Cure Models for Health Technology Assessment: Can They Be Trusted for Decision-Making?

ISPOR Europe 2024, Issue Panel 129, 18th November 2024

Moderator: Megan John, NHS and Chair of NICE Appraisal Committee D

Panelists: Federico Felizzi, Menarini GMBH
Nick Latimer, University of Sheffield, Delta Hat
Melanie Whittington, Tufts Medical Center

University of Sheffield

Delta Hat
Powered by Petauri™

Cure Models for Health Technology Assessment: Can They Be Trusted for Decision-Making?

The Case for Flexible Parametric Non-Mixture Cure Models

ISPOR Europe 2024, Issue Panel 129, 18th November 2024

Professor Nick Latimer
University of Sheffield and Delta Hat Ltd.
n.latimer@sheffield.ac.uk

A WORLD TOP 100 UNIVERSITY

Disclosures

- I was a member of a NICE Appraisal Committee for 5 years
- I'm a member of NICE's Decision Support Unit and am an author on technical support documents about survival analysis (TSD 16) and flexible survival models (TSD 21)
- I work part-time for Delta Hat Ltd, a consultancy company
- I learnt a lot about these flexible parametric non-mixture cure models while working on a project for BMS, and then learnt even more about them while writing a tutorial for *PharmacoEconomics* on cure models (with Mark Rutherford)[1]
  - This is what motivated me to propose this issue panel
- **These are my own opinions, not necessarily those of NICE, the DSU, Delta Hat, BMS, or Mark!**

1. Latimer NR, Rutherford MJ. Mixture and non-mixture cure models for health technology assessment: What you need to know. *PharmacoEconomics* (2024) 42:1073-1090.

Abbreviations: DSU, Decision Support Unit; NICE, National Institute for Health and Care Excellence; TSD, Technical support document

Frameworks for cure models

University of Sheffield

- What a cure model means, or represents, depends on the framework in which it is fitted
- In this talk, I am assuming that we are fitting cure models in a **relative survival framework**
- **This means that we model the difference between the hazard function observed in the trial, and the hazard function in the (age and sex-matched) general population**
- **Cure occurs when the all-cause hazard function for the modelled patient group converges with the general population hazard function: this indicates that the disease-specific hazard has fallen to zero**

Hazard function: the rate at which death occurs over time

*There are alternative frameworks for cure models (all-cause, disease specific – see extra slides at the end of the deck) but these are generally more problematic

[It's important to be clear about the framework being used – often when cure models are used in HTA this is not done!]

Mixture Cure Models (in a relative survival framework)

University of Sheffield

- **Key assumption:** there are two groups of individuals – cured and uncured. Cured patients are cured from the baseline time-point
- **What the model does:** We fit a parametric model to the trial data, and the point at which the modelled hazards are predicted to converge with the (age- and sex-matched) general population hazards dictates the cure fraction
- **Interpretation:** MCMs are primarily about the **cure fraction**. The model estimates this fraction and assumes it is present from the baseline time-point. Thus, 'cured' patients are cured from time zero. Their survival is dictated by general population hazards, and they never experience an excess risk of death. The parametric model represents survival in the uncured group

Abbreviations: MCM, Mixture Cure Model

Non-mixture Cure Models (in a relative survival framework)

University of Sheffield

- **Key assumption:** Does not split the population into cured and uncured groups, instead assuming that cure is apparent after a specific time-point
- **What the model does:** We fit a parametric model to the trial data, and the point at which the modelled hazards are predicted to converge with the (age- and sex-matched) general population hazards dictates the cure time-point
- **Interpretation:** NMCs are primarily about the **cure time-point**. Patients who reach the cure time-point are cured; survival beyond this point is determined by general population hazards. Before this time-point all patients are at an excess risk of death, and deaths from other causes can also occur

Abbreviations: NMC, Non-mixture Cure Model

Model options

- For both MCMs and NMCs we have some modelling options
- In particular, we need to choose the **parametric distribution** we are going to use
- **For MCMs, this distribution represents the uncured group of patients**
  - Typically standard parametric distributions are used (Weibull, log-normal, etc.)
  - Flexible parametric models are possible, but are seldom used
  - A model that is appropriate for the hazards expected in the uncured group should be chosen

Abbreviations: MCM, Mixture cure model; NMC, Non-mixture Cure Model

Model options

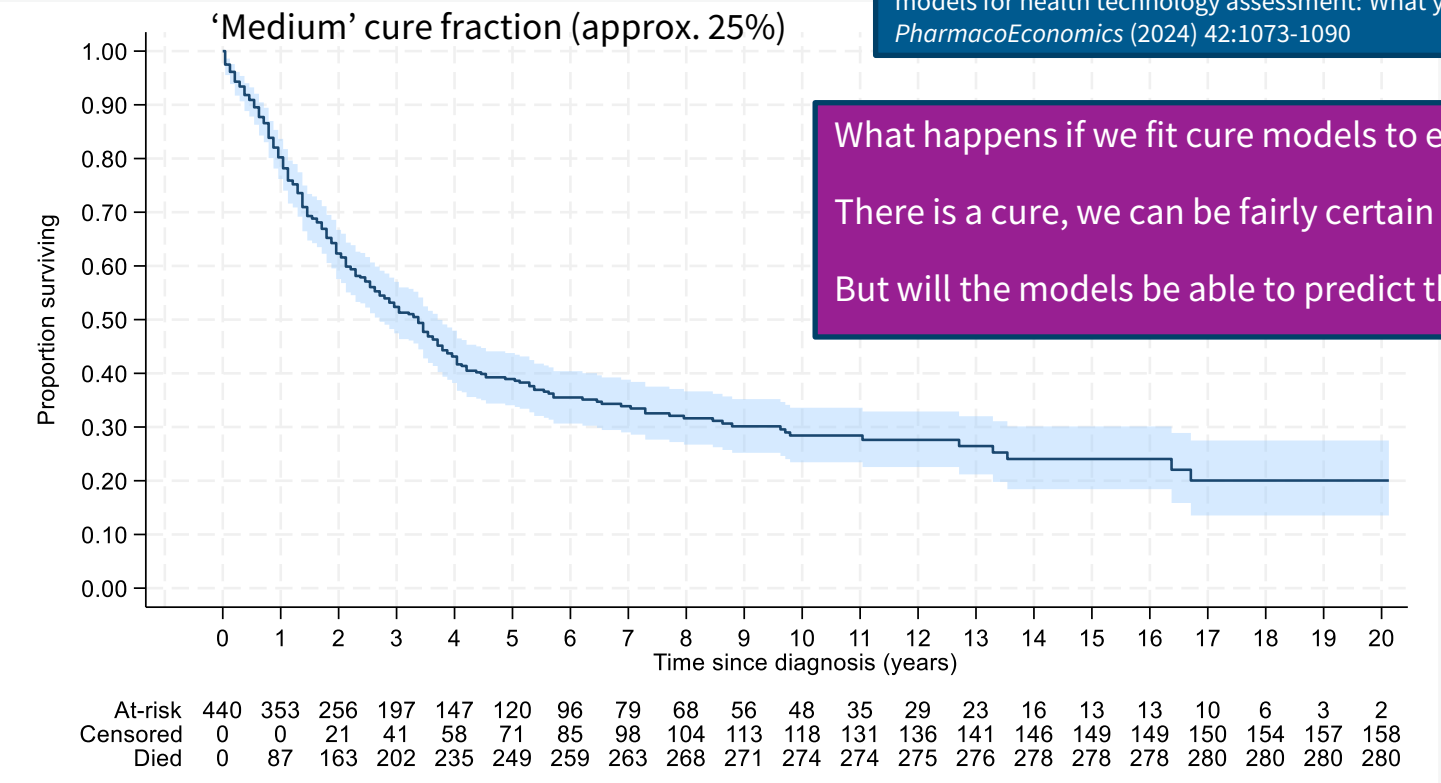
- For both MCMs and NMCs we have some modelling options
- In particular, we need to choose the **parametric distribution** we are going to use
- **For MCMs, this distribution represents the uncured group of patients**
  - Typically standard parametric distributions are used (Weibull, log-normal, etc.)
  - Flexible parametric models are possible, but are seldom used
  - A model that is appropriate for the hazards expected in the uncured group should be chosen
- **For NMCs, this distribution represents the cohort prior to the cure time-point**
  - Standard parametric distributions can be used (Weibull, log-normal, etc.)
  - But **flexible parametric NMCs** have also been developed[1]
  - **These provide analysts with an additional tool to control when the cure time-point will occur, by placement of a 'boundary knot'**
  - **The ability to 'control' the cure time-point is extremely helpful in the HTA context, where we often have relatively small RCTs with limited follow-up**

1. Andersson TML *et al.* Estimating and modelling cure in population-based cancer studies within the framework of flexible parametric survival models. *BMC Med Res Methodol*. 2011;11(1):96.

Abbreviations: HTA. Health technology assessment; MCM, Mixture cure model; NMC, Non-mixture Cure Model; RCT, randomised controlled trial

Demonstrating flexible parametric NMCs

Example derived from: Latimer NR, Rutherford MJ. Mixture and non-mixture cure models for health technology assessment: What you need to know. *PharmacoEconomics* (2024) 42:1073-1090



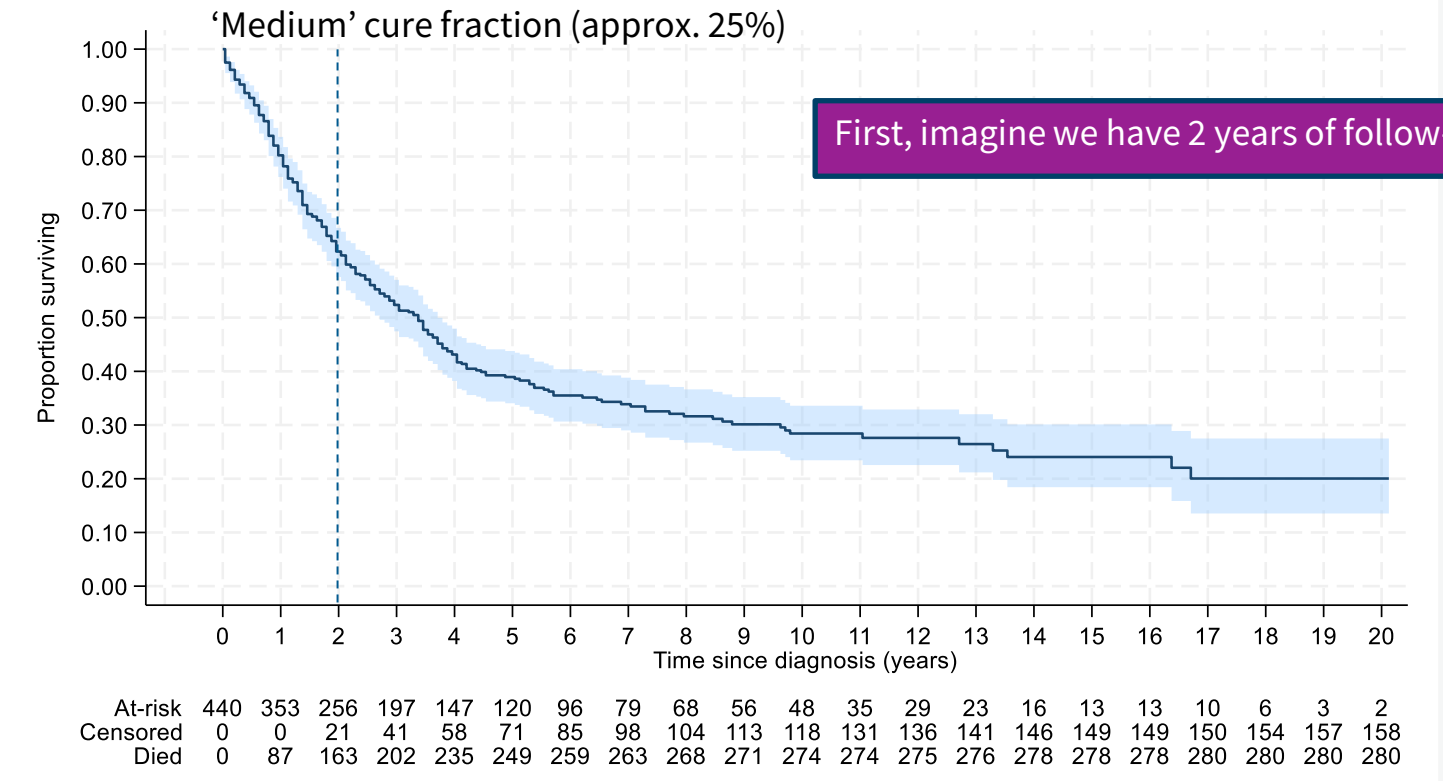
What happens if we fit cure models to early cuts of this data?

There is a cure, we can be fairly certain of that

But will the models be able to predict this accurately?

Abbreviations: NMC, Non-mixture Cure Model

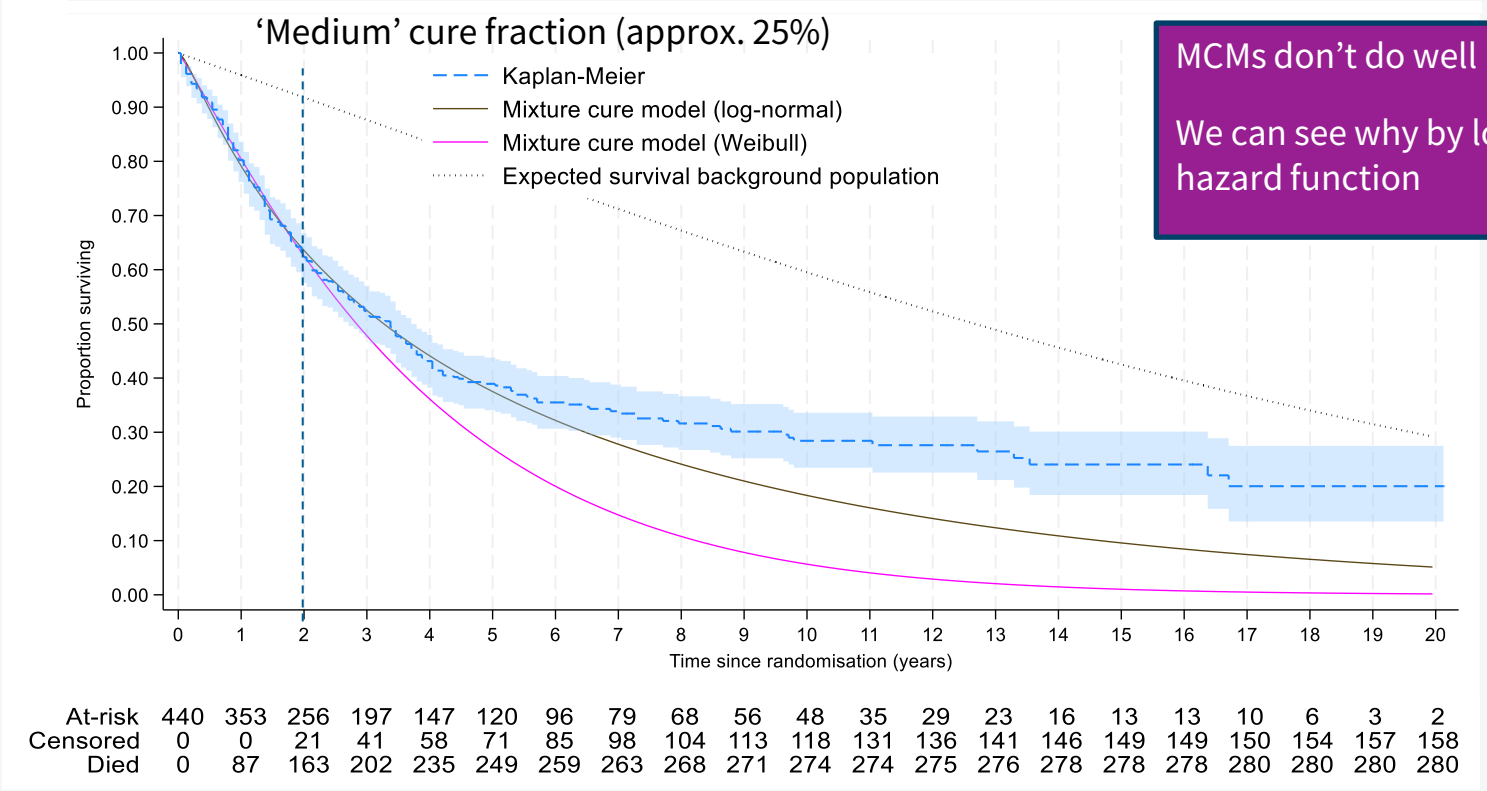
Demonstrating flexible parametric NMCs



First, imagine we have 2 years of follow-up

Abbreviations: NMC, Non-mixture Cure Model

Demonstrating flexible parametric NMCs

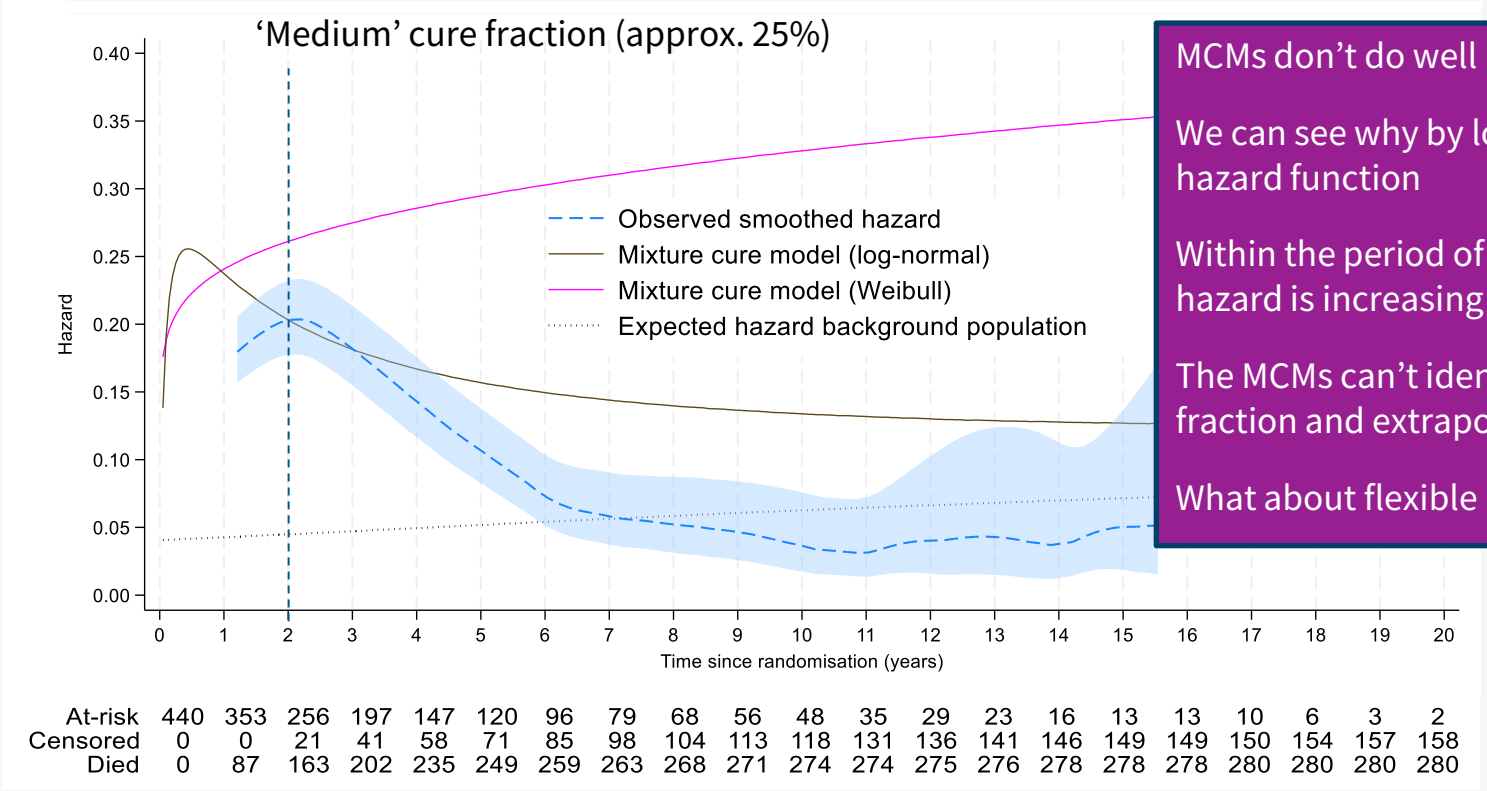


MCMs don’t do well

We can see why by looking at the hazard function

Abbreviations: MCM, Mixture cure model, NMC, Non-mixture cure model

Demonstrating flexible parametric NMCs



MCMs don’t do well

We can see why by looking at the hazard function

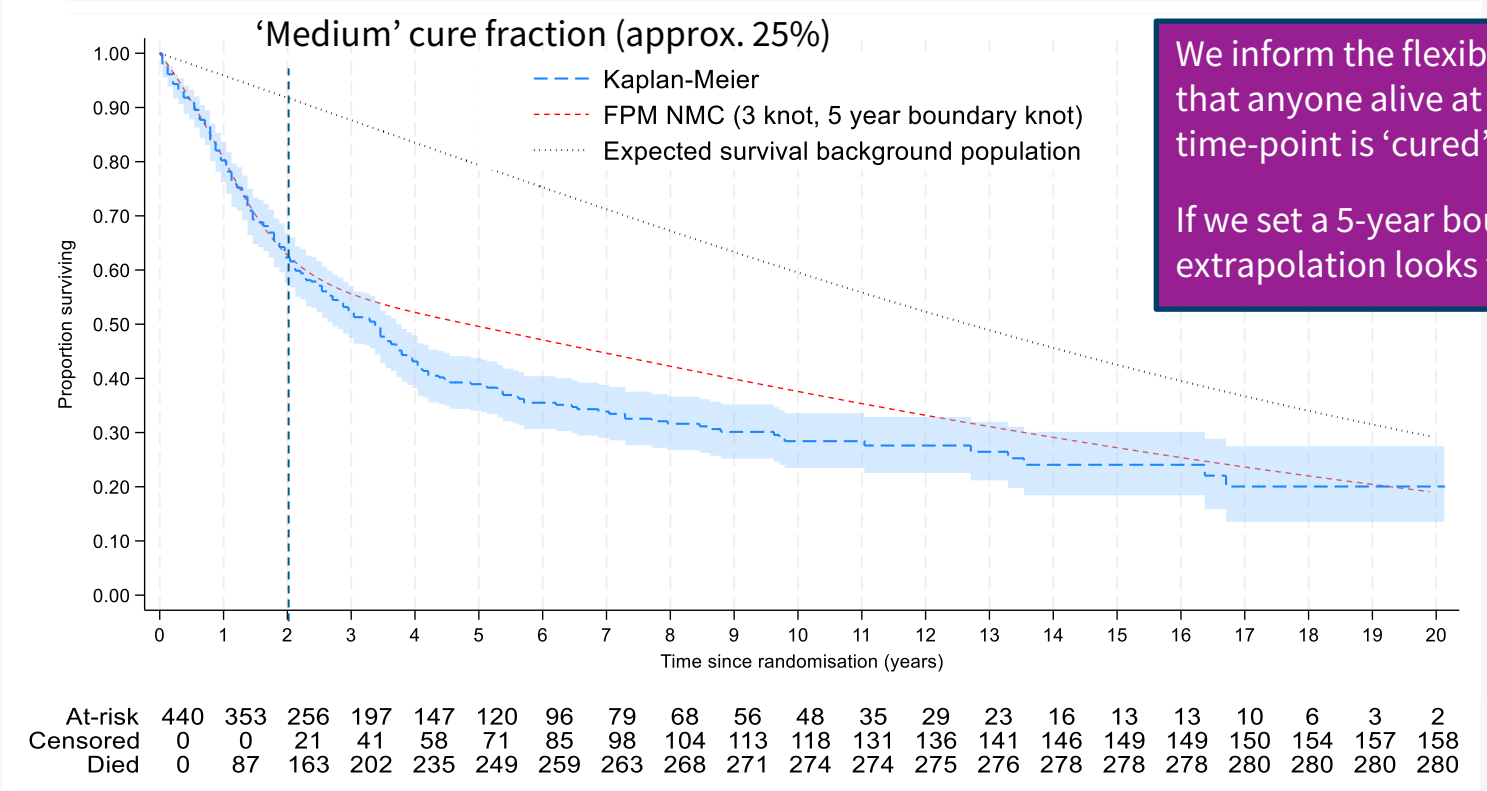
Within the period of the data, the hazard is increasing

The MCMs can’t identify the cure fraction and extrapolate poorly

What about flexible parametric NMCs?

Abbreviations: MCM, Mixture cure model; NMC, Non-mixture cure model

Demonstrating flexible parametric NMCs

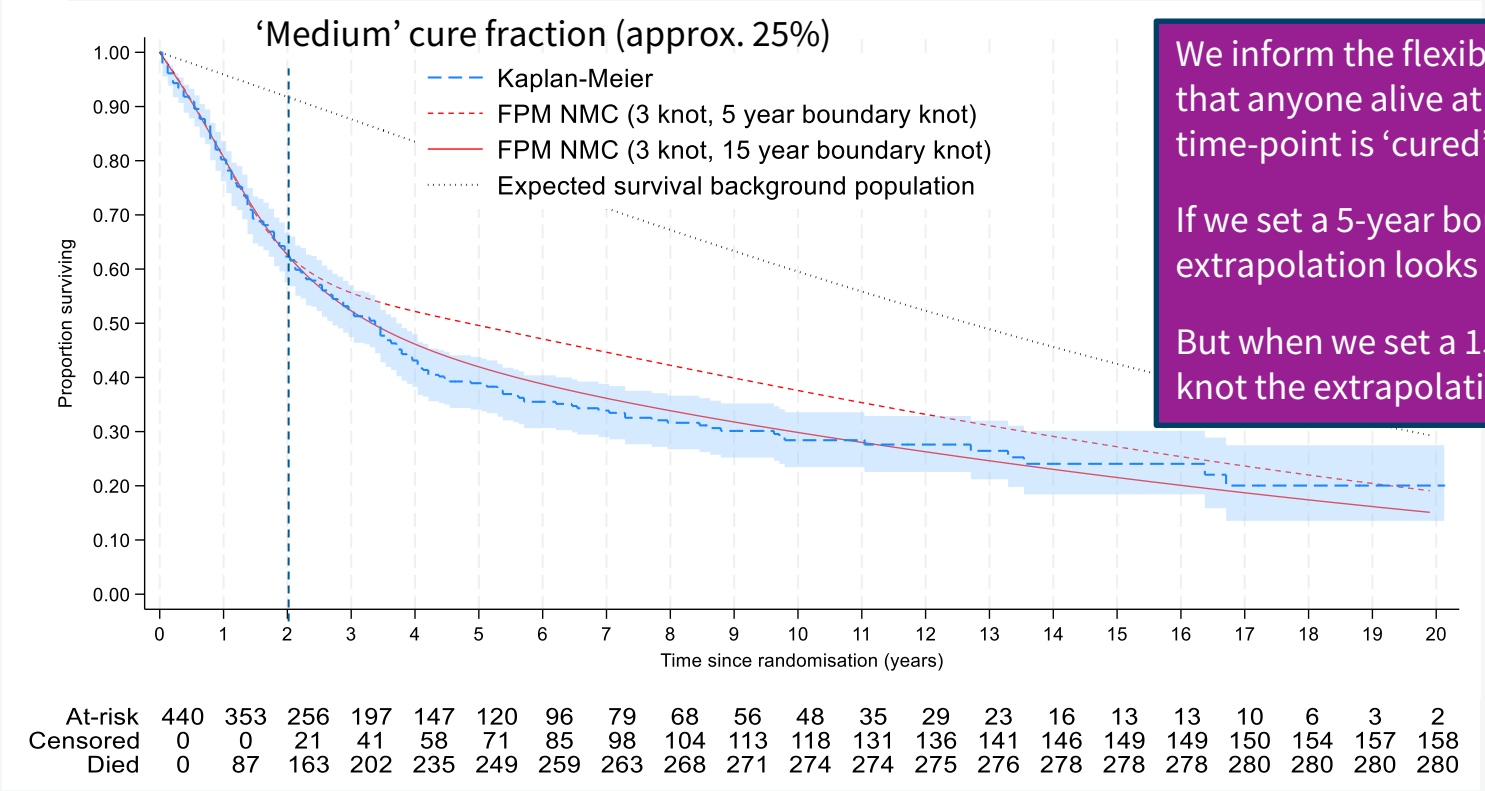


We inform the flexible parametric NMC that anyone alive at the boundary knot time-point is ‘cured’

If we set a 5-year boundary knot the extrapolation looks too optimistic

Abbreviations: FPM, Flexible parametric model; NMC, Non-mixture cure model

Demonstrating flexible parametric NMCs



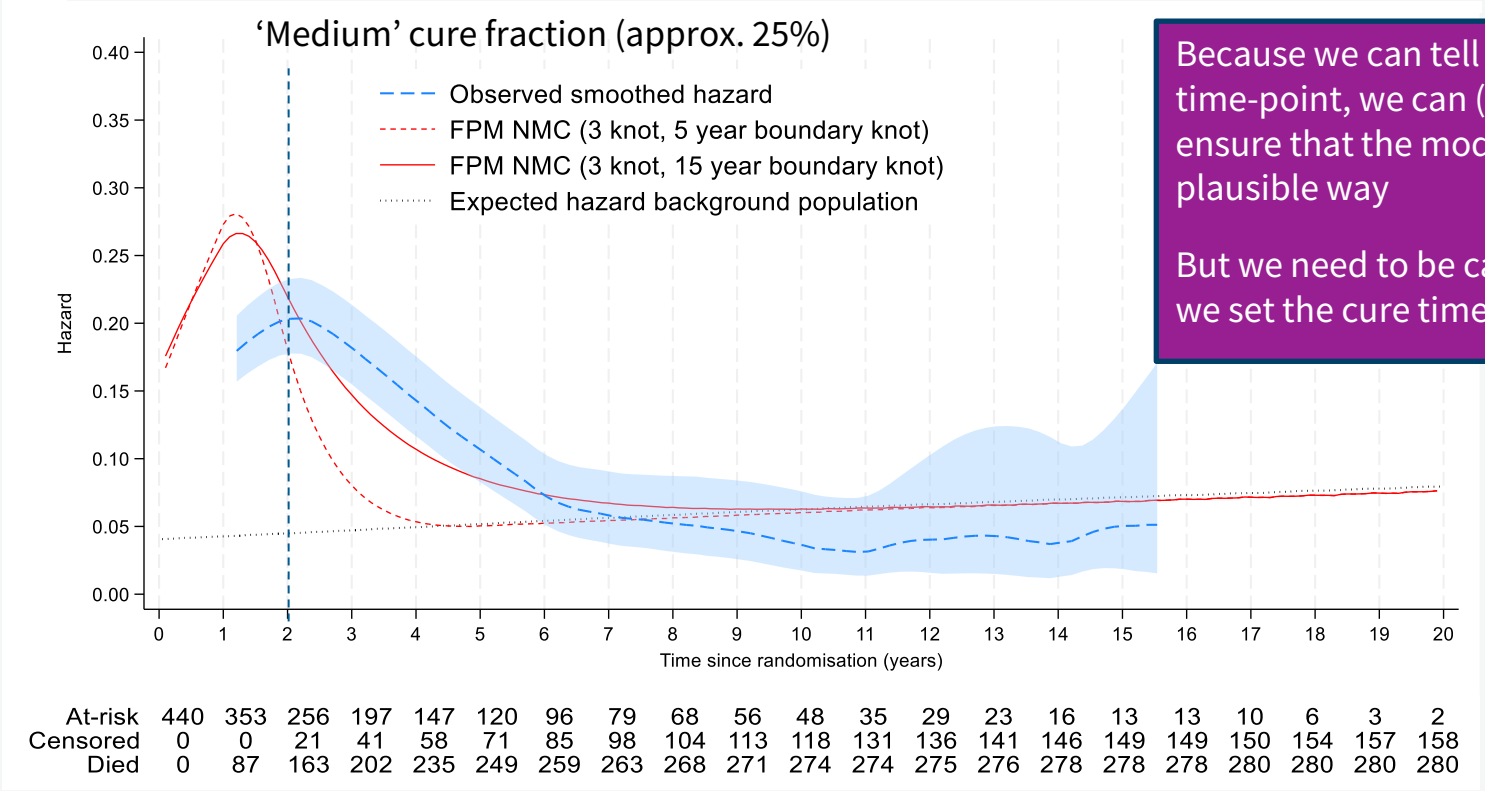
We inform the flexible parametric NMC that anyone alive at the boundary knot time-point is ‘cured’

If we set a 5-year boundary knot the extrapolation looks too optimistic

But when we set a 15-year boundary knot the extrapolation is better

Abbreviations: FPM, Flexible parametric model; NMC, Non-mixture cure model

Demonstrating flexible parametric NMCs

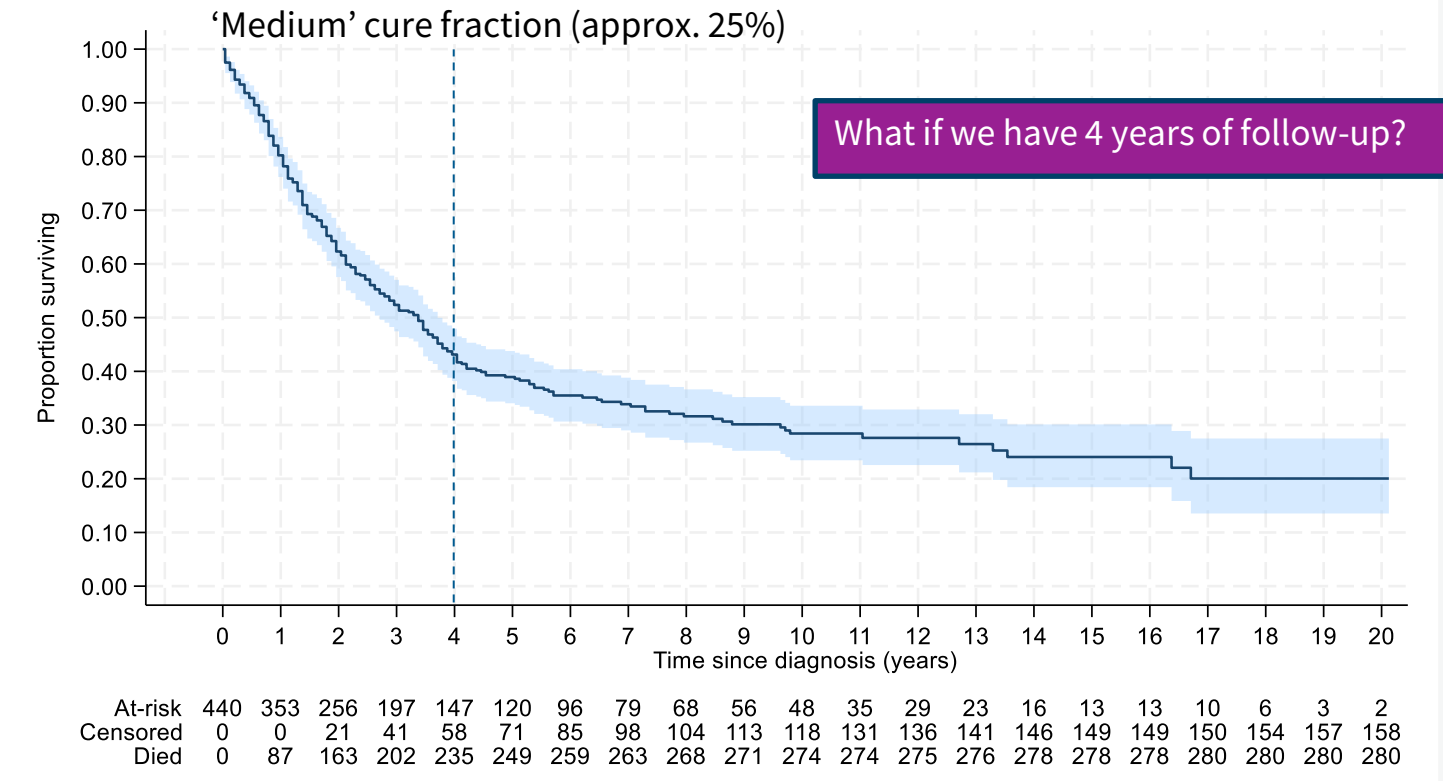


Because we can tell the model the cure time-point, we can (to some extent) ensure that the model extrapolates in a plausible way

But we need to be careful about where we set the cure time-point!

Abbreviations: FPM, Flexible parametric model; NMC, Non-mixture cure model

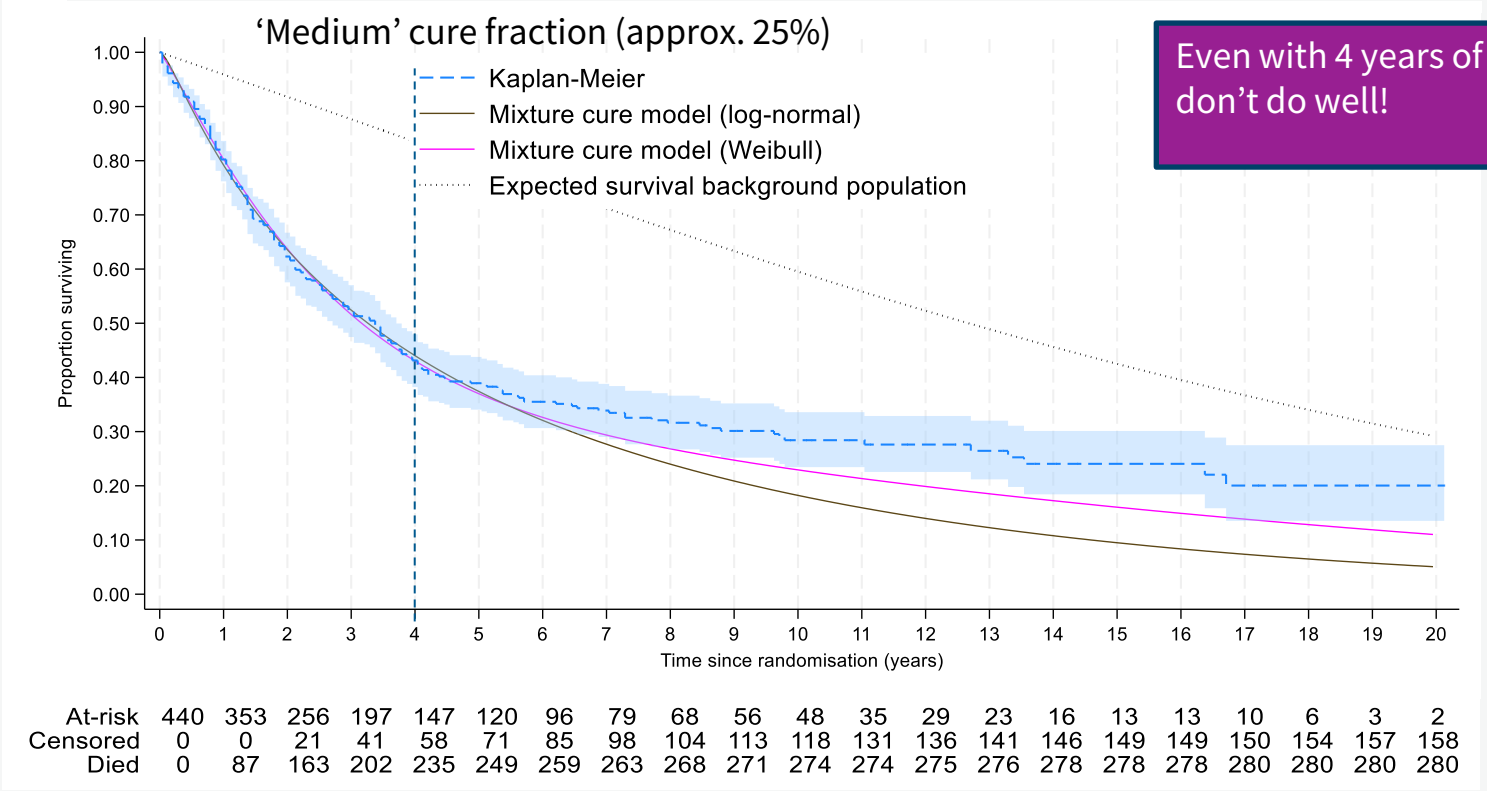
Demonstrating flexible parametric NMCs



What if we have 4 years of follow-up?

Abbreviations: NMC, Non-mixture Cure Model

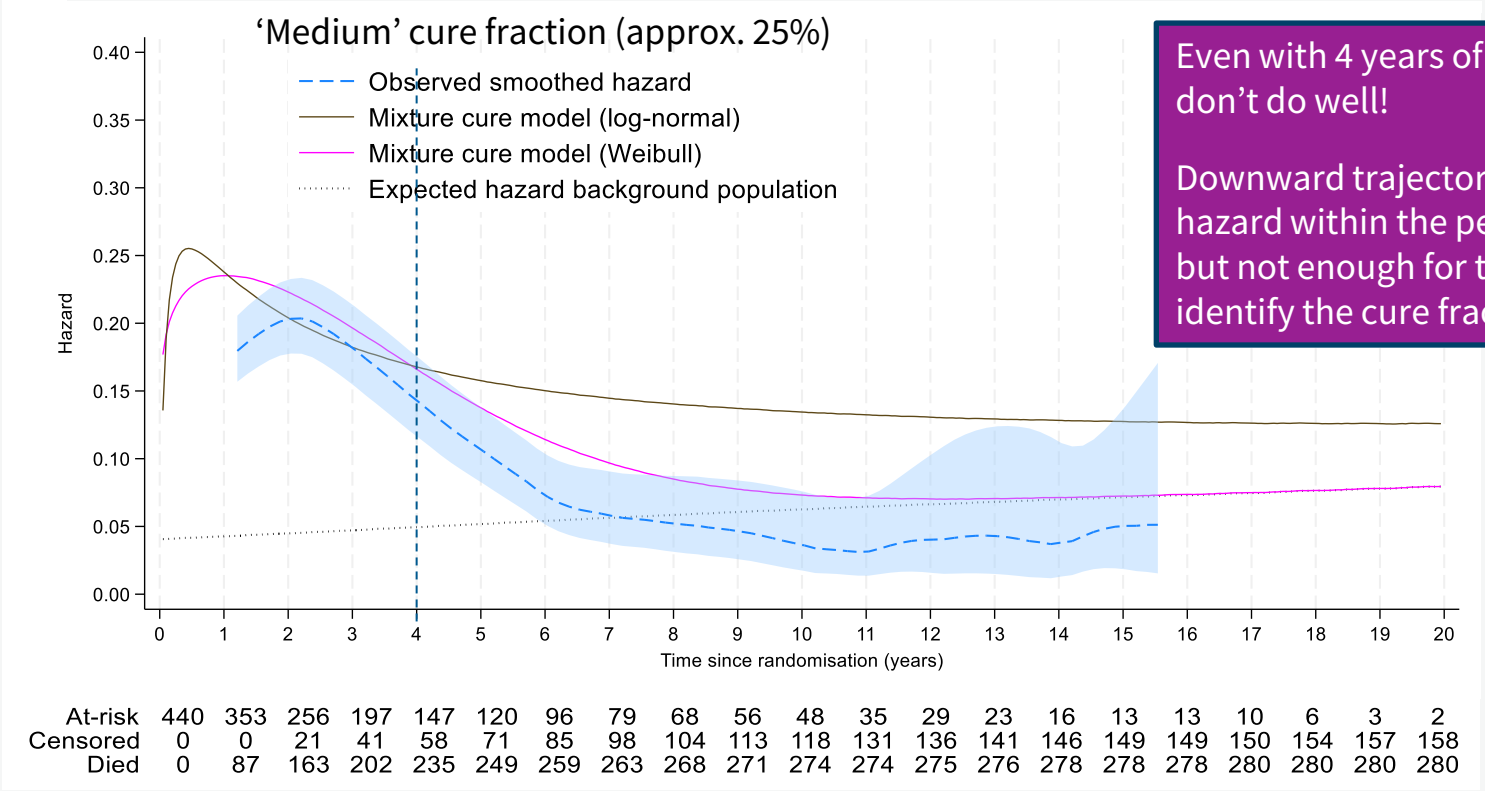
Demonstrating flexible parametric NMCs



Even with 4 years of data the MCMs don’t do well!

Abbreviations: MCM, Mixture cure model; NMC, Non-mixture cure model

Demonstrating flexible parametric NMCs

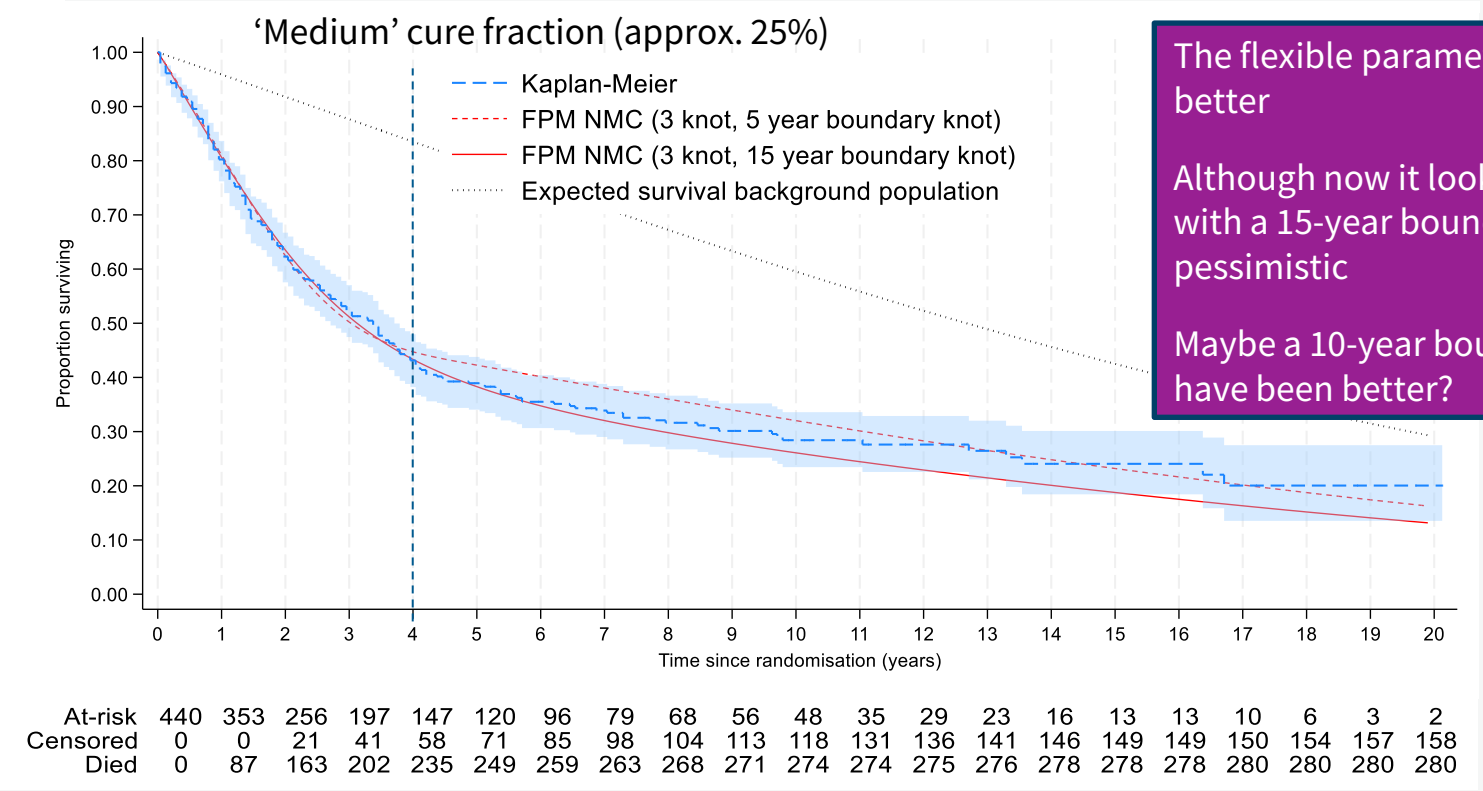


Even with 4 years of data the MCMs don't do well!

Downward trajectory is observed in the hazard within the period of the data, but not enough for the MCMs to identify the cure fraction accurately

Abbreviations: MCM, Mixture cure model; NMC, Non-mixture cure model

Demonstrating flexible parametric NMCs



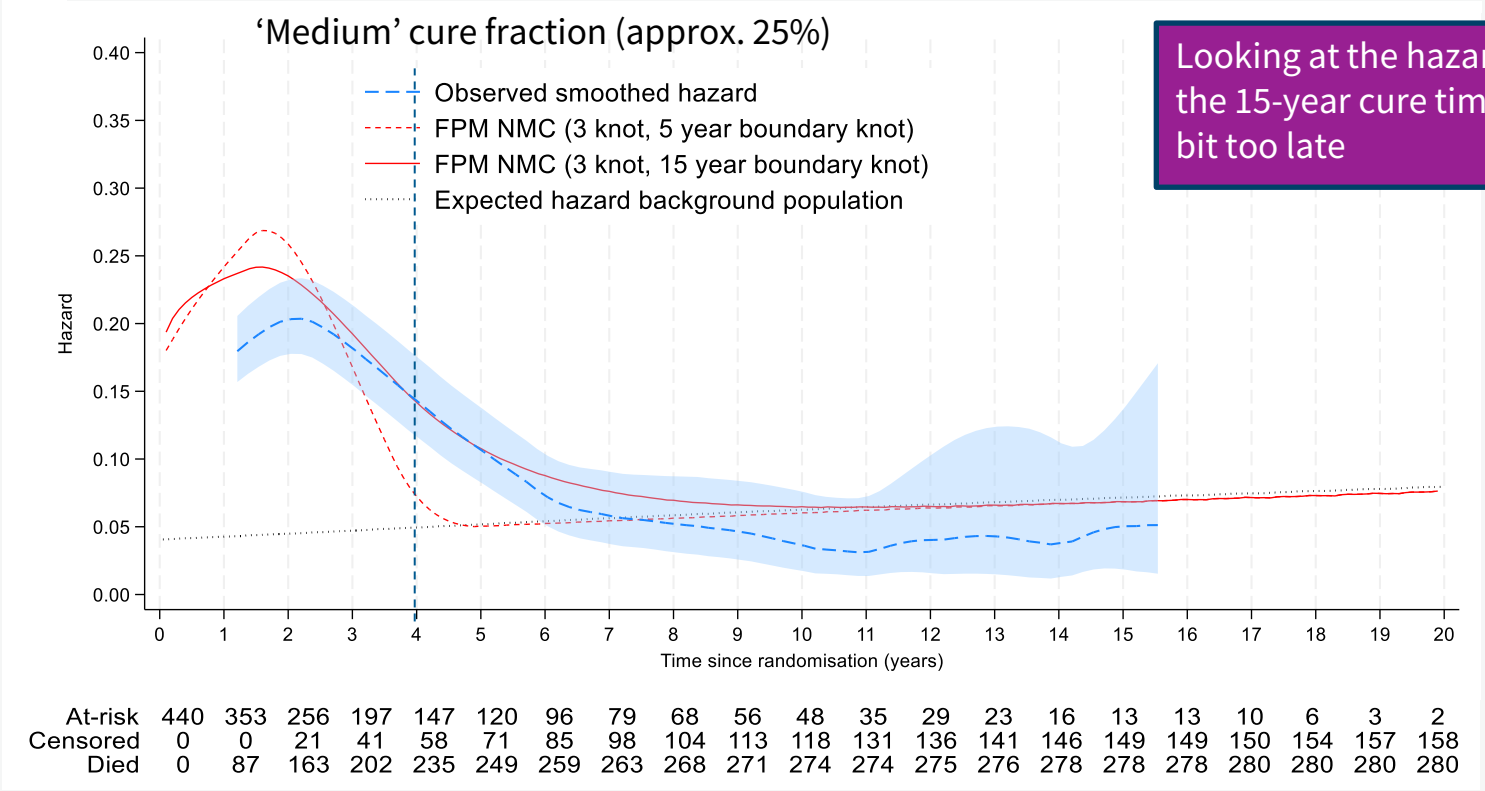
The flexible parametric NMCs again do better

Although now it looks like the model with a 15-year boundary knot may be pessimistic

Maybe a 10-year boundary knot would have been better?

Abbreviations: FPM, Flexible parametric model; NMC, Non-mixture cure model

Demonstrating flexible parametric NMCs



Looking at the hazards, it appears that the 15-year cure time-point might be a bit too late

Abbreviations: FPM, Flexible parametric model; NMC, Non-mixture cure model

Another example

In the previous example, MCMs performed poorly because they could not accurately identify the cure fraction. The decrease in the hazard was not well established in the observed period, and the MCMs under-estimated the cure fraction

Sometimes the opposite can happen, especially if Kaplan-Meier curves appear to plateau during observed follow-up periods…

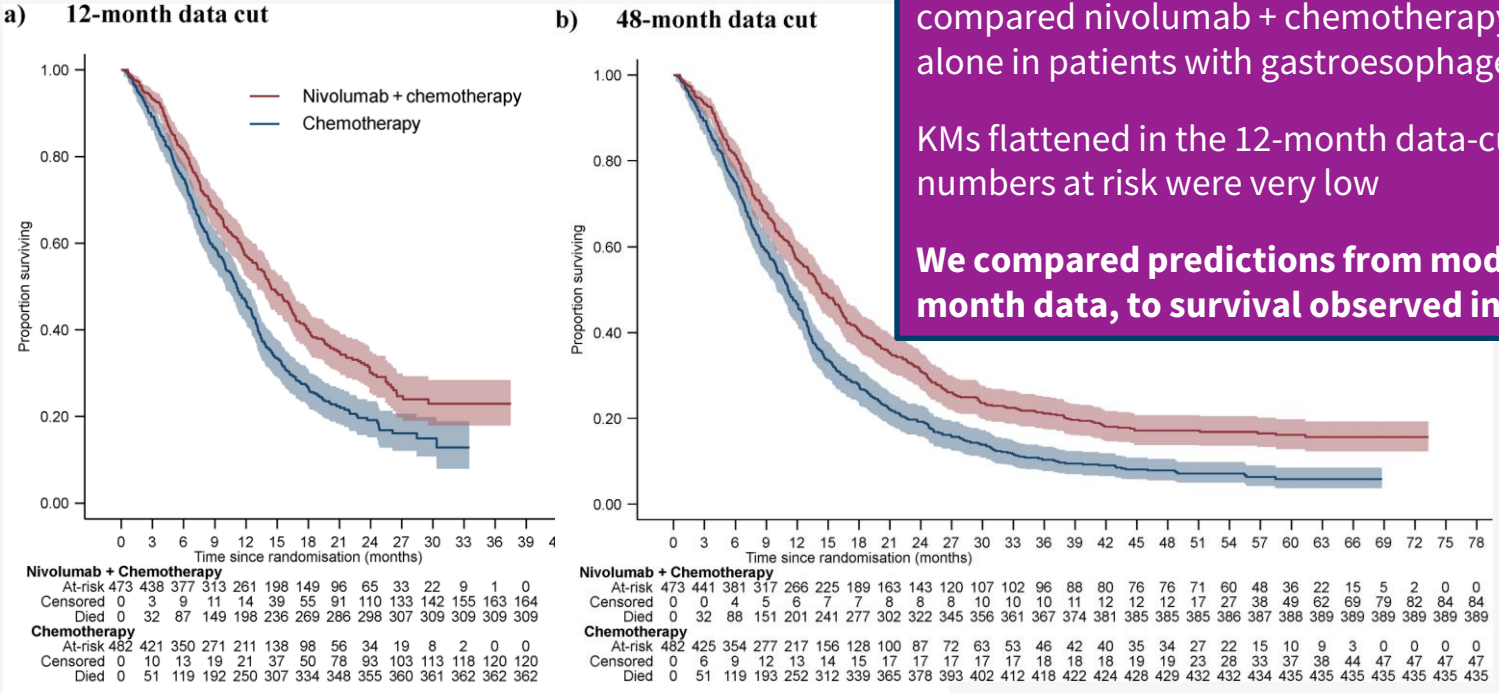
Abbreviations: MCM, Mixture cure model

Another example

Latimer *et al.* tested different cure models by fitting them to the 12-month data-cut from the CheckMate-649 Study, which compared nivolumab + chemotherapy to chemotherapy alone in patients with gastroesophageal adenocarcinoma

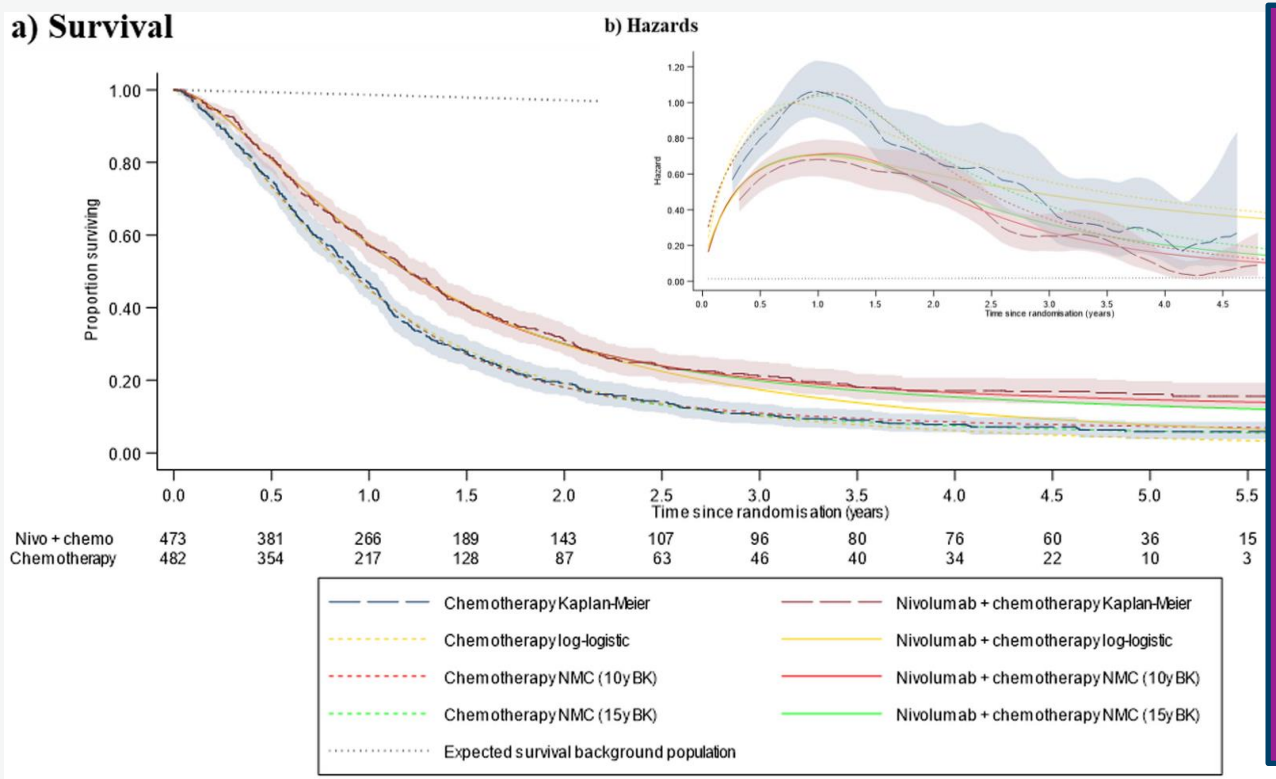
KMs flattened in the 12-month data-cut, but at points where numbers at risk were very low

We compared predictions from models fitted to the 12-month data, to survival observed in the 48-month data-cut



Figures adapted from: Latimer NR, Taylor K, Hatswell AJ, Ho S, Okorogheye G, Chen C, Kim I, Borrill J, Bertwistle D. An Evaluation of an Algorithm for the Selection of Flexible Survival Models for Cancer Immunotherapies: Pass or Fail? *PharmacoEconomics* (2024) https://doi.org/10.1007/s40273-024-01429-0

Another example



In this case, the MCMs again failed to accurately identify the cure fraction – it was generally over-estimated, probably due to the early flattening of the KM curves

This resulted in unrealistic survival estimates (plots were not included on the published graphs, because they were implausible)

In this case, only flexible parametric NMCs were considered to have produced plausible extrapolations

Figures adapted from: Latimer NR, Taylor K, Hatswell AJ, Ho S, Okorogheye G, Chen C, Kim I, Borrill J, Bertwistle D. An Evaluation of an Algorithm for the Selection of Flexible Survival Models for Cancer Immunotherapies: Pass or Fail? *PharmacoEconomics* (2024) https://doi.org/10.1007/s40273-024-01429-0

Abbreviations: BK, boundary knot; MCM, Mixture cure model; NMC, non-mixture cure model

Conclusions

- If a treatment is likely to ‘cure’ some patients, this can have a crucial impact on effectiveness and cost-effectiveness estimates
- When trials have short follow-up, we usually don’t see the cure in the observed data
- But if a cure can confidently be predicted, or is plausible, it makes sense to try to model this

Conclusions

- If a treatment is likely to ‘cure’ some patients, this can have a crucial impact on effectiveness and cost-effectiveness estimates
- When trials have short follow-up, we usually don’t see the cure in the observed data
- But if a cure can confidently be predicted, or is a plausible, it makes sense to try to model this
- I think decision-makers don’t trust cure models, because the MCMs usually used often result in implausible extrapolations
- Flexible parametric NMCs represent a useful alternative that have not been used in HTA
  - They allow some degree of control so that implausible extrapolations can be avoided
  - They offer the possibility of sensitivity analysis, testing a range of boundary knots
  - And they don’t require us to ‘inform’ the cure fraction – we just need to define the cure time-point (which seems easier?) [think about this during Federico’s talk]

Abbreviations: MCM, Mixture cure model; NMC, Non-mixture cure model

Conclusions

- **Caution is required! No models are perfect**
  - Fitting cure models when there is not a cure can lead to extremely misleading results
  - Even when there is a cure, MCMs and NMCs can both extrapolate badly
  - → **But this is easier to protect against with a flexible parametric NMC**

➔ **If cure is plausible, flexible parametric NMCs are more useful than MCMs (even 'informed' MCMs…) and should be considered by HTA decision-makers**

Abbreviations: HTA, Health technology assessment; MCM, Mixture cure model; NMC, Non-mixture cure model

Thanks for listening!

Back-up slides

Frameworks for cure models

University of Sheffield

- What a cure model means, or represents, depends on the framework in which it is fitted
- Assume we have a trial dataset and want to fit a cure model. Possible frameworks are:

All-cause framework

- Fit a cure model to the all-cause survival function observed in our trial
- But we need to build in background 'general population' mortality rates, so that people don't end up living forever

→ When do we build these in?

- If we include them from time 0, we will double count early deaths and our survival model will predict survival that is a bit lower than we observe in the trial
- If we build background mortality in from a later time-point we need to justify the time chosen, and we are essentially saying that no 'other cause' deaths could have occurred before that point

→ So this framework is problematic!

Frameworks for cure models

- What a cure model means, or represents, depends on the framework in which it is fitted
- Assume we have a trial dataset and want to fit a cure model. Possible frameworks are:

Disease-specific framework

- Fit a cure model to the disease-specific survival function observed in our trial, using cause of death information
- But we need to build in background 'general population' mortality rates, so that people don't end up living forever
- This could be done using data on deaths from 'other causes' during the trial, and using lifetables after the trial period

→ But fitting models to 'other cause' deaths observed during the trial may be problematic
→ And we might not have reliable information on the cause of death
→ So this framework is also problematic!

Frameworks for cure models

University of Sheffield

- What a cure model means, or represents, depends on the framework in which it is fitted
- In this talk, I am assuming that we are fitting cure models in a **relative survival framework**

Relative survival framework

- Healthcare interventions often aim to prevent people from dying from the disease the treatment is for
- Logical to consider cure as occurring when the all-cause hazard function for death (the rate at which death occurs over time) for the modelled patient group converges with the general population hazard function

→ This is referred to as the 'relative survival' or ' excess mortality' framework

- We model the difference between the hazard function observed in the trial, and the hazard function in the (age and sex-matched) general population

→ General population mortality rates are used directly in relative survival cure models
→ Do not require data on cause of death
→ Do not require assumptions around when to being incorporating general population mortality

Cure Models for Health Technology Assessment: Can They Be Trusted for Decision-Making?

Federico Felizzi
"informed" mixture-cure models

ISPOR Europe 2024 Issue Panel 129, 18th November 2024

Disclosure

- Former employee of Roche, Novartis
- Shareholder of Novartis
- Employee of the Menarini Group

- **All statements and opinions are my personal views**

Structure

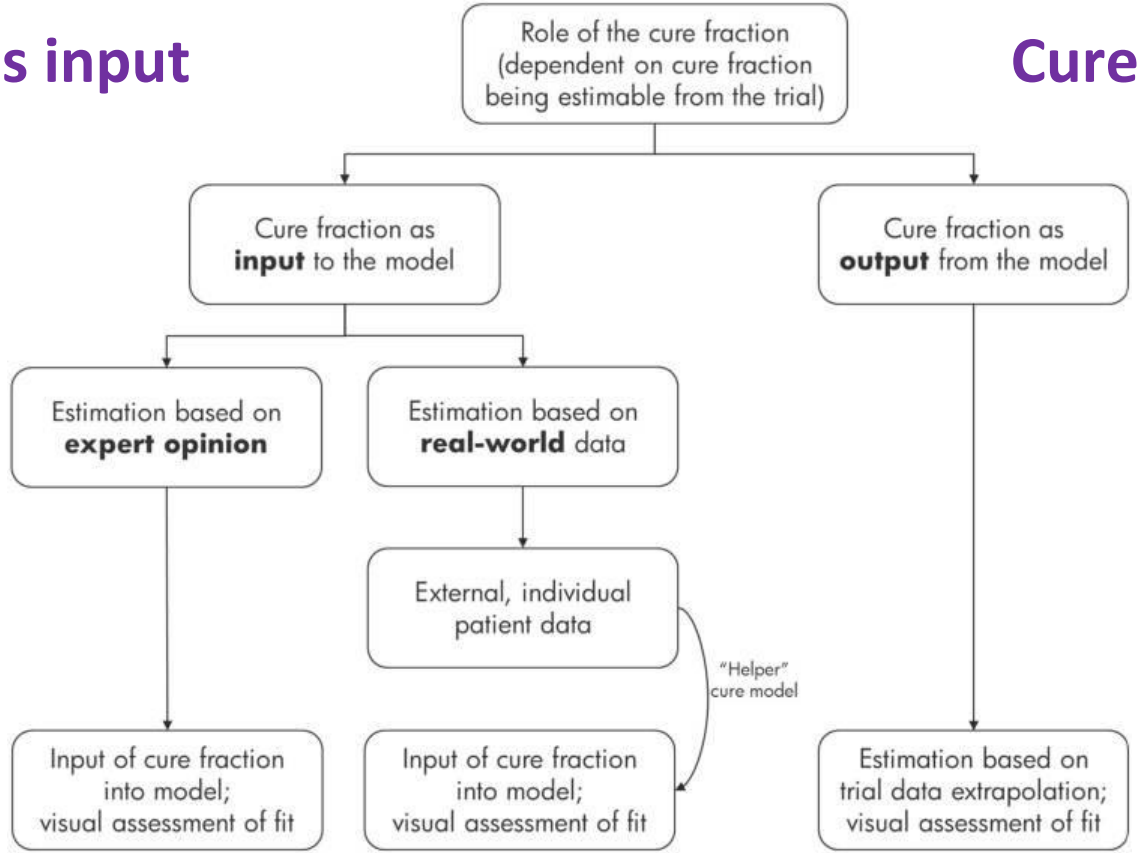
1. Ingredients of mixture-cure models
2. Use of external sources (RWE) to inform the cure proportion
3. Use of intermediate endpoints (PFS) to inform the cure proportion

Mixture-cure models

- Age, Gender and Nationality used to estimate the «background hazard» with the use of country-specific mortality tables
- Clinical trial data for the specific time to event endpoint of interest
- Algorithms to estimate the cure fraction

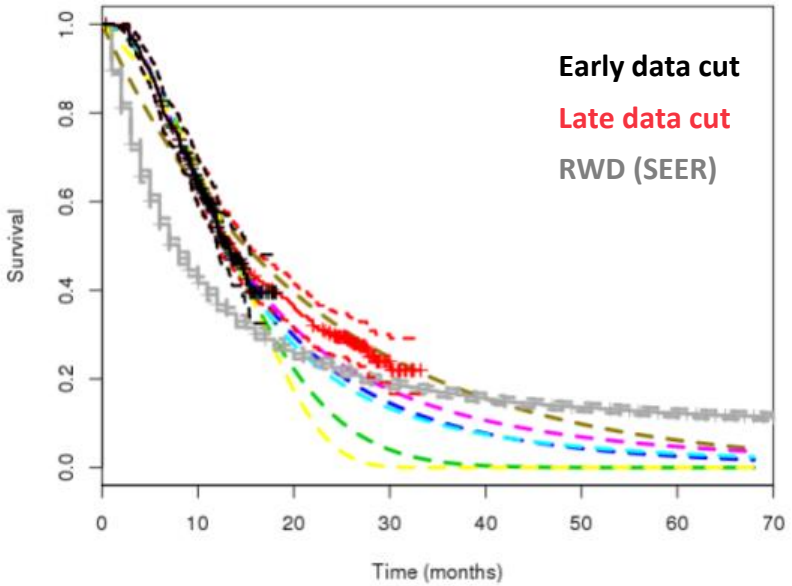
Cure as input

Cure as output



Felizzi F, Paracha N, Pöhlmann J, Ray J. Mixture Cure Models in Oncology: A Tutorial and Practical Guidance. PharmacoEconomics Open. 2021;5(2):143–155.

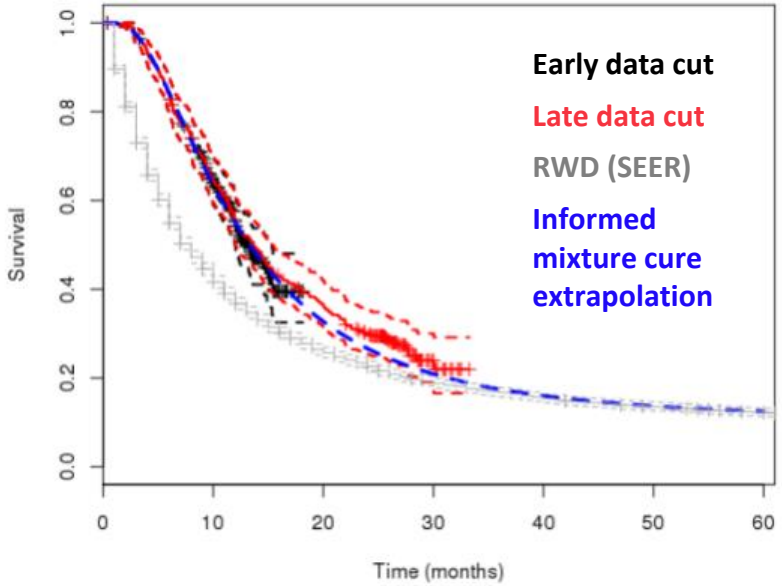
The case for the use of the «informed» mixture-cure



- Case of metastatic melanoma
- Parametric functions systematically under-estimate the new datacut
- RWE (SEER) proves to be a valuable benchmark for evaluating the accuracy of the prediction

CADTH symposium,
Ottawa, Apr 2017

The case for the use of the «informed» mixture-cure



1. The cure proportion extracted from SEER «informs» the OS extrapolation
2. Increased accuracy in the prediction from the early data cut

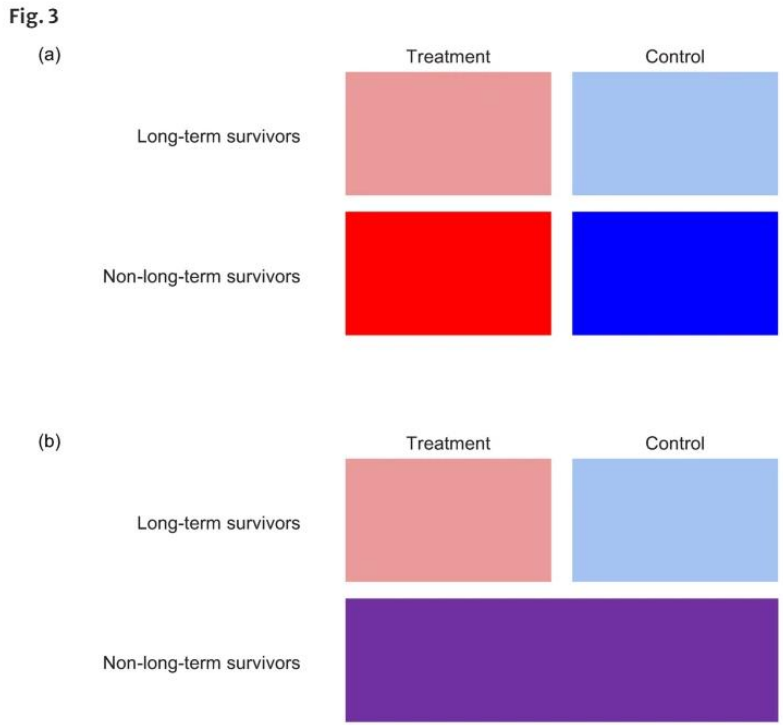
CADTH symposium,
Ottawa, Apr 2017

Intermediate endpoints to inform the cure

a) Assuming non-cured (long-term survivors) proportion (mortality hazard) may vary between intervention and control arm;

b) Assuming no difference in mortality hazard between arms for non-cured

Fit a mixture-cure model on PFS using treatment as covariate on the cure estimate



Felizzi F, Launonen A, Thuresson P-O. Approximation of Long-Term Surv with Polatuzumab Vedotin Plus Bendamustine and Rituximab for Patients with Relapsed/Refractory Diffuse Large B-Cell Lymphoma: Results Based The GO29365 Trial. PharmacoEconomics Open. 2022;7(1):37–46

Independent cured and uncured

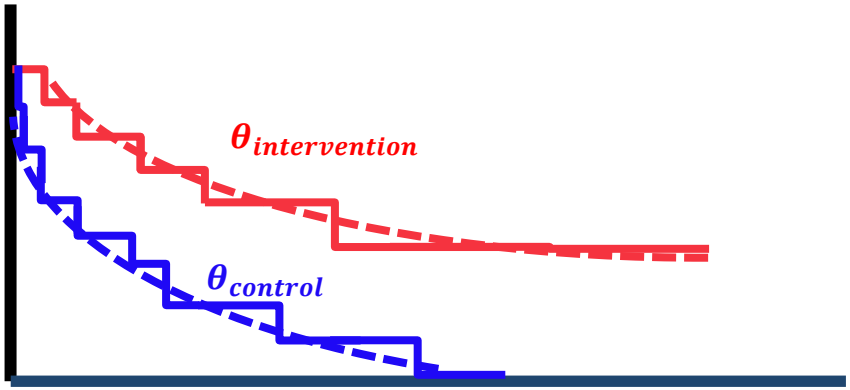
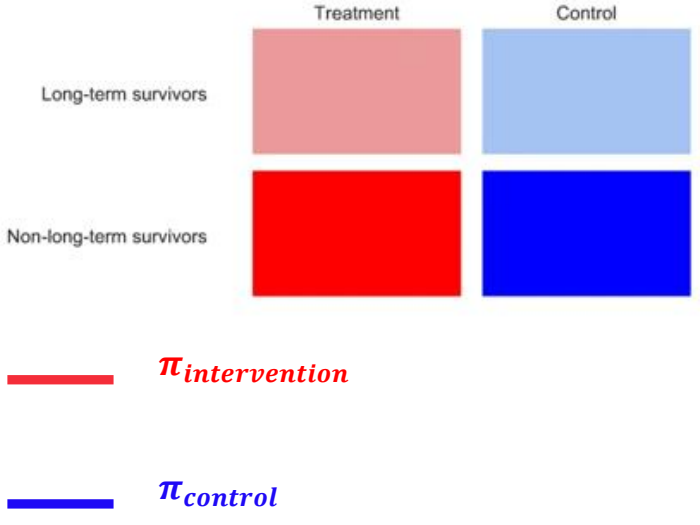
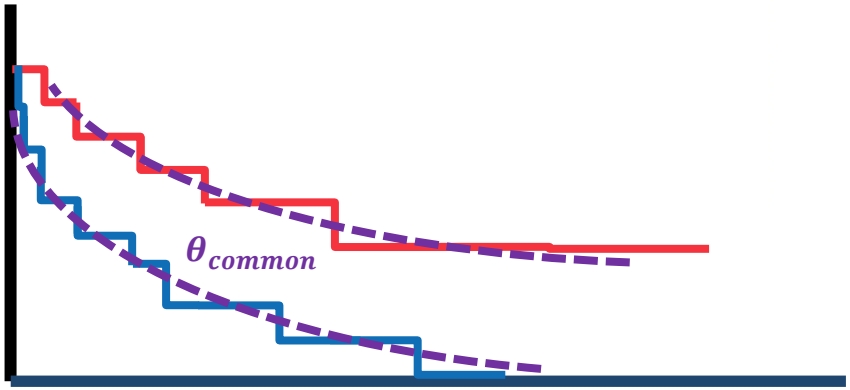


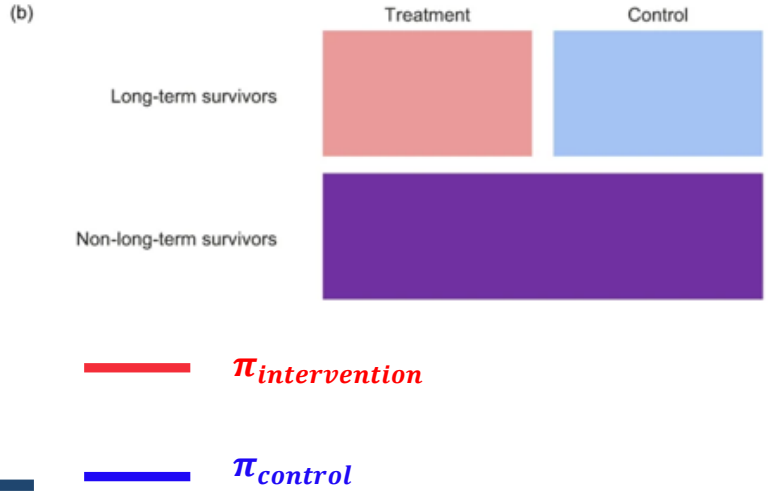
Fig. 3

(a)

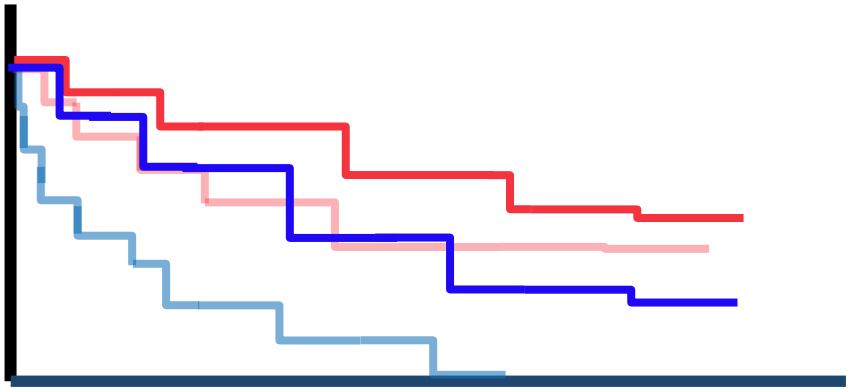


Independent cured, common uncured

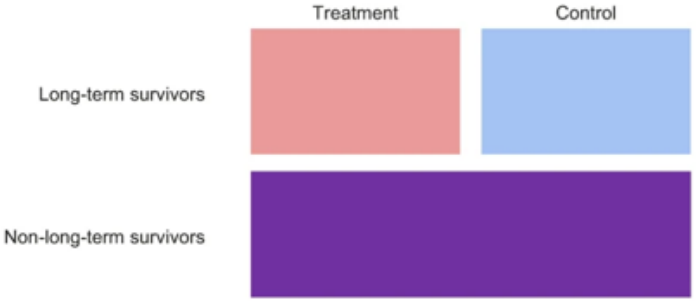




OS extrapolation



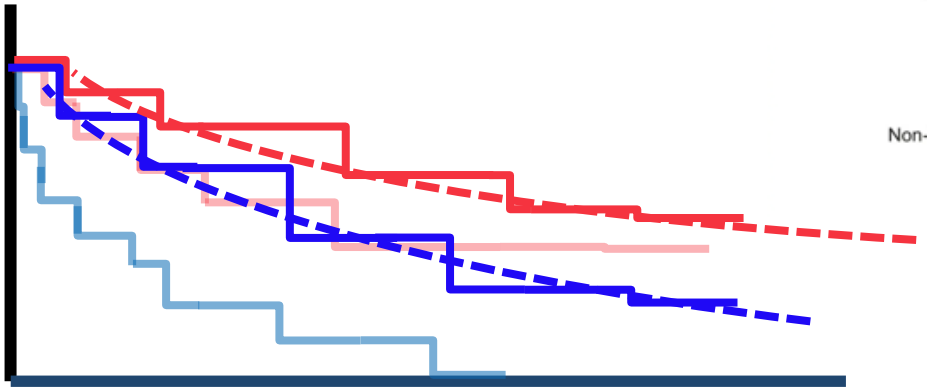
(b)

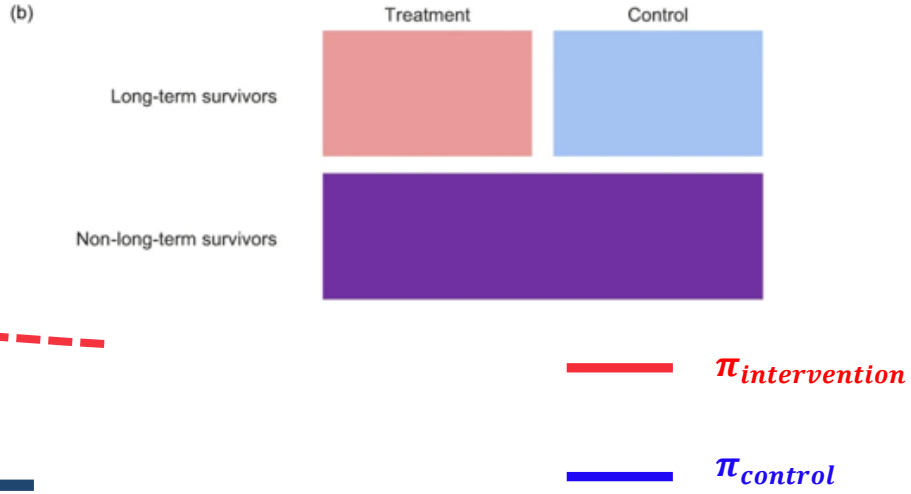


$\pi_{intervention}$

$\pi_{control}$

OS extrapolation, PFS informed





A NICE appraisal example

1. The POLARIX trial evaluated the efficacy and safety of polatuzumab + R-CHP in untreated patients with DLBCL, aiming to improve PFS compared to the R-CHOP regimen.
2. The study provided initial evidence suggesting a potential "cure" fraction for patients in remission at 24 months, supporting the use of mixture cure models in assessing long-term survival and cost-effectiveness for this patient population.

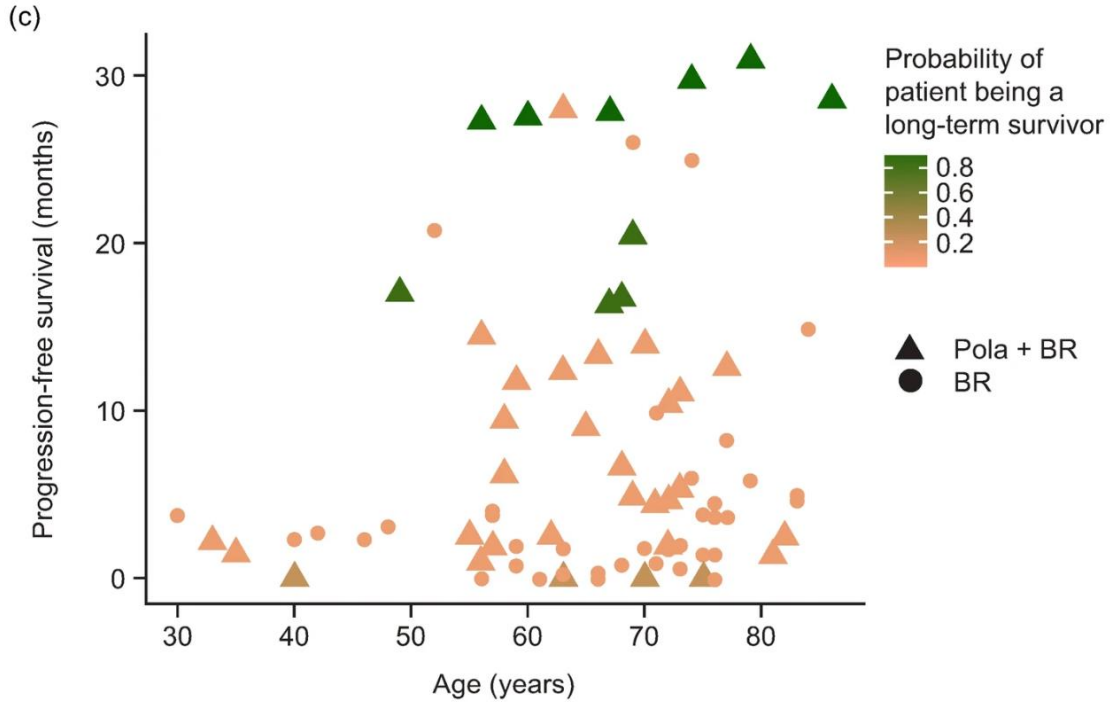
Cooper K, Maund E, Takahashi MT, Shepherd J. Using Cure Modelling for Cost Effectiveness in the NICE Technology Appraisal of Polatuzumab Vedotin in Combination for Untreated Diffuse Large B Cell Lymphoma: An External Assessment Group Perspective. PharmacoEconomics. 2024;42(11):1177–1179.

A NICE appraisal example

1. The Committee was open to the use of a mixture cure model, given that evidence suggested some patients could achieve long-term remission (or be “cured”) if they remained disease-free for 24 months
2. The ERG questioned the application of the cure model in the absence of statistically confirmed (OS), given the immature OS data. The ERG recommended aligning the cure fraction across both treatment arms to reflect the uncertainty

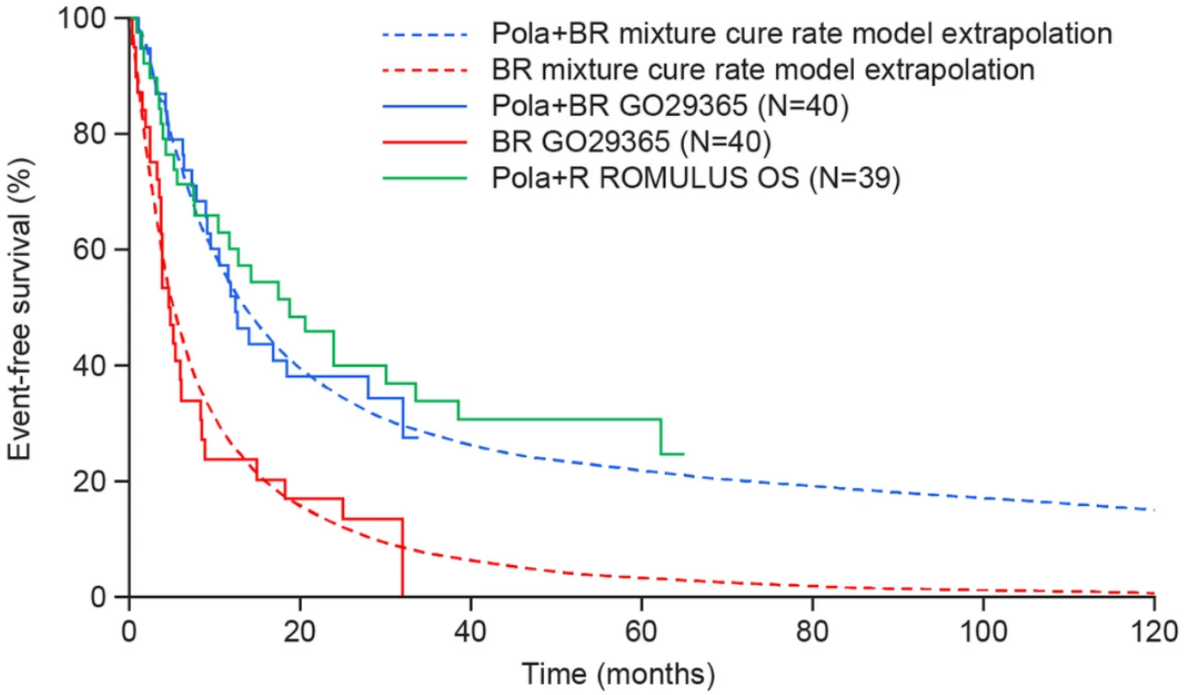
Cooper K, Maund E, Takahashi MT, Shepherd J. Using Cure Modelling for Cost Effectiveness in the NICE Technology Appraisal of Polatuzumab Vedotin in Combination for Untreated Diffuse Large B Cell Lymphoma: An External Assessment Group Perspective. PharmacoEconomics. 2024;42(11):1177–1179.

A side-note, towards individual cured probabilities



Thank you!

Results for OS (PFS-informed) cure



LEERINK
CENTER FOR PHARMACOECONOMICS

Cure Models to Support Decision Making in the US

Melanie D. Whittington, PhD, MS
Managing Director and Head
Leerink Center for Pharmacoeconomics

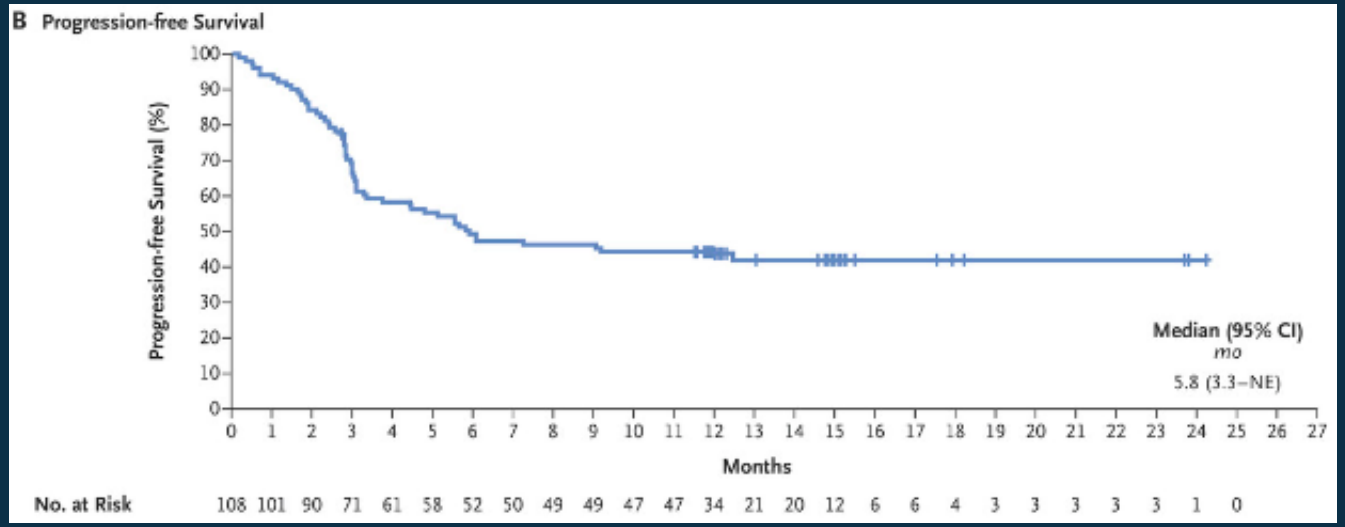
Disclosures

I am a Managing Director and Head of the Leerink Center for Pharmacoeconomics, which is a division of MEDACorp and an affiliate of Leerink Partners.

Example 1

ICER’s Assessment for CAR-T Therapies for Leukemia and Lymphoma

LEERINK
CENTER FOR PHARMACOECONOMICS



Source: Neelapu et al., 2017. NEJM.

Feature	Standard Parametric	Flexible Parametric Non-Cure	Mixture Cure
Parametric curve for downward slope	✓	✓	✓
Knot at curve flattening		✓	
Separate model for cured vs not cured			✓

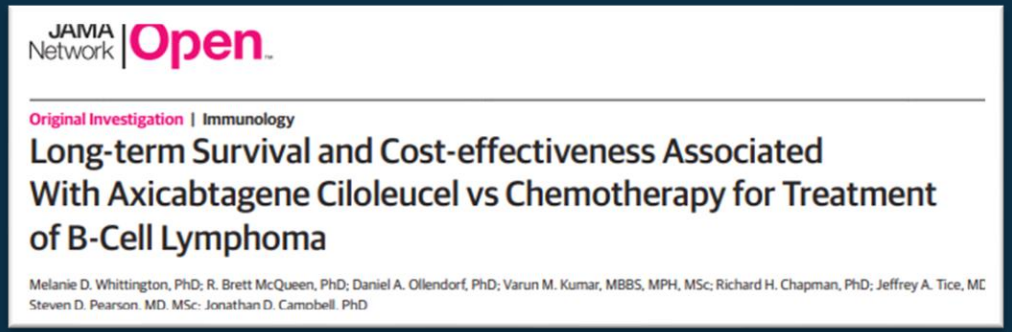
Life Years (discounted)	Standard Parametric	Flexible Parametric Non-Cure	Mixture Cure
Intervention	2.83	7.35	7.66
Comparator	0.94	3.21	3.17
Incremental	1.89	4.14	4.49

Which was “chosen” in the assessment?

- Flexible parametric was the “base-case”
- Presented standard parametric as a lower bound in a scenario analysis

Technical and Practical Considerations

- All can be programmed relatively easily
- Flexible parametric models require determining at what time point to introduce a knot or fit a new curve
- Cure models require defining a “cure” (e.g., who and when)
- Structural uncertainty is not captured well in traditional sensitivity analyses



JAMA Network | Open

Original Investigation | Immunology

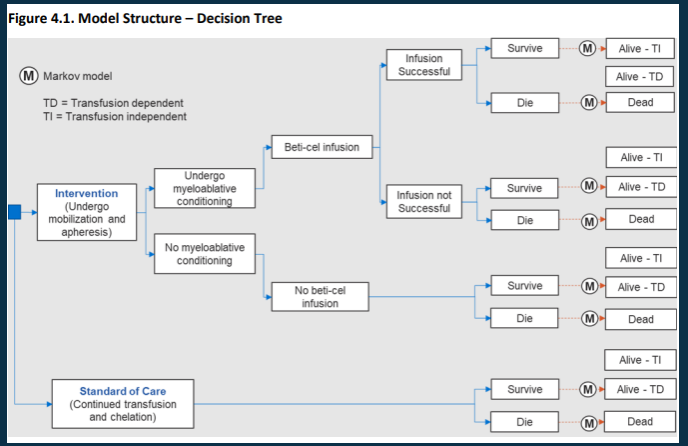
Long-term Survival and Cost-effectiveness Associated With Axicabtagene Ciloleucel vs Chemotherapy for Treatment of B-Cell Lymphoma

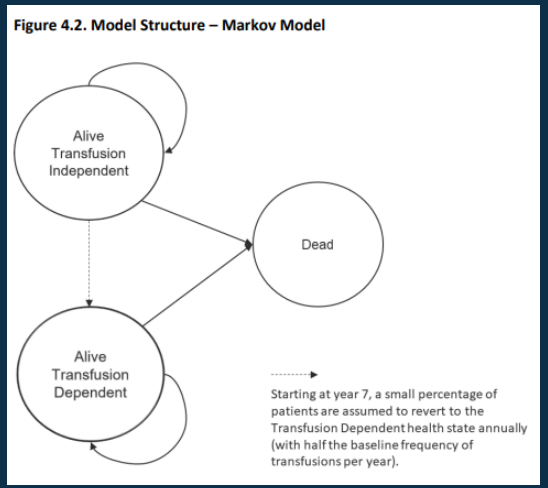
Melanie D. Whittington, PhD; R. Brett McQueen, PhD; Daniel A. Ollendorf, PhD; Varun M. Kumar, MBBS, MPH, MSc; Richard H. Chapman, PhD; Jeffrey A. Tice, MD; Steven D. Pearson, MD, MSc; Jonathan D. Campbell, PhD

Example 2

ICER’s Assessment for Betibeglogene Autotemcel for Beta Thalassemia

LEERINK
CENTER FOR PHARMACOECONOMICS





Source: ICER's assessment on betibeglogene autotemcel for beta thalassemia.

Melanie.Whittington@leerink.com

LEERINK
CENTER FOR PHARMACOECONOMICS