



This is a repository copy of *R and Shiny for cost-effectiveness analyses : why and when? A hypothetical case study.*

White Rose Research Online URL for this paper:
<https://eprints.whiterose.ac.uk/161083/>

Version: Accepted Version

Article:

Hart, R., Burns, D., Ramaekers, B. et al. (8 more authors) (2020) R and Shiny for cost-effectiveness analyses : why and when? A hypothetical case study. *PharmacoEconomics*, 38. pp. 765-776. ISSN 1170-7690

<https://doi.org/10.1007/s40273-020-00903-9>

This is a post-peer-review, pre-copyedit version of an article published in *PharmacoEconomics*. The final authenticated version is available online at:
<https://doi.org/10.1007/s40273-020-00903-9>

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

Title: R and *Shiny* for cost-effectiveness analyses: why and when? A hypothetical case study

Authors: Rose Hart¹, Darren Burns¹, Bram Ramaekers², Shijie Ren³, Daniel Gladwell¹, Will Sullivan¹, Niall Davison⁴, Owain Saunders⁴, Indeg Sly¹, Theresa Cain¹, Dawn Lee¹

¹ BresMed Health Solutions, Sheffield, UK

² Department of Clinical Epidemiology and Medical Technology Assessment, Maastricht University Medical Center, Maastricht, Netherlands

³ School of Health and Related Research, University of Sheffield, Sheffield, UK

⁴ BresMed Health Solutions, Manchester, UK

Corresponding author: Rose Hart; email: rhart@bresmed.com; Tel: +44 (0)114 309 4372

Author ORcIDs: Rose Hart, 0000-0001-7302-2559; Darren Burns, 0000-0002-5209-8041; Bram Ramaekers, 0000-0001-5785-9228; Shijie Ren, 0000-0003-3568-7124; Will Sullivan, 0000-0001-5221-7074; Niall Davison, 0000-0002-9447-0825.

Word limit: 4,000 words (Original Research Article) – current = 3,982 (excluding in-text citations)

Abstract: max 250 words – current = 208

Abstract

Introduction: Health economics models are typically built in Microsoft Excel[®] owing to its wide familiarity, accessibility and perceived transparency. However, given the increasingly rapid and analytically complex decision-making needs of both the pharmaceutical industry and the field of health economics and outcomes research (HEOR), the demands of cost-effectiveness analyses may be better met by the programming language R.

Objective: This case study provides an explicit comparison between Excel and R for contemporary cost-effectiveness analysis.

Methods: We constructed duplicate cost-effectiveness models using Excel and R (with a user interface built using the *Shiny* package) to address a hypothetical case study typical of contemporary health technology assessment.

Results: We compared R and Excel versions of the same model design to determine the advantages and limitations of the modelling platforms in terms of (i) analytical capability, (ii) data safety, (iii) building considerations, (iv) usability for technical and non-technical users and (v) model adaptability.

Conclusions: The findings of this explicit comparison are used to produce recommendations for when R might be more suitable than Excel in contemporary cost-effectiveness analyses. We conclude that selection of appropriate modelling software needs to consider case-by-case modelling requirements, particularly (i) intended audience, (ii) complexity of analysis, (iii) nature and frequency of updates and (iv) anticipated model run-time.

Key points

- We present duplicate health economics models for a hypothetical chimeric antigen receptor treatment, built in Excel and R with an interface using the *Shiny* package
- We provide an explicit comparison between the two software platforms for contemporary cost-effectiveness analysis
- We discuss when R might be preferable to Excel in contemporary cost-effectiveness analysis and the implications of a shift from Excel towards R for transparency and efficiency within HEOR

1. Introduction

1.1. R as an alternative to Excel for cost-effectiveness modelling

Since health technology assessment (HTA)-informed decision-making came to prominence in the late 20th century, the software of choice for health economic models has been Microsoft Excel[®]. This is understandable, Microsoft Office[®] has been almost as ubiquitous as the desktop computer since the 1990s. For cohort-level models, Excel's spreadsheet interface is typically sufficient to calculate deterministic results in near-real time. The low cost and fluency benefits of Excel in this context are not limited to the model builder. A model built in Excel is more accessible to reviewers, managers, clinicians and patients than one presented in code. For an alternative platform to be preferential to Excel in a health economic modelling application, its relative build and execution advantages should be considered alongside accessibility, fluency and transparency.

R has been available as open-source software since 2000 [1, 2], and the use of R for health economic modelling is not new. Examples of note include the cost-effectiveness model for the 2006 HTA monograph for etanercept and efalizumab to treat psoriasis, the model developed by the University of Sheffield School of Health and Related Research for the assessment of rituximab for the first-line treatment of Stage III–IV follicular lymphoma in 2011, and the assessment of opioid agonists by the US Institute for Clinical and Economic Review [3-5]. Recently, R has been increasingly considered as an alternative to Excel for health economic modelling. In July 2018, University College London hosted its first annual event with presentations from academic and industry modellers working in health economic evaluation, exploring and promoting the use of R as an alternative to Excel for cost-effectiveness analysis [6]. In addition, numerous papers have been published demonstrating the benefits of statistical software over Excel [7-12].

For health economic modelling, R holds a key advantage over Excel. With its primary function as a statistical software, R allows the user to consolidate and calculate in real time the underlying statistics informing health economic decision problems (referred to henceforth as 'end-to-end' functionality) – for example, survival analysis of patient-level registrational trial data to inform a three-state oncology model. End-to-end functionality is of considerable advantage as it allows both internal and external validity to be demonstrated all the way from input data to decision analytic endpoint, with uncertainty analyses and reporting all in one place. Given a comparable level of modelling experience, construction in R with end-to-end functionality provides superior consistency and efficiency compared to Excel, an argument that has been made elsewhere [12].

1.2. Comparative transparency of R versus Excel

The advantages of Excel over R in current HTA modelling include perception of transparency and audience familiarity [13-15]. Transparency can be defined as provision of documentation on a model's structure, equations, parameter values and assumptions. To be fully transparent, the model needs to be understandable to both non-modellers and modellers who may want to replicate the model and findings [16, 17]. Models implemented using R are transparent in ways unfamiliar to Excel. End-to-end functionality lends itself to more transparent HTA submissions, as functions can be tested in isolation, and internal and external validation throughout statistical and modelling steps can be incorporated into the model itself. R packages can be freely validated as they are open-source, while R scripts force the analyst to construct models in a linear fashion, which are arguably easier to grasp than often disjointed and sometimes incoherent Excel implementations. For additional transparency, outputs of analyses conducted in R can be embedded within pre-formatted reports via the text conversion tool, R Markdown, such that documents are readily reproducible and a clear trail from analysis to output can be followed [18]. An additional benefit of R is that code development can be managed in tracked stages using readily accessible version control software, allowing parallel workstreams with staged reviews and merges [19]. However, those same HTA stakeholders who likely recognize Excel spreadsheets – reviewers, managers, clinicians and patients – are unlikely to be familiar with R code. In this sense, an Excel-based health economic model has clear transparency advantages over a similar model in R. In late 2019, this translates as a hard barrier for R for global HTA cost-effectiveness models: some key HTA bodies do not accept models in R, including the Dutch National Health Care Institute, Zorginstituut Nederland [20-22]. For multinational pharmaceutical companies developing submissions to HTA bodies globally, this can understandably be enough to preclude R as a choice for late-stage cost-effectiveness modelling. However, R is becoming more familiar, and HTA bodies are increasingly moving to accept cost-effectiveness models in R [23].

1.3. Incorporation of *Shiny* within R models to improve user-accessibility and produce applications for online distribution

Shiny is a package for R that allows construction of an interactive graphical user-interface associated with the accompanying R code [24]. *Shiny* is increasingly being used by health economic modellers [25]. A model built with *Shiny* comprises two parts: a user-interface and a server. The *user-interface* is designed to be *user-friendly*; the server operates as the engine. *Shiny* allows the user to amend inputs in a similar manner to Excel without needing to access the background R code. In this sense, an R model using *Shiny* is comparable to an Excel model developed in part within Visual Basic for Applications® (VBA) code, but with capability for complex statistical analysis, model flexibility and high-quality visuals. Several *Shiny*-based health economic models have been disseminated to date, including two open-source value tools from the *Innovation and Value Initiative* (IVI) [26]. Additionally, initiatives such as those undertaken by the US Food and Drug Administration (FDA) tool for searching drug safety reporting ratios, the Personalized Medicine Economics Research tool developed by the University of Washington, and the Global health cost-effectiveness analysis registry produced by the Center for the Evaluation of Value and Risk demonstrate the application and relevance of *Shiny* from a broader HTA perspective [27-29].

Of course, a *Shiny* interface is not sufficient for a reviewer to understand the underlying mechanisms of a model. However, its use serves as a bridge between complex functionality and model accessibility and may make R more appealing to health economic audiences.

1.4. The increasing need for analytical complexity to estimate long-term treatment effectiveness: a case study for our times

The benefit of R over Excel as a platform for both statistical analysis (e.g. survival analysis) of patient-level data (PLD) and cost-effectiveness modelling is becoming increasingly important, owing to changes in the development, delivery and funding of medicines. For technologies addressing high unmet need, the European Medicines Agency (EMA) and the FDA have lowered the evidence generation hurdle, allowing interventions to receive accelerated approval based on single-arm Phase II study data with limited follow-up [30-32]. Between 1 January 1999 and 8 May 2014, 76 unique licences were granted by the EMA and FDA in the absence of randomized controlled trial data [33]. Consequently, a far greater scope of statistical analysis is required to appropriately reflect the uncertainty in the expected treatment benefit estimates from such limited data.

In 2018, the first EMA and FDA chimeric antigen receptor (CAR) T-cell therapy licences were granted for use of tisagenlecleucel and axicabtagene ciloleucel in haematological malignancies [34-37]. When assessing the cost effectiveness of a CAR T-cell therapy at time of regulatory approval, characteristic issues include (i) potential benefits that far exceed the relatively short duration of trial follow up, and (ii) a lack of randomized evidence versus the comparator [38]. Owing to the novelty of CAR T-cell therapy, combined with the complexity of statistical analyses required to adequately assess cost and survival outcomes, this technology was selected as the intervention in our hypothetical case study as an exemplar of a contemporary problem of our time [39, 40]. This case study was used to produce an explicit comparison between Excel and R with *Shiny* interface in cost-effectiveness analysis. We have chosen a hypothetical case study using simulated data because the objective of this paper is to compare modelling platforms, rather than deriving a cost-effectiveness ‘answer’, and because the authors do not own PLD from regulatory trials. The decision problem presented here uses a simulated paediatric B-cell acute lymphoblastic leukaemia patient dataset, mimicking typical data limitations of the indication as described in Hettle et al. [38], to model the intervention’s incremental benefit versus chemotherapy.

1.5. Aims

Our primary aim is to present novel, illustrative evidence on the relative appropriateness of R versus Excel for contemporary cost-effectiveness modelling. A secondary aim is to add to the growing demonstrative evidence that provision of a graphical user-interface (such as *Shiny*) can improve R model accessibility. To this end, we built equivalent cost-effectiveness models in both R with *Shiny* interface and Excel to produce an explicit comparison between modelling platforms.

2. Methods

2.1. Model software and access

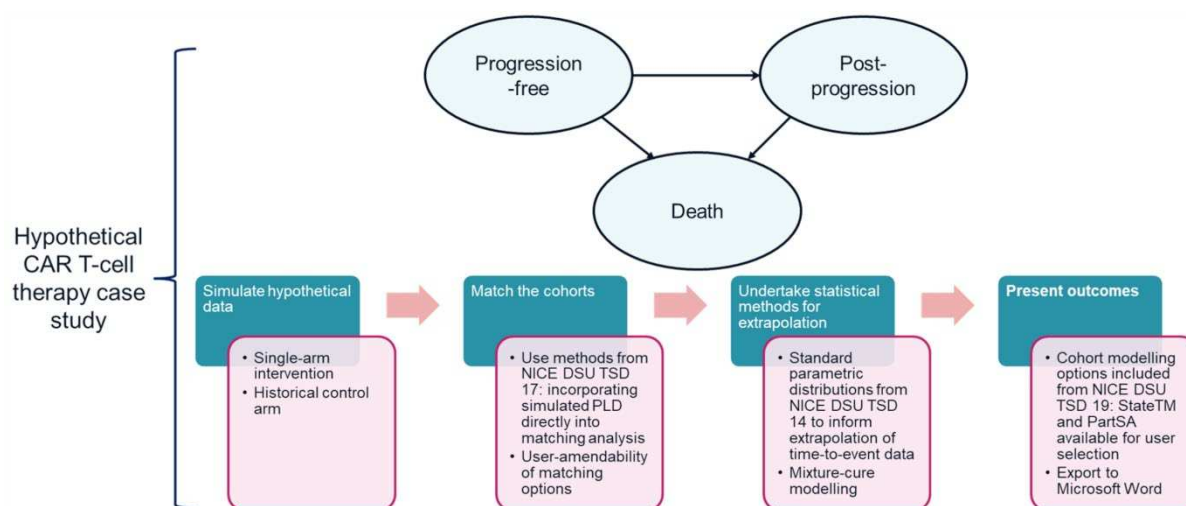
A model was constructed in R with *Shiny* and is referred to here as the intRface™ model (Interactive R: Flexible Applications for Cost-effectiveness – available to view via the following link: [intRface](#)). The purpose of the subsequent Excel model build was twofold: (i) for validation and (ii) for comparison. The Excel model is available to view in Online Resource 1. Both models used the same base case inputs and produced the same base case results, fulfilled the objective of addressing the decision problem and adhered to the National Institute for Health and Care Excellence (NICE) Decision Support Unit (DSU) modelling guidelines. The methods, assumptions and measured outcomes of the models are summarized in Online Resource 2.

2.2. Cost-effectiveness model structure

The hypothetical case study was approached from a UK perspective, using guidelines from the NICE DSU Technical Support Documents (TSDs). Broadly in line with the 2018 NICE single technology appraisals (STAs) of tisagenlecleucel and axicabtagene ciloleucel within their licensed indications [41-43], we used a cohort-level cost-effectiveness model, structured with three health states: progression-free (PF), post-progression (PP) and death. The methods described in the NICE DSU TSD 19 to determine cohort health state membership over time were used to include both state-transition modelling (StateTM) and partitioned survival analysis (PartSA) structures. This was owing to the recommendation that where PartSA is used, StateTM should also be used to assist in verifying the plausibility of extrapolations and to address uncertainties in the extrapolation period [44].

The model objective was to capture and summarize lifetime patient health outcomes and UK National Health Service costs. To measure this, an 80-year time horizon and a 28-day cycle length were used with a 3.5% annual discount rate for costs and quality-adjusted life years [45]. Fig. 1 summarizes the model structure and analysis steps taken to capture cost-effectiveness outcomes for the decision problem.

Fig. 1: Schematic of objectives for cost-effectiveness model end-to-end functionality



CAR, chimeric antigen receptor; DSU, Decision Support Unit; NICE, National Institute for Health and Care Excellence; PartSA, partitioned survival analysis; PLD, patient-level data; StateTM, state-transition modelling; TSD, technical support document.

2.3. Analysis requirements

The technical complexities of the decision problem were considered when simulating the PLD and scoping the analysis requirements. PLD were simulated for CAR T-cell therapy and clofarabine as two single-arm trials, each with two data-cuts. A simulated 24-month data-cut was used as the base case for all analyses, with the option to switch to a 36-month data-cut to demonstrate the inclusion of an 'updated data-cut'. Prognostic patient characteristics as reported by studies in paediatric B-cell acute lymphoblastic leukaemia [46, 47] were used alongside available Kaplan–Meier (KM) data reported by Hettle et al. and in the axicabtagene ciloleucel NICE STA [48, 38] to simulate PLD, consisting of patient characteristics as well as censoring and event data for progression-free survival (PFS) and overall survival (OS).

Matching analysis functionality was included in the model as the simulated single-arm trials necessitated that an analysis be included to account for the lack of randomization and potential for bias of treatment effect. NICE DSU TSD 17 guidance [49] outlines several methods to address the problems presented by non-randomized data. Based on assumptions that the simulated data observed all factors that influence relevant patient outcomes, and that there was sufficient overlap of similar patients between the two trials, propensity score matching analysis was included. Matching aims to mimic trial randomization, creating a subset of the data by selecting patients from the treated and control groups who are similar to each other in baseline characteristics [50, 51].

To extrapolate health outcomes over the time horizon, NICE DSU TSD 14 guidance [52] was consulted to determine the parametric survival models used for analysing the time-to-event data. In line with this guidance, functionality was included that recommends suitable outcomes for assessing the validity of distributions, including visual inspection of the extrapolations versus the KM and providing the Akaike information criterion and Bayesian information criterion statistics. Parametric model outcomes were then used to determine health state membership over time for both StateTM and PartSA.

Different patient groups may be expected to experience different survival rates depending on their response to immunotherapies, and it has been reported that CAR T-cell therapy has the potential to induce long-term remission [41-43]. A frequent limitation of emerging therapies is that patient follow-up is insufficient and traditional parametric extrapolations of survival may not be representative of real-world patient outcomes. “Mixture-cure” functionality was therefore included as a modelling option to allow the production of outcomes where a proportion of the patient population is ‘cured’ and therefore experiences a probability of death equal to that of the calculated age-matched general population estimates [53, 54]. The ‘cured’ and ‘uncured’ patient populations are then combined to produce curves for the overall population [55, 56].

3. Results

An explicit comparison between the constructed intRface and Excel models was undertaken to consider key issues often encountered within health economic modelling surrounding transparency, build and adaptation efficiency, or analytical limitations: (i) analytical capability, (ii) data safety, (iii) model building, (iv) usability for technical and non-technical users and (v) model adaptability. We use this comparison to recommend scoping considerations required to inform the suitability of R for cost-effectiveness modelling on a case-by-case basis.

3.1. Comparative model capability

The analytical capabilities of R are such that a model built using this platform can include all the required stages of analysis from using the PLD to reporting outcomes, giving models end-to-end functionality. This is not possible in Excel because of the more limited analysis capabilities. Some inputs for the Excel model needed to be derived outside of Excel, such as the propensity score matching results and parametric survival curve parameters. In this case study, they were extracted from the R model. The analytical capabilities of each model are presented in Table 1.

Table 1: Comparative analytical capabilities between intRface and Excel models

Functionality	intRface model	Excel model
Live propensity score	Included with the ability to add or	Not possible to include: only

Functionality	intRface model	Excel model
matching analysis	remove existing covariates from the analysis. The unmatched data can also be used if propensity score matching is not required	scenarios that have been previously run can be explored. Matching analysis needs to be run externally, and results copied into model
Live fitting of parametric models	All parametric models are fitted to the active dataset	Parametric models need to be fitted to the active dataset externally, and results copied into model – a laborious task for updates to data-cut or subgroup exploration
PartSA (including MCM) and StateTM modelling	Model includes PartSA and StateTM modelling strategies. These are informed by the internally calculated parametric fits	Model includes PartSA and StateTM modelling strategies. These are informed by models fit outside of Excel with estimates pasted in
Probabilistic sensitivity analysis – time taken for 1,000 PSA runs using base case settings	1.42 minutes	13.2 minutes
One-way sensitivity analysis – time taken to run 109 parameter scenarios	0.27 minutes	2.4 minutes
Automatic report generation	Report template is set up within RMD to automatically populate tables and figures with active modelling analyses when selected	Highly challenging to include; not included
Quality control	<ul style="list-style-type: none"> Table included with selected diagnostic checks Linear code with vectors and data frames produced by single calculations that need to be checked once. However, tracing an individual calculation from start to finish can take longer than in Excel Packages used are open-source: version to be used needs to be defined to ensure stability over time 	<ul style="list-style-type: none"> Diagnostic checks included in patient flow sheet Cell-by-cell checks were required across all sheets because of individual calculations, meaning there was potential for drag down error and inconsistency within columns and data frames
Model size	5.1 MB – Includes R scripts and Excel input workbooks containing simulated PLD, general population survival statistics and cost inputs	30.9 MB – Single workbook
Version control	Managed by the version control software Git to allow tracked changes, code reversions and parallel work streams	Manual change log. Multiple versions required to allow reversions. Difficult to work in parallel
Key: MB, megabytes; MCM, mixture-cure modelling; PartSA, partitioned survival analysis; PLD, patient-level data; RMD, R Markdown; StateTM, state transition model.		

3.2. Data safety

The incorporation of externally derived values into cost-effectiveness models is widely practised to preserve the confidentiality of trial data; however, this is not possible for a model with end-to-end functionality. Where PLD are incorporated within end-to-end models, they can be used both locally and online via hosting software and secure servers with user authentication [57]. However, in stages of model development where external distribution and validation are required, it is understood that having any PLD contained within the model may breach data security. As within Excel, in R it is possible to remove the PLD from a model and save the results and statistics of analyses requiring PLD for later use in cost-effectiveness analysis. This delinking from PLD dependency should be done at the last moment prior to external dissemination to allow full functionality to be maintained as long as required.

3.3. Model building

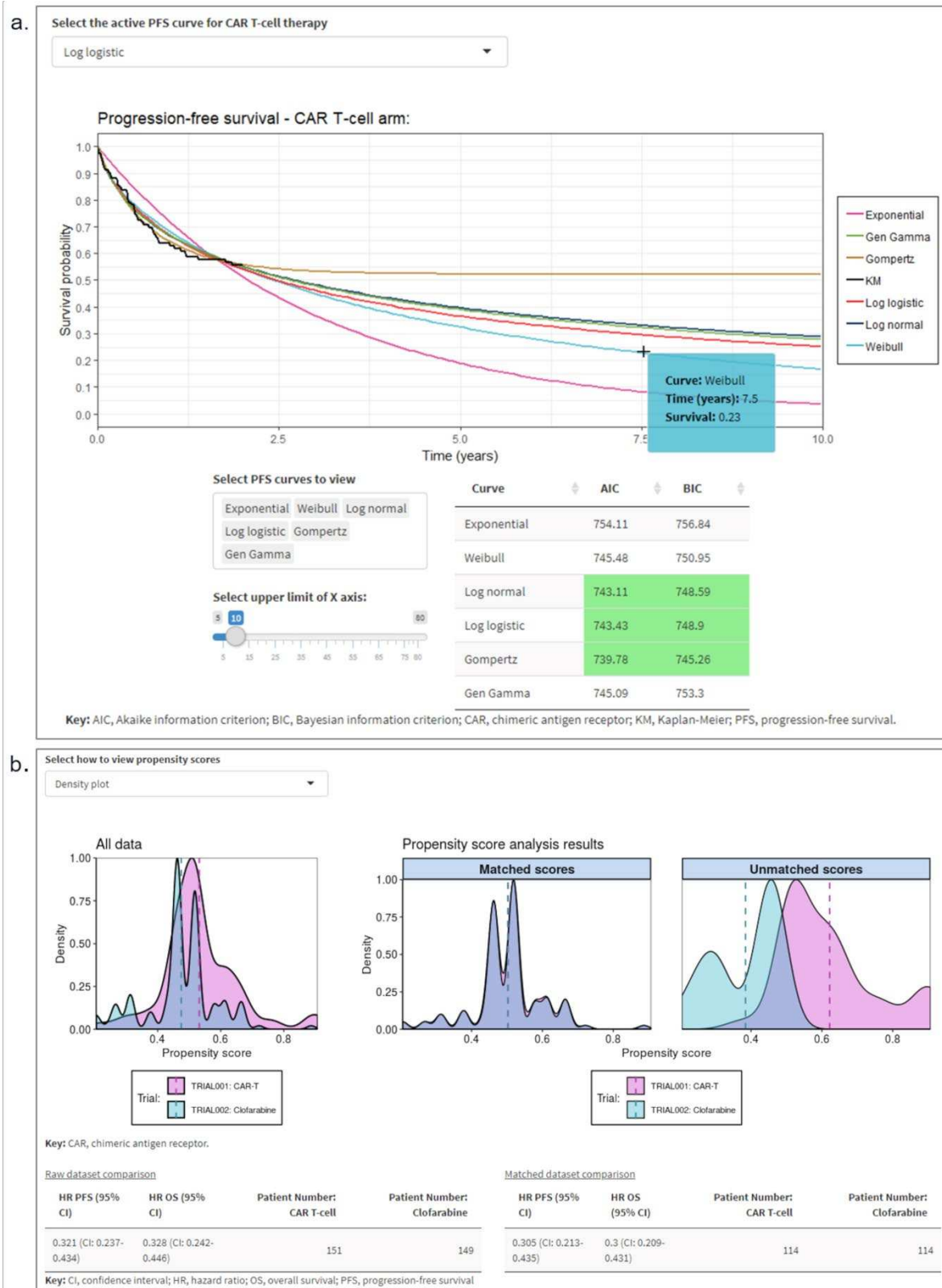
Developing the intRface model presented coding challenges that were not required in producing the Excel model, owing to the differences in the included functionality. Consequently, the intRface model build took significantly longer than the Excel model. However, the Excel model was dependent on the matching analysis and survival extrapolation analysis from R. This would remain so even if the full intRface model was not constructed, and there would still be the need to revisit R to obtain new estimates upon any change in data. It should also be acknowledged that the flow of information within an R model wrapped in *Shiny* functionality is not as linear as a model produced in R alone, which adds additional coding complexity. It is for these reasons that when considering the platform for model build, functionality requirements should be considered alongside potential model building challenges, requirements for review and audience familiarity, and subsequent adaptability to ensure overall project efficiency.

3.4. Comparative usability

The usability of a cost-effectiveness model in R is dependent on the user's familiarity with R language. In this application, the *Shiny* package was used to create a user-interface to allow multi-stakeholder user interaction with underlying R functionality without being exposed to the script. This combination of user-friendly interactivity with powerful R functionality ensures that users can view and amend complex analyses where previously there was little accessibility. The propensity score matching analysis within intRface is a good example of this, where users can select covariates, instantaneously view the results of the matching analysis and incorporate these into cost-effectiveness calculations. Whilst useful, such accessibility should be treated with caution to ensure that the functionality options available cannot be abused by users to push unfounded or biased assumptions. Consideration of audience knowledge, statistical validation and anticipated adaptability requirements should be balanced in informing available options.

Both the Excel and intRface model presented value input tables and dropdown options in clear model 'pages'. The only significant difference to user experience from this perspective is the model reactivity. The Excel model reactivity is such that amending values automatically updates the model. Within intRface, updating and running the model is triggered by buttons. Because of the ready availability of 'update' buttons within intRface, this was not found to be detrimental to user-experience. An advantage of R that has been included within intRface is the ability to produce interactive graphics that are responsive to hovering and convenient graph display options that inform users of relevant amendable options (Fig. 2).

Fig. 2: User-interactive outcome display options available within interface



A) Survival analysis graph display with hover functionality, user-amendable axis slider, curve selection options and AIC/BIC tables within 'Efficacy analysis: Select curves' page; B) Propensity

score matching results graph, display options, and table presenting hazard ratios for matched and unmatched datasets within ‘Survival data: Propensity score matching’ page of intRface.

AIC, Akaike information criterion; BIC, Bayesian information criterion; CAR, chimeric antigen receptor; KM, Kaplan–Meier; PFS, progression-free survival.

3.5. Comparative adaptability

An essential consideration for any model is that it has the required level of flexibility to react to anticipated adaptation. Modelling for HTA often involves adaptation of existing models to different localities and amendment to include updated data. Sufficient flexibility to amend base case modelling options and costs and to update trial data-cuts may be required. We therefore consider the suitability of both models in light of common adaptation requirements. We considered hypothetical adaptation scenarios in Table 2. The main difference between the two models is in the derivation and use of statistical analysis. In the intRface model, the changes to the raw data or matching analysis automatically feed through to the cost-effectiveness model. However, because the Excel model is dependent on externally derived estimates, updating these with a new analysis or data-cut would involve extensive copying and pasting of all results for each of the modelling strategies, significantly increasing the potential for error. The intRface model limits the potential for copy-and-paste error because the functionality to pass analysis outputs to model inputs already exists.

Table 2: Hypothetical model adaptation scenarios

Adaptation	intRface model	Excel model
Update to data cut	<ul style="list-style-type: none"> • Ensure that the new data are in the same format as previous data file • Link the file containing the data to the correct switch option • Sense check to ensure data are feeding through 	<ul style="list-style-type: none"> • Ensure that the new data are in the same format as previous data file • Re-run to produce new results (accounting for propensity score matching if required) in R and then extract these covariates • Copy and paste covariates for PartSA (including MCM) and StateTM into the model • QC check required to ensure correct curves are being produced
Updating propensity score matching analysis	<ul style="list-style-type: none"> • Select new covariate selection, calliper, and data-cut • Select button to update analysis 	<ul style="list-style-type: none"> • Re-run analysis and extract new covariates from R • Copy and paste covariates for PartSA (including MCM) and StateTM into the model • QC check required to ensure correct curves are being produced
Adaptation of inputs for local affiliates	<p>Costs:</p> <ul style="list-style-type: none"> • Fill in costs sheet by sending the Excel template to affiliate contact • Link the Excel sheet to the country switch in the model 	<p>Costs:</p> <ul style="list-style-type: none"> • Fill in costs sheet by sending the Excel template to affiliate contact • Copy and paste values into the ranges required

Adaptation	intRface model	Excel model
	Model options: <ul style="list-style-type: none"> • Overwrite current switch values 	Model options: <ul style="list-style-type: none"> • Overwrite current switch values
Reporting changes	A report template is written in RMD where the figures and tables reference the active model settings and outcomes. A button is then selected from within intRface and produces a Microsoft Word® document with required figures, tables, styles, headings and bibliography. For any subsequent change in results, a new Word document can be produced with updated tables and figures automatically incorporated. Required changes to the text for interpretation will need to be made manually in Word or RMD	A report is written in Microsoft Word where the figures and tables are copied from Excel. For any subsequent change in results, the figures and tables are re-copied from Excel. This increases the chance of error and will require a QC cross-check. Required changes to the text for interpretation will need to be made manually in Word
New functionality (e.g. addition of new health state or tunnel states)	Case-by-case consideration required. If previous R functionality exists in a modular format, then time and effort to incorporate functionality will most likely be faster in R or at worst similar between the two software options. This is particularly the case for implementation of tunnel states, which is not a short job in Excel.	
Removing functionality	Case-by-case consideration required. Anticipated to be similar time and effort required in most instances.	
Key: MCM, mixture-cure modelling; PartSA, partitioned survival analysis; QC, quality control; RMD, R Markdown; StateTM, state transition model.		

4. Discussion

We have provided an illustrative example of the differences in the complexity, functionality and adaptability of a similar cost-effectiveness model build in Excel and R. When choosing the software for an economic model, we advise consideration of the lifetime purpose, audience and technical requirements of the analysis, particularly as alternatives to Excel as a cost-effectiveness modelling platform become more widely accepted by HTA bodies. The hypothetical case study selected here presented modelling issues and analyses that are not universal across health economic decision problems; however, when they do arise, awareness of the relative merits of different software platforms for cost-effectiveness modelling may be valuable.

In general, model developers and reviewers are likely to remain more comfortable constructing and reviewing models that are programmed in Excel. When an analysis can be programmed in Excel and remain ‘fit for purpose’ in terms of computational efficiency and analytic transparency, there seems little benefit in moving to an alternative approach. Moving forward, increasing acceptance by external reviewers will be driven by increased training as part of health economics courses and using accepted and validated code for as many functionalities as feasible, with new analyses being well highlighted and documented [58].

When would we anticipate that programming in R will lead to substantive improvements in efficiency and transparency? In which contexts are the benefits likely to outweigh the costs to programmers and

reviewers of working in a less familiar platform? It is anticipated that such economic evaluations will meet the first of the following criteria, and are likely to meet one or more of criteria 2–4:

1. The intended audience (e.g. the company undertaking early evaluation of product potential or the HTA body) accepts economic evaluations in R. Furthermore, there is confidence that any reviewers of central importance have the capacity and inclination to review an evaluation undertaken in R. There is substantial heterogeneity in the willingness of different HTA bodies to review analyses that are more analytically demanding than cohort-based Excel models [7]. NICE accepts models written in R [59]; the use of R can be accepted by other HTA bodies but typically requires pre-approval [60-62].
2. A number of complex statistical analytical methods are required to generate the inputs that drive patient transitions within the economic model, such that there is a substantial transparency benefit to being able to review how these statistical analyses and the resulting uncertainty are undertaken and incorporated into the economic evaluation.
3. There are likely to be late and/or multiple changes to either the underlying data informing the economic evaluation or the approach taken to statistically analyse these data, and these are likely to be required with quick turnaround.
4. The core economic model would have a long run-time if undertaken in Excel *front-end*. This includes analyses such as patient-level simulations. However, if a simulation is coded in VBA, comparable reviewer-transparency issues may be faced across R and Excel, such that a key benefit of using Excel is lost even if the speed is similar [63].

Widespread adoption of R with or without a user-interface would bring benefits to the field of health economic modelling, namely transparent, reproducible and adaptable models that have reusable functionality and are more capable of quantifying decision uncertainty [7]. If the trend to undertake HTA in parallel to the regulatory process becomes more widespread, the third stated criterion (the need to update underlying data) will become increasingly relevant. If regulatory and HTA evaluations are undertaken earlier in the drug development process, when clinical effectiveness data are less mature and there is an increased reliance on complex statistical methods to estimate the ‘true’ benefit and uncertainty surrounding this for an intervention, the likelihood of the fourth stated criterion increases. Of course, if a greater number of models are constructed using R, the technical barrier to programmers and reviewers will continue to fall, as a result of both accumulated experience and the standardization of methods, functions and frameworks [15]. The capacity to readily incorporate standardized functions means it is not difficult to envision a future in which R-based models are considered easier to appraise than Excel-based models.

In this study, we have considered R versus Excel for cost-effectiveness models with the design specified to HTA requirements. Earlier in the drug development process, the comparative advantages of R versus Excel to inform go/no-go decision-making are related to those exposed here, but worthy of further consideration. R may also have a growing role in model validation and quality control of Excel models [64].

R is not a silver bullet for the wider issues faced by the HTA community. The tension between accelerated access and a desire for robust HTA decision making will remain. At a more mundane level, analysts will always need to follow conventions of good practice when annotating their code and clearly report modelling methods, assumptions and implementation decisions. An additional consideration is the alteration and updating of R packages, which would require models to be regularly reviewed and maintained to ensure update compatibility if distributed via R and not a controlled server. However, by providing a user-friendly platform that allows all crucial analyses to be undertaken in one location, R models can serve analysts and the wider community in generating analyses that are more efficient, less error prone and more transparent than would be the case if

R and *Shiny* for cost-effectiveness analyses: why and when? A hypothetical case study

undertaken in Excel. It is hoped our efforts here can help others in the field make informed software selection decisions for HTA cost-effectiveness models, on a case-by-case basis.

Data Availability Statement

The datasets generated and analysed during the current study are available from the corresponding author on reasonable request. The Microsoft Excel[®] model presented in this publication is included in the supplementary material files. The R code for the intRface[™] model produced for this study is not publicly available due to commercial interest but is available from the corresponding author on reasonable request. A demo version of the front-end of the model can be found here: https://bresmed-intRface-hypothetical-car-t-model.shinyapps.io/IntRface_Model-PharmacoEconomics/

Compliance with Ethical Standards

Author contributions: Authors RH, WS, DG, ND and DL were involved with the conception, planning and writing of this manuscript. Authors RH, DB and ND programmed the model demonstrated within this manuscript. Authors OS, IS and TC were involved in programming methods and functions adapted for use within the model. Authors SR, BR and DL were involved in quality control and validation of methods used in the model demonstrated within this manuscript. All authors edited and commented on drafts and approved the submitted versions.

Funding: No external funding was received for the development of the model or manuscript.

Conflict of Interest: All of the authors declare that they have no conflict of interest relevant to the content of this manuscript

Supplementary Material

Online Resource 1: Excel model

Online Resource 2: Consolidated Health Economic Evaluation Reporting Standards (CHEERS) checklist

References

1. Microsoft R Application Network. Introduction to R. 2019. <https://mran.microsoft.com/documents/what-is-r>. Accessed 6 Feb 2019.
2. R Core Team. R: A language and environment for statistical computing. In: R Foundation for Statistical Computing. Vienna, Austria 2018. <https://www.R-project.org/>. Accessed 10 May 2019.
3. Woolacott N, Hawkins N, Mason A, Kainth A, Khadjesari Z. Etanercept and efalizumab for the treatment of psoriasis: a systematic review. *Health Technol Assess*. 2006;10(46).
4. National Institute for Health and Care Excellence (NICE). Rituximab for the first-line treatment of stage III-IV follicular lymphoma (TA243). 2012. <https://www.nice.org.uk/guidance/ta243>. Accessed 11 June 2019.
5. Institute for Clinical and Economic Review (ICER). Extended-Release Opioid Agonists and Antagonist Medications for Addiction Treatment (MAT) in Patients with Opioid Use Disorder: Effectiveness and Value. 2018. https://icer-review.org/wp-content/uploads/2018/04/ICER_OUD_Final_Evidence_Report_120318.pdf. Accessed 11 June 2019.
6. Baio G. R for trial and model-based cost-effectiveness analysis. 2019. <http://www.statistica.it/gianluca/teaching/r-hta-workshop/>. Accessed 30 Jan 2019.
7. Incerti D, Thom H, Baio G, Jansen JP. R You Still Using Excel? The Advantages of Modern Software Tools for Health Technology Assessment. *Value Health*. 2019;22(5):575-9. doi:10.1016/j.jval.2019.01.003.
8. Krijkamp EM, Alarid-Escudero F, Enns EA, Jalal HJ, Hunink MGM, Pechlivanoglou P. Microsimulation Modeling for Health Decision Sciences Using R: A Tutorial. *Med Decis Making*. 2018;38(3):400-22. doi:10.1177/0272989X18754513.
9. Hollman C, Paulden M, Pechlivanoglou P, McCabe C. A Comparison of Four Software Programs for Implementing Decision Analytic Cost-Effectiveness Models. *Pharmacoeconomics*. 2017;35(8):817-30. doi:10.1007/s40273-017-0510-8.
10. Williams C, Lewsey JD, Briggs AH, Mackay DF. Cost-effectiveness Analysis in R Using a Multi-state Modeling Survival Analysis Framework: A Tutorial. *Med Decis Making*. 2017;37(4):340-52. doi:10.1177/0272989X16651869.
11. Jalal H, Pechlivanoglou P, Krijkamp E, Alarid-Escudero F, Enns E, Hunink MGM. An Overview of R in Health Decision Sciences. *Med Decis Making*. 2017;37(7):735-46. doi:10.1177/0272989X16686559.
12. Baio G, Heath A. When Simple Becomes Complicated: Why Excel Should Lose its Place at the Top Table. *Global & Regional Health Technology Assessment*. 2017.
13. Sampson CJ, Arnold R, Bryan S, Clarke P, Ekins S, Hatswell A et al. Transparency in Decision Modelling: What, Why, Who and How? *Pharmacoeconomics*. 2019. doi:10.1007/s40273-019-00819-z.
14. Incerti D, Curtis JR, Shafrin J, Lakdawalla DN, Jansen JP. A Flexible Open-Source Decision Model for Value Assessment of Biologic Treatment for Rheumatoid Arthritis. *Pharmacoeconomics*. 2019;37(6):829-43. doi:10.1007/s40273-018-00765-2.
15. Alarid-Escudero F, Krijkamp EM, Pechlivanoglou P, Jalal H, Kao SZ, Yang A et al. A Need for Change! A Coding Framework for Improving Transparency in Decision Modeling. *Pharmacoeconomics*. 2019. doi:10.1007/s40273-019-00837-x.
16. Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB et al. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--7. *Value Health*. 2012;15(6):843-50. doi:10.1016/j.jval.2012.04.012.
17. Tappenden P, Caro JJ. Improving Transparency in Decision Models: Current Issues and Potential Solutions. *Pharmacoeconomics*. 2019;37(11):1303-4. doi:10.1007/s40273-019-00850-0.
18. Allaire J, Xie Y, McPherson J, Luraschi J, Ushey K, Atkins A et al. *rmarkdown: Dynamic Documents for R*. R package version 1.13. 2019. <https://rmarkdown.rstudio.com>. Accessed 3 Dec 2019.
19. RStudio. Version Control with Git and SVN. 2019. <https://support.rstudio.com/hc/en-us/articles/200532077-Version-Control-with-Git-and-SVN>. Accessed 3 Dec 2019.
20. National Health Care Institute Zorginstituut Nederland (ZIN). Template Pharmacoeconomic dossier. 2016.

- <https://www.zorginstituutnederland.nl/binaries/zinl/documenten/publicatie/2016/07/25/template-farmaco-economisch-dossier/Template+Pharmacoeconomic+dossier.doc>. Accessed 12 Nov 2019.
21. Pharmaceutical Management Agency (PHARMAC). Prescription for Pharmacoeconomic Analysis - Methods for Cost-Utility Analysis. 2015. <https://www.pharmac.govt.nz/assets/pfpa-2-2.pdf>. Accessed 12 Nov 2019.
22. Canadian Agency for Drugs and Technologies in Health (CADTH). Procedure and Submission Guidelines for the CADTH Common Drug Review. 2019. https://cadth.ca/sites/default/files/cdr/process/Procedure_and_Guidelines_for_CADTH_CDR.pdf. Accessed 12 Nov 2019.
23. HealthEconomics.com. ICER Using New Economic Modeling Platform for Value Assessment. 2018. <https://www.healtheconomics.com/industry-news/icer-using-new-economic-modeling-platform-for-value-assessment>. Accessed 12 Nov 2019.
24. RStudio. Shiny. 2019. <https://shiny.rstudio.com/>. Accessed 2 Dec 2019.
25. Chang W, Cheng J, Allaire J, Xie Y, McPherson J. shiny: Web Application Framework for R. R package version 1.3.2. 2019. <https://CRAN.R-project.org/package=shiny>. Accessed 2 July 2019.
26. Innovation and Value Initiative. Open-Source Value Platform. 2019. <https://www.thevalueinitiative.org/open-source-value-project/>. Accessed 6 Feb 2019.
27. Food and Drug Administration (FDA). RR-Drug. 2019. https://openfda.shinyapps.io/RR_D/. Accessed 12 Nov 2019.
28. (CHOICE). TUoW-TCHOPaEI. PriMER Platform: Diffusion of Precision Medicine Technologies. 2019. <https://uwchoice.shinyapps.io/primer/>. Accessed 12 Nov 2019.
29. Center for Evaluation of Value and Risk in Health (CEVR). Global Health Cost Effectiveness Analysis Registry. 2019. <http://healtheconomics.tuftsmedicalcenter.org/ghcearegistry/>. Accessed 12 Nov 2019.
30. Eichler HG, Oye K, Baird LG, Abadie E, Brown J, Drum CL et al. Adaptive licensing: taking the next step in the evolution of drug approval. *Clin Pharmacol Ther.* 2012;91(3):426-37. doi:10.1038/clpt.2011.345.
31. Food and Drug Administration (FDA). Accelerated Approval Program. 2018. <https://www.fda.gov/drugs/resourcesforyou/healthprofessionals/ucm313768.htm>. Accessed 2 July 2019.
32. European Medicines Agency. Conditional Marketing Authorisation: Report on Ten Years of Experience at the European Medicines Agency. 2016. https://www.ema.europa.eu/en/documents/report/conditional-marketing-authorisation-report-ten-years-experience-european-medicines-agency_en.pdf. Accessed 14 May 2019.
33. Hatswell AJ, Baio G, Berlin JA, Irs A, Freemantle N. Regulatory approval of pharmaceuticals without a randomised controlled study: analysis of EMA and FDA approvals 1999–2014. *BMJ Open.* 2016;6(6):e011666. doi:10.1136/bmjopen-2016-011666.
34. European Medicines Agency. Yescarta. 2018. <https://www.ema.europa.eu/en/medicines/human/EPAR/yescarta>. Accessed 9 Apr 2019.
35. European Medicines Agency. Kymriah. 2018. <https://www.ema.europa.eu/en/medicines/human/EPAR/kymriah>. Accessed 9 Apr 2019.
36. Food and Drug Administration (FDA). Yescarta (axicabtagene ciloleucel). 2018. <https://www.fda.gov/vaccines-blood-biologics/cellular-gene-therapy-products/yescarta-axicabtagene-ciloleucel>. Accessed 7 June 2019.
37. Food and Drug Administration (FDA). Kymriah (tisagenlecleucel). 2019. <https://www.fda.gov/vaccines-blood-biologics/cellular-gene-therapy-products/kymriah-tisagenlecleucel>. Accessed 6 June 2019.
38. Hettle R, Corbett M, Hinde S, Hodgson R, Jones-Diette J, Woolacott N et al. Exploring the assessment and appraisal of regenerative medicines and cell therapy products. 2015. <https://www.nice.org.uk/Media/Default/About/what-we-do/Science%20policy%20and%20research/final-york-report-march-16.pdf>. Accessed 2 July 2019.
39. Picanco-Castro V, Goncalves Pereira C, Swiech K, Ribeiro Malmegrim KC, Tadeu Covas D, Silveira Porto G. Emerging CAR T cell therapies: clinical landscape and patent technological routes. *Hum Vaccin Immunother.* 2019;1-10. doi:10.1080/21645515.2019.1689744.

40. Abreu TR, Fonseca NA, Goncalves N, Moreira JN. Current challenges and emerging opportunities of CAR-T cell therapies. *J Control Release*. 2020;319:246-61. doi:10.1016/j.jconrel.2019.12.047.
41. National Institute for Health and Care Excellence (NICE). Tisagenlecleucel for treating relapsed or refractory B-cell acute lymphoblastic leukaemia in people aged up to 25 years [TA554]. 2018. <https://www.nice.org.uk/guidance/ta554/documents/final-appraisal-determination-document>. Accessed 17 Apr 2019.
42. National Institute for Health and Care Excellence (NICE). Axicabtagene ciloleucel for treating diffuse large B-cell lymphoma and primary mediastinal large B-cell lymphoma after 2 or more systemic therapies [TA559]. 2019. <https://www.nice.org.uk/guidance/ta559/documents/final-appraisal-determination-document>. Accessed 23 Jan 2019.
43. National Institute for Health and Care Excellence (NICE). Tisagenlecleucel for treating relapsed or refractory diffuse large B-cell lymphoma after 2 or more systemic therapies (TA567). 2019. <https://www.nice.org.uk/guidance/ta567/documents/final-appraisal-determination-document>. Accessed 05 July 2019.
44. Woods B, Sideris E, Palmer S, Latimer N, Soares M. NICE DSU Technical Support Document 19. Partitioned Survival Analysis for Decision Modelling in Health Care: A Critical Review. 2017. <http://scharr.dept.shef.ac.uk/nicedsu/wp-content/uploads/sites/7/2017/06/Partitioned-Survival-Analysis-final-report.pdf>. Accessed 2 May 2019.
45. National Institute for Health and Care Excellence (NICE). Guide to the methods of technology appraisal. 2013. <https://www.nice.org.uk/guidance/pmg9/resources/guide-to-the-methods-of-technology-appraisal-2013-pdf-2007975843781>. Accessed 2 May 2019.
46. Sousa D, Ferreira F, Félix F, Lopesa M. Acute lymphoblastic leukemia in children and adolescents: prognostic factors and analysis of survival. *Rev Bras Hematol Hemoter*. 2015;37(4):223-9.
47. Fuster JL. Current approach to relapsed acute lymphoblastic leukemia in children. *World J Hematol*. 2014;3(3):49-70. doi:10.5315/wjh.v3.i3.49.
48. National Institute for Health and Care Excellence. Single Technology Appraisal Axicabtagene ciloleucel for treating diffuse large B-cell lymphoma and primary mediastinal B-cell lymphoma after 2 or more systemic therapies (TA559): Committee Papers. 2019. <https://www.nice.org.uk/guidance/ta559/evidence/appraisal-consultation-committee-papers-pdf-6661404973>. Accessed 11 Apr 2019.
49. Faria R, Hernandez Alava, M., Manca, A., Wailoo, A.J. NICE DSU Technical Support Document 17: The use of observational data to inform estimates of treatment effectiveness for Technology Appraisal: Methods for comparative individual patient data. 2015. <http://scharr.dept.shef.ac.uk/nicedsu/wp-content/uploads/sites/7/2016/03/TSD17-DSU-Observational-data-FINAL.pdf>. Accessed 10 May 2019.
50. Rosenbaum PR, Rubin DB. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*. 1983;70(1):41-55.
51. Stuart EA. Matching methods for causal inference: A review and a look forward. *Stat Sci*. 2010;25(1):1-21. doi:10.1214/09-STS313.
52. Latimer N. NICE DSU Technical Support Document 14: Undertaking survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data. 2011. <http://nicedsu.org.uk/wp-content/uploads/2016/03/NICE-DSU-TSD-Survival-analysis.updated-March-2013.v2.pdf>. Accessed 2 May 2019.
53. Lambert PC, Thompson JR, Weston CL, Dickman PW. Estimating and modeling the cure fraction in population-based cancer survival analysis. *Biostatistics*. 2007;8(3):576-94. doi:10.1093/biostatistics/kxl030.
54. Othus M, Barlogie B, Leblanc ML, Crowley JJ. Cure models as a useful statistical tool for analyzing survival. *Clin Cancer Res*. 2012;18(14):3731-6. doi:10.1158/1078-0432.CCR-11-2859.
55. Lambert PC. Modeling of the cure fraction in survival studies. *The Stata Journal*. 2007;7(3):351-75.
56. Grant TS, Burns D, Kiff C, Lee D. A Case Study Examining the Usefulness of Cure Modelling for the Prediction of Survival Based on Data Maturity. *Pharmacoeconomics*. 2019. doi:10.1007/s40273-019-00867-5.

57. RStudio. RStudio Connect. 2019. <https://rstudio.com/products/connect/>. Accessed 3 Dec 2019.
58. Ghabri S, Stevenson M, Moller J, Caro JJ. Trusting the Results of Model-Based Economic Analyses: Is there a Pragmatic Validation Solution? *Pharmacoeconomics*. 2019;37(1):1-6. doi:10.1007/s40273-018-0711-9.
59. National Institute for Health and Care Excellence (NICE). Guide to the processes of technology appraisal. 2014. <https://www.nice.org.uk/process/pmg19>. Accessed 11 June 2019.
60. Canadian Agency for Drugs and Technologies in Health (CADTH). CADTH Common Drug Review (CDR). 2019. <https://www.cadth.ca/about-cadth/what-we-do/products-services/cdr>. Accessed 11 June 2019.
61. National Centre for Pharmacoeconomics (NCPE). HTA Guidelines. 2019. <http://www.ncpe.ie/submission-process/hta-guidelines/>. Accessed 11 June 2019.
62. Pharmaceutical Benefits Advisory Committee (PBAC). The Pharmaceutical Benefits Advisory Committee Guidelines. 2016. <https://pbac.pbs.gov.au/information/printable-version-of-guidelines.html>. Accessed 11 June 2019.
63. Moller J, Davis S, Stevenson M, Caro JJ. Validation of a DICE Simulation Against a Discrete Event Simulation Implemented Entirely in Code. *Pharmacoeconomics*. 2017;35(10):1103-9. doi:10.1007/s40273-017-0534-0.
64. Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R et al. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technol Assess*. 2004;8(36):iii-iv, ix-xi, 1-158. doi:10.3310/hta8360.