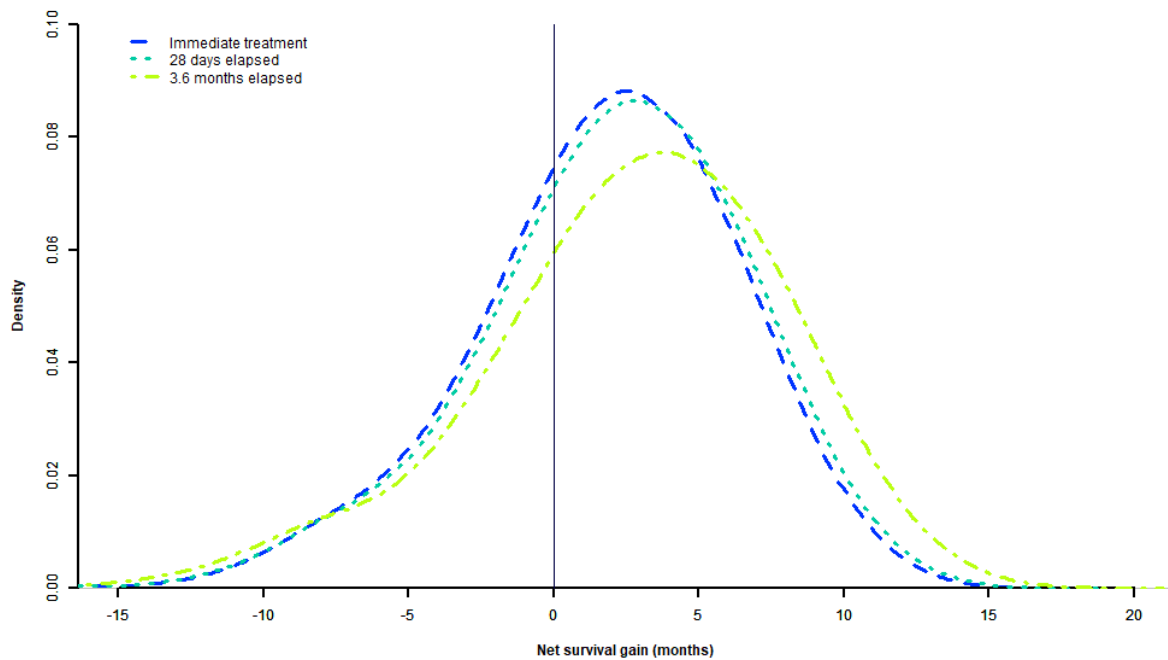
The resulting analysis shows that in a comparison against survival from disease progression, ofatumumab is estimated to provide a mean survival benefit of 2.0 months, with a 70% probability of survival being improved compared to the care patients received prior to the licensing of ofatumumab. If the likely delay in receiving the next line of therapy was 28 days, we estimate ofatumumab is estimated to improve survival by a mean of 2.2 months (with 71% probability of improving survival). With a delay of 3.6 months before ofatumumab treatment initiation, this survival gain is estimated at 2.8 months, with an 73% probability of improving survival. This final estimate is likely the most realistic in the context of publicly available information.

3.4. Assumptions and limitations associated with the approach

As with the approach considered in Section 2, there are several assumptions inherent in the approach explored in this Section, most notably the suitability of the data from the external, prior-line study. If the study for the previous line of treatment was conducted in a jurisdiction with different treatment options to that of the new treatment, then the post-disease-free survival results from this study are unlikely to be generalisable. A further limiting factor is that the trial data for the

previous line are unlikely to be available to the company who developed the subsequent

intervention. The company may therefore only be able to access information that is in the public

domain, which may not include for example details on patient demographics, subgroups, or the care

they received – all of which increase the level of uncertainty.

In terms of the comparability of patient populations, whilst there may be differences in patient

characteristics across different datasets these may be able to be adjusted for using methods such as

matching adjusted indirect comparison. Such adjustment is more complicated than in standard cross

trial comparisons, as data on patient characteristics are not routinely collected (or reported) at

disease progression. For example, the trial of ofatumumab required patients to have ECOG score of

0 or 1 on study entry (as did the alemtuzumab study) - if some patients had deteriorated by the time

of progression with alemtuzumab, this may in turn bias the results as they would have been

ineligible for entry to the ofatumumab study. There may be situations where such PLD are available

which would allow for further analysis, though this would need data to be made available by the

original study authors, or an older study conducted by the same company.

A further limitation of the approach is determining the point at which patients would be likely to

receive treatment, particularly in conditions where patients do not move immediately to a new

therapy. In the motivating example the trial protocol specified a minimum of 28 days since prior

therapy (which is loosely linked to disease progression), with a mean of 3.6 months from disease

progression until patients received ofatumumab. Ensuring time since previous treatment is collected

in studies where this approach may be needed may reduce some of this uncertainty, which may

need to be formally incorporated in to a probabilistic model.

4. Discussion

4.1. Merit of the approaches

The two methods presented may allow an additional estimate of the benefit of a novel intervention

in specific scenarios. Such analysis may be useful supplement to a more traditional controlled

comparison (where the control is either internal to the trial, or via a more traditional historical control), or serve as a comparison where no other data are available. Such comparisons are of relevance to regulators who in both motivating examples had difficulty in quantifying the benefit of the products. The approaches are likely also of interest to payers who have difficulty in estimating the value for money of products similar to those used as examples, and may also be of interest to clinicians who need to understand the comparative effectiveness of different products to make treatment recommendations.

Whilst the approaches proposed help to characterise the uncertainty in comparisons where neither direct nor indirect evidence is available, their validity will need to be assessed in each case prior to the methods being applied. The trials used in the analysis and disease course must be carefully scrutinized for appropriateness, and the structural uncertainty understood. Ultimately, estimates based on epidemiological and statistical analysis cannot substitute for randomised controlled trials, and will always be subject to a degree of unquantifiable uncertainty.

Due to the uncertainty around results from both approaches, we recommend that when the approaches are used, careful justification is given to survival curve fitting, particularly the long-term plausibility which has the potential to drive the results. This is especially the case when projecting PPS times (approach 2), as this is using the data in the tail of the curve, where the underlying data are most scarce.

The companion R file we provide should make the application of these methods more straightforward. In the file we simulate data, and then make comparisons using the approaches highlighted – showing the steps required. Loading patient data in place of the simulated data should greatly shorten the time required to understand and implement the methods.

4.2. Implications

The work performed has implications not only for the conduct of analysis, but also for data collection. The approach presented in Section 2 requires that data be collected within the pivotal trial regarding the previous line of treatment. Should full details not be available, the analyses that can be performed become more limited. Prior to commencing a study where such an approach would be required, sponsors should therefore carefully consider the data that need to be collected regarding prior treatment(s).

The approach presented in Section 3 also requires data collection, but in a different way. Firstly, in literature searching, efforts should be made to identify data from the prior line of treatment as well as for data from more traditional historical controls. The available data (and desire to make studies comparable) may also shape the design of the subsequent study in terms of inclusion/exclusion criteria, or data collected. For example, even if a patient characteristic is not thought to be prognostic or relevant for the novel intervention, it may be prudent to collect such data so the similarity of patients between the planned study and historical control (or potential historical control) can be assessed. Companies with multiple assets in a given disease area may also wish to continue data collection beyond disease progression, with future comparisons in mind.

In order to make the most compelling case possible for the use of an intervention, companies should therefore carefully consider their study design. Although they are unlikely to change the nature of their study from an uncontrolled/single arm trial to an RCT, they may still wish to consider additional data collection to enable estimates of comparative effectiveness to be made.

4.3. Future research

The approaches presented demonstrate how control arms may be synthesized in their absence, both methods however require validation of their predictive accuracy. We would suggest the ideal validation for the approaches would be to use the methods to estimate the outcomes of control arms for RCTs, and compare the predicted outcomes to those observed in the non-interventional

arm of the RCT. Such an approach would ideally be performed to show the validity of the approach in a given disease, prior to the use of the method for decision making.

A development that could be considered for each approach is a fully Bayesian model, with priors specified using expert opinion. This would likely reduce the distribution of predicted net benefit for each of the models as whilst some extreme values of survival for both arms are technically possible, they may be implausible (which would be reflected in the prior used). The major advantage of a Bayesian model however would be the ability to parameterise the additional assumptions required – for example how much shorter remission would be with each subsequent line of treatment in the TTP approach, or the elapsed time before treatment is commenced as in the OS minus PFS approach. Instead of sensitivity analyses, best estimates with an appropriate distribution could be included in the model.

A second area where further research would be welcome is how multiple historical controls should be handled, when available. It is entirely conceivable that in the second approach presented (the use of PPS from the prior line) multiple trials may be available, each of which would yield a potential historical control. Alternatively, there may be published historical controls, as well as estimates made using the method(s) presented – it is not clear how such a data package should be synthesized or used in decision making. A recent focus of work has been that of dynamic 'borrowing' from historical trials (adequately down-weighted), to increase the sample size of contemporary control arms (26). Whilst in our examples no historical data exists (nor internal controls for data to be added to), similar methods with each source weighted according to its utility may be of use.

4.4. Conclusions

The work presented shows two non-standard approaches which analysts can use to construct estimates of comparative effectiveness in the absence of direct or indirect comparator evidence. These approaches give additional options for analysts in estimating comparative evidence, and

whilst these do not substitute for the availability of controlled trials, they may help inform decision making.

The implications of the work are that companies conducting uncontrolled studies, or studies in area where there is no network of evidence should consider how they intend to use data from studies (particularly uncontrolled studies) to produce estimates of comparative effectiveness. Although the study design is unlikely to change, exact data collected may need to be reviewed should either of these methods be required.

References

1. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The Results of Direct and Indirect Treatment Comparisons in Meta-Analysis of Randomized Controlled Trial. J Clin Epidemiol. 1997;50(6):683–91.

2. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Stat Med. 2004 Oct 30;23(20):3105–24.

3. Pocock SJ. The combination of randomized and historical controls in clinical trials. J Chronic Dis. 1976;29(3):175–188.

4. Caro JJ, Ishak KJ. No head-to-head trial? Simulate the missing arms. Pharmacoeconomics. 2010;28(10):957–967.

5. Signorovitch JE, Sikirica V, Erder MH, Xie J, Lu M, Hodgkins PS, et al. Matching-Adjusted Indirect Comparisons: A New Tool for Timely Comparative Effectiveness Research. Value Health. 2012 Sep;15(6):940–7.

6. Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available: Table 1. Am J Epidemiol. 2016 Mar 18;kwv254.

7. Dreyling M, Ghielmini M, Rule S, Salles G, Vitolo U, Ladetto M. Newly diagnosed and relapsed follicular lymphoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up†. Ann Oncol. 2016 Sep;27(suppl_5):v83–90.

8. Treatment for low grade non Hodgkin lymphoma | Non-Hodgkin lymphoma | Cancer Research UK [Internet]. [cited 2017 Sep 27]. Available from: http://www.cancerresearchuk.org/about-cancer/non-hodgkin-lymphoma/treatment/low-grade

9. Gopal AK, Kahl BS, de Vos S, Wagner-Johnston ND, Schuster SJ, Jurczak WJ, et al. PI3Kδ Inhibition by Idelalisib in Patients with Relapsed Indolent Lymphoma. N Engl J Med. 2014 Jan 22;370(11):1008–18.

10. Scottish Medicines Consortium. Idelalisib SMC advice number 1039/15 [Internet]. 2015 [cited 2017 Dec 21]. Available from: https://www.scottishmedicines.org.uk/files/advice/idelalisib_Zydelig_FINAL_April_2015_for_website.pdf

11. Salles G, Schuster SJ, de Vos S, Wagner-Johnston ND, Viardot A, Blum KA, et al. Efficacy and safety of idelalisib in patients with relapsed, rituximab- and alkylating agent-refractory follicular lymphoma: a subgroup analysis of a phase 2 study. Haematologica. 2017 Apr;102(4):e156–9.

12. Guyot P, Ades A, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. BMC Med Res Methodol. 2012;12:9.

13. PBAC. Public summary document - November 2015 Meeting - Idelalisib [Internet]. 2015 [cited 2017 Dec 21]. Available from: http://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2015-11/files/idelalisib-follicular-lymphoma-psd-november-2015.pdf

14. Edlin R, McCabe C, Hulme C. Cost Effectiveness Modelling for Health Technology Assessment: A Practical Course. 1st ed. 2015 edition. ADIS; 2015. 224 p.

15. Briggs A. Decision Modelling for Health Economic Evaluation. 1 edition. Oxford: Oxford University Press, USA; 2006. 250 p.

16. Baio G, Dawid AP. Probabilistic sensitivity analysis in health economics. Stat Methods Med Res. 2015 Dec 1;24(6):615–34.

17. Diehl LF, Perry DJ. A comparison of randomized concurrent control groups with matched historical control groups: are historical controls valid? J Clin Oncol. 1986;4(7):1114–1120.

18. MacLehose RR, Reeves BC, Harvey IM, Sheldon TA, Russell IT, Black AM. A systematic review of comparisons of effect sizes derived from randomised and non-randomised studies. Health Technol Assess Winch Engl. 2000;4(34):1–154.

19. Moroz V, Wilson JS, Kearns P, Wheatley K. Comparison of anticipated and actual control group outcomes in randomised trials in paediatric oncology provides evidence that historically controlled studies are biased in favour of the novel treatment. Trials. 2014;15(1):481.

20. Latimer NR. Survival Analysis for Economic Evaluations Alongside Clinical Trials—Extrapolation with Patient-Level Data: Inconsistencies, Limitations, and a Practical Guide. Med Decis Making. 2013 Aug 1;33(6):743–54.

21. Wierda WG, Kipps TJ, Mayer J, Stilgenbauer S, Williams CD, Hellmann A, et al. Ofatumumab As Single-Agent CD20 Immunotherapy in Fludarabine-Refractory Chronic Lymphocytic Leukemia. J Clin Oncol. 2010 Apr 1;28(10):1749–55.

22. Hatswell AJ, Thompson GJ, Maroudas PA, Sofrygin O, Delea TE. Estimating outcomes and cost effectiveness using a single-arm clinical trial: ofatumumab for double-refractory chronic lymphocytic leukemia. Cost Eff Resour Alloc. 2017 May 26;15:8.

23. Keating MJ, Flinn I, Jain V, Binet J-L, Hillmen P, Byrd J, et al. Therapeutic role of alemtuzumab (Campath-1H) in patients who have failed fludarabine: results of a large international study. Blood. 2002 May 15;99(10):3554–61.

24. Almond C, Stevens J, Ren K, Batty A. The estimated survival of patient with double refractory chronic lymphocytic leukaemia: A reanalysis of NICE TA202 using bayesian methodology to model observed data. ISPOR 16th Annual European Congress; 2013 Nov; Dublin, Ireland.

25. GlaxoSmithKline. Interim Clinical Trial Report - Hx-CD20-406 [Internet]. 2008 [cited 2018 Apr 13]. Available from: https://www.gsk-clinicalstudyregister.com/files2/gsk-111773-clinical-study-report-redact.pdf

26. Viele K, Berry S, Neuenschwander B, Amzal B, Chen F, Enas N, et al. Use of historical control data for assessing treatment effects in clinical trials. Pharm Stat. 2014 Jan;13(1):41–54.