Methods for Causal Inference Using Real-World Data

The Use of Causal Inference Methods in Health Technology Assessment: Addressing Hypothetical Estimands in Clinical Trials

ISPOR Europe 2024, Real World Evidence Summit, 17th November 2024

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

Introduction – Setting the Scene

- Barbra has introduced some complex methods (the ‘g-methods’)

→ We may all need to become more familiar with these if the use of real-world data (RWD) in health technology assessment (HTA) continues to increase as we expect

- For agencies and decision-makers who are most familiar with using evidence from randomised controlled trials (RCTs), this might seem like a big step

- But actually, several of these methods have been used by some HTA agencies for a long time already (but in the analysis of RCTs rather than RWD)

1. Maybe familiarity with methods is not such a big step
2. We can learn from past use of these methods to predict challenges that will create barriers to their acceptance in an RWD setting

Abbreviations: HTA, Health Technology Assessment; RCT, Randomised Controlled Trial; RWD, Real-world data

Treatment switching in RCTs

- **Treatment switching:** Patients switch onto a treatment that they were not randomised to
  - → In oncology trials, often patients in the control arm switch onto the experimental treatment
  - → Or patients in either randomised group switch onto some other non-study treatment
- This has implications for analyses conducted for HTA agencies

HTA decision problem: What is the effectiveness and cost-effectiveness of inserting the new treatment into the treatment pathway at the specified line of therapy, compared to retaining the current standard treatment pathway?

- RCTs are usually analysed using the intention-to-treat (ITT) principle, making no adjustment for treatment changes that occur over time
- If patients switch onto treatments that are not representative of the treatment pathways that would be used in clinical practice, the ITT analysis does not directly address the decision problem
- An adjustment analysis that estimates counterfactual outcomes that would have been observed in the absence of switching is more relevant – a hypothetical estimand

Abbreviations: HTA, Health Technology Assessment; ITT, Intention-to-treat; RCT, Randomised Controlled Trial

Treatment switching in RCTs

- **Treatment switching:** Patients switch onto a treatment that they were not randomised to

7 Key components of the Target Trial protocol:

- Eligibility criteria
- Treatment strategies
- Assignment procedures
- Follow-up period (and time zero)
- Outcomes
- Estimands
- Analysis plan

One of the crucial parts of a Target Trial analyses is defining the treatment strategies we're interested in.

If people deviate from the treatment strategy we're interested in **we censor them and use weighting to remove bias**

→ **We analyse the data ensuring we estimate the treatment effect between the treatment strategies we're interested in**
→ **We're doing the same thing when we adjust for treatment switching in RCTs**

analysis does not directly address the decision problem

- An adjustment analysis that estimates counterfactual outcomes that would have been observed in the absence of switching is more relevant – a hypothetical estimand

Abbreviations: RCT, Randomised Controlled Trial

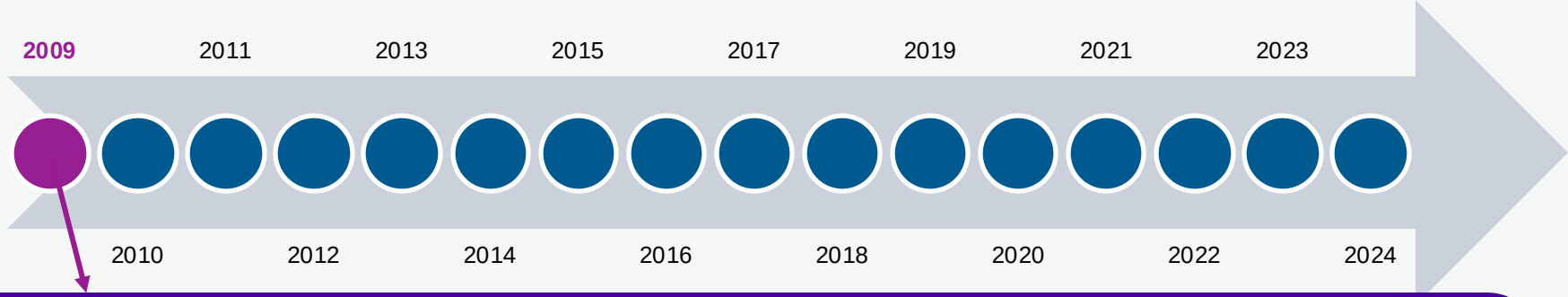
Treatment switching in RCTs

- Using statistical methods to adjust for treatment switching in RCTs is not a new thing in HTA

Abbreviations: HTA, Health Technology Assessment; RCT, Randomised Controlled Trial

Treatment switching in RCTs

- Using statistical methods to adjust for treatment switching in RCTs is not a new thing in HTA



NICE TA179 (2009), sunitinib for GIST

https://www.nice.org.uk/guidance/ta179

- Sunitinib (n=243) vs Placebo (n=118)
- At an interim analysis (max. follow-up 54 weeks): HR for Overall Survival = 0.49
- Patients in control group then allowed to switch to sunitinib → 84% switched
- At the later follow-up: **ITT analysis**: OS HR = 0.88, ICER = £77k
  Rank preserving structural failure time model (with g-estimation): OS HR = 0.51, ICER = £32k

→ **RPSFTM considered acceptable** and sunitinib recommended

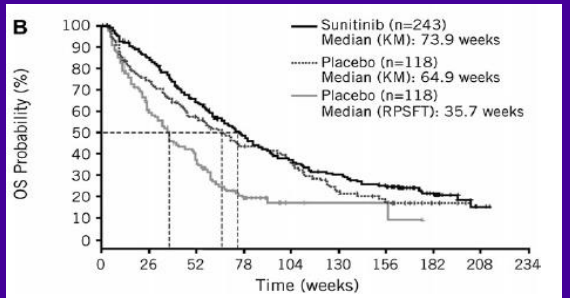
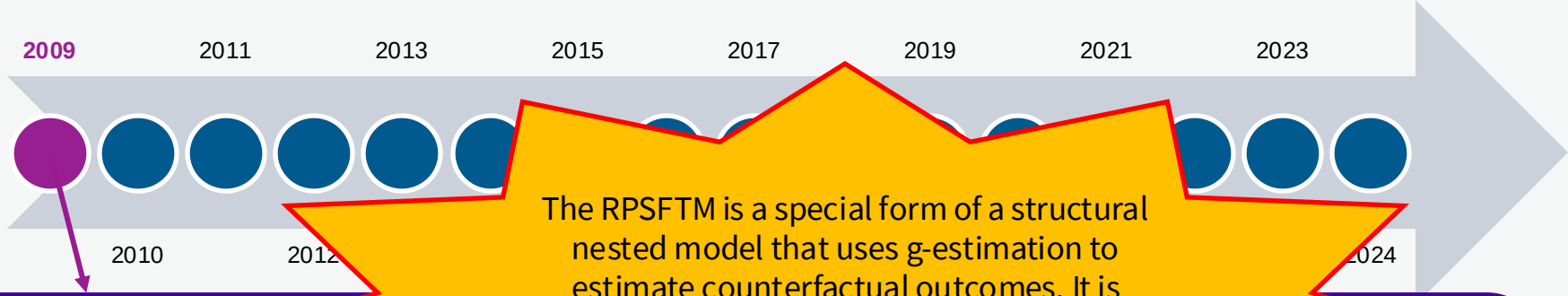


Figure from Blay, 2010, Annals of Oncology. Numbers not identical to those in NICE submission but very close

Abbreviations: GIST, Gastro-intestinal stromal tumours; HR, Hazard Ratio; HTA, Health Technology Assessment; ICER, Incremental cost-effectiveness ratio; ITT, Intention-to-treat; KM, Kaplan-Meier; NICE, National Institute for Health and Care Excellence; OS, Overall survival; RCT, Randomised Controlled Trial; RPSFTM, Rank preserving structural failure time model; TA, Technology Appraisal

Treatment switching in RCTs

- Using statistical methods to adjust for treatment switching in RCTs is not a new thing in HTA



The RPSFTM is a special form of a structural nested model that uses g-estimation to estimate counterfactual outcomes. It is 'special' because it is designed for RCTs, allowing the g-estimation to leverage randomisation to avoid having to make the 'no unmeasured confounding' assumption (Robins and Tsiatis, 1991)[1]

NICE TA179 (2009), su
https://www.nice.org.uk/gu

- Sunitinib (n=243) vs Placeb
- At an interim analysis (ma
- Patients in control grou
- At the later follow-up: **ITT analysis: C**
  Rank preservi
  (with g-estimat HR = 0.51;

→ **RPSFTM considered acceptable** and sunitinib recommended

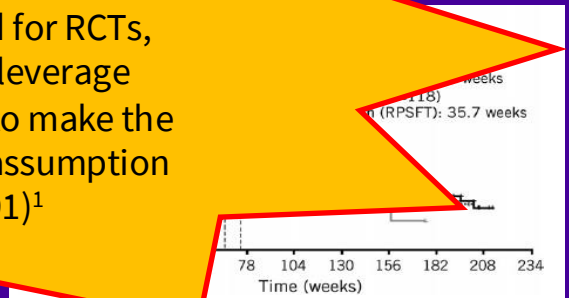
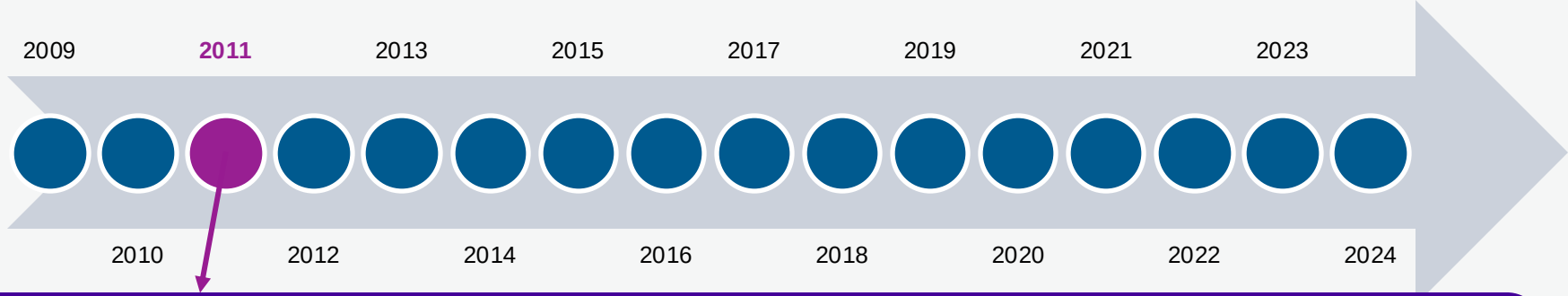


Figure from Blay, 2010, Annals of Oncology. Numbers not identical to those in NICE submission but very close

[1] Robins, J.M., Tsiatis, A.A. Correcting for non-compliance in randomized trials using rank preserving structural failure time models. Comm. Stats-Theory Methods 1991; 20(8):2609-2631.
Abbreviations: GIST, Gastro-intestinal stromal tumours; HR, Hazard Ratio; HTA, Health Technology Assessment; ICER, Incremental cost-effectiveness ratio; ITT, Intention-to-treat; KM, Kaplan-Meier; NICE, National Institute for Health and Care Excellence; OS, Overall survival; RCT, Randomised Controlled Trial; RPSFTM, Rank preserving structural failure time model ; TA, Technology Appraisal

Treatment switching in RCTs

- Using statistical methods to adjust for treatment switching in RCTs is not a new thing in HTA



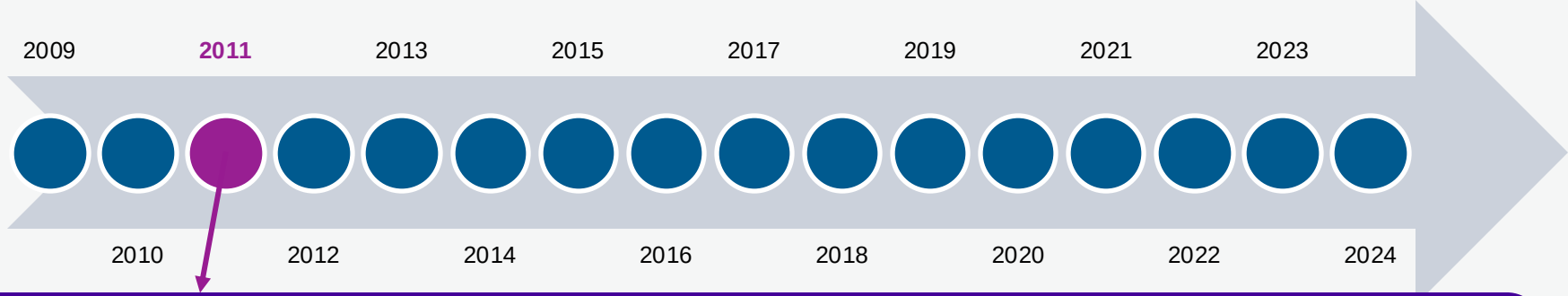
NICE TA215 (2011), pazopanib for renal cell carcinoma (51% of control group patients switched onto pazopanib)
https://www.nice.org.uk/guidance/ta215

- RPSFTM with g-estimation and IPCW with marginal structural models both used (including analyses by Prof Jamie Robins)
- RPSFTM preferred by Appraisal Committee

Abbreviations: HTA, Health Technology Assessment; IPCW, Inverse probability of censoring weights; NICE, National Institute for Health and Care Excellence; RCT, Randomised Controlled Trial; RPSFTM, Rank preserving structural failure time model; TA, Technology Appraisal

Treatment switching in RCTs

- Using statistical methods to adjust for treatment switching in RCTs is not a new thing in HTA



NICE TA215 (2011), pazopanib for renal cell carcinoma (51% of control group patients switched onto pazopanib)
https://www.nice.org.uk/guidance/ta215

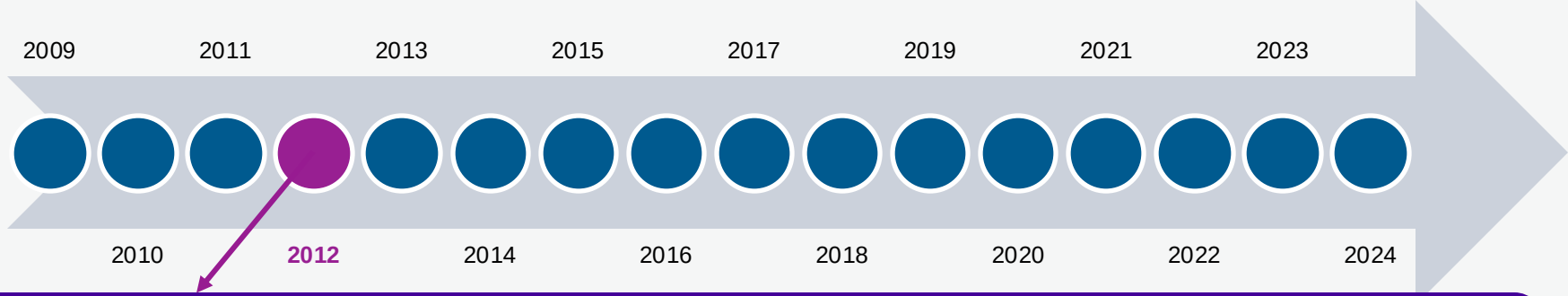
- RPSFTM with g-estimation and IPCW with marginal structural models bot… …e Robins)
- RPSFTM preferred by Appraisal Committee

IPCW in this context involves censoring switchers, and then applying weights to remaining patients to avoid bias. This requires the 'no unmeasured confounding' assumption

Abbreviations: HTA, Health Technology Assessment; IPCW, Inverse probability of censoring weights; NICE, National Institute for Health and Care Excellence; RCT, Randomised Controlled Trial; RPSFTM, Rank preserving structural failure time model; TA, Technology Appraisal

Treatment switching in RCTs

- Using statistical methods to adjust for treatment switching in RCTs is not a new thing in HTA



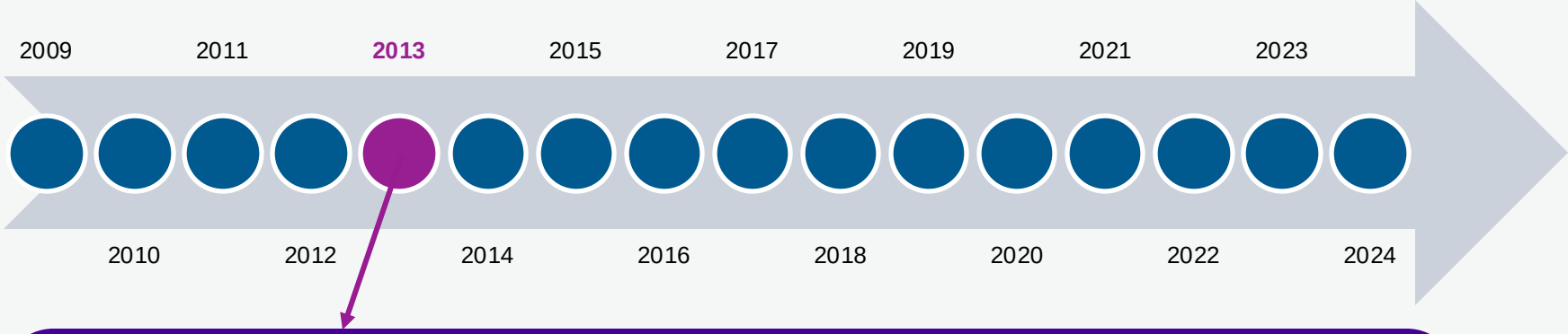
NICE TA263 (2012), bevacizumab for metastatic breast cancer (52% of control group switched onto bevacizumab)
https://www.nice.org.uk/guidance/ta263

- RPSFTM used but considered unreliable by the Appraisal Committee

Abbreviations: HTA, Health Technology Assessment; NICE, National Institute for Health and Care Excellence; RCT, Randomised Controlled Trial; RPSFTM, Rank preserving structural failure time model; TA, Technology Appraisal

Treatment switching in RCTs

- Using statistical methods to adjust for treatment switching in RCTs is not a new thing in HTA



Treatment switching first mentioned in **NICE 'Guide to the Methods of Technology Appraisal'**:
https://www.nice.org.uk/process/pmg9/resources/guide-to-the-methods-of-technology-appraisal-2013-pdf-2007975843781

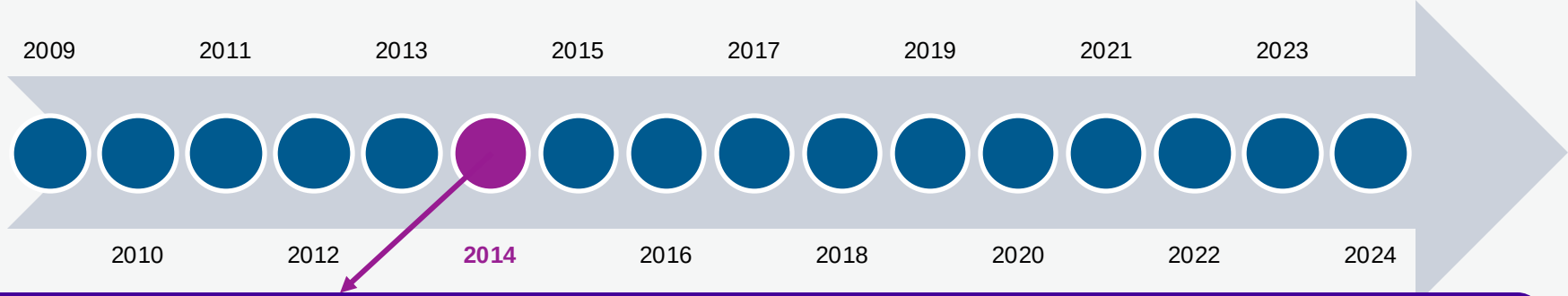
"…when intention-to-treat analysis is considered inappropriate, statistical methods that adjust for treatment switching can also be presented…"

- 'Simple' methods (such as excluding or censoring switchers) should be avoided
- Chosen methods should be well justified

Abbreviations: HTA, Health Technology Assessment; NICE, National Institute for Health and Care Excellence; RCT, Randomised Controlled Trial

Treatment switching in RCTs

- Using statistical methods to adjust for treatment switching in RCTs is not a new thing in HTA



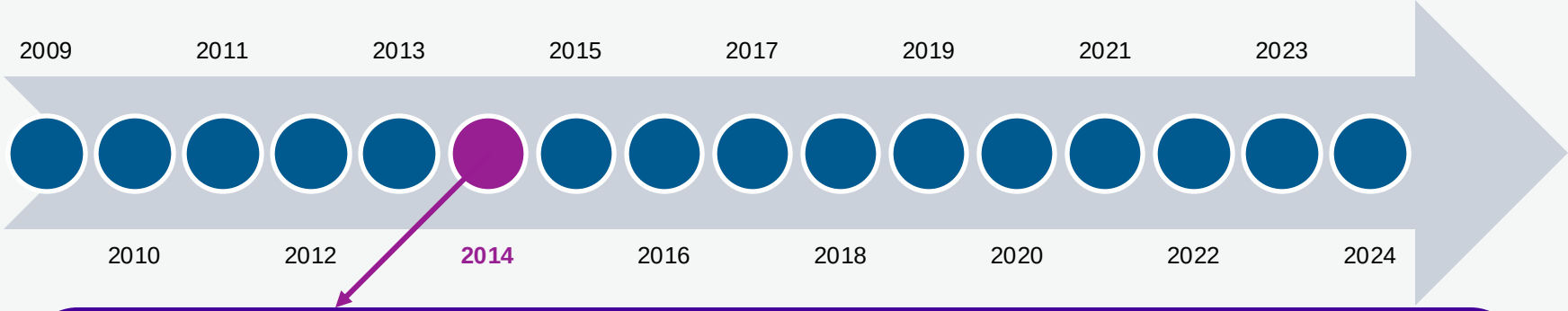
NICE TA321 (2014), dabrafenib for melanoma (57% of control group switched onto dabrafenib)
https://www.nice.org.uk/guidance/ta321

- RPSFTM used but the evidence review group was not convinced

Abbreviations: HTA, Health Technology Assessment; NICE, National Institute for Health and Care Excellence; RCT, Randomised Controlled Trial; RPSFTM, Rank preserving structural failure time model; TA, Technology Appraisal

Treatment switching in RCTs

- Using statistical methods to adjust for treatment switching in RCTs is not a new thing in HTA



NICE Decision Support Unit published **technical support document (TSD) 16**: 'Adjusting survival time estimates in the presence of treatment switching' https://www.sheffield.ac.uk/media/34229/download?attachment

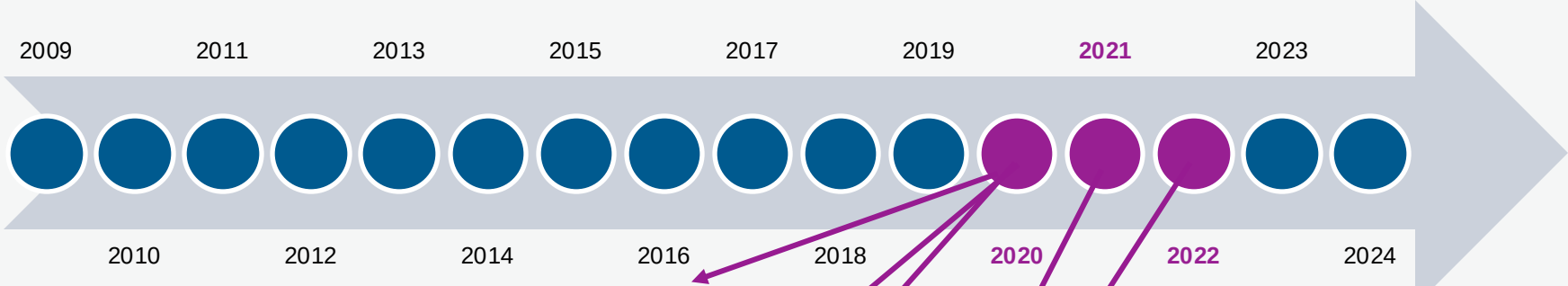
- Highlights that methods need to be able to deal with time-dependent confounding
- Rank preserving structural failure time models (RPSFTM) (i.e. structural nested models with g-estimation)
- Inverse probability of censoring weights (IPCW) in marginal structural models
- Two-stage estimation (TSE) (non- g-method, still requires no unmeasured confounding)
- How to choose between methods

g-methods

Abbreviations: HTA, Health Technology Assessment; IPCW, Inverse probability of censoring weights; NICE, National Institute for Health and Care Excellence; RCT, Randomised Controlled Trial; RPSFTM, Rank preserving structural failure time model; TA, Technology Appraisal; TSD, technical support document; TSE, Two-stage estimation

Treatment switching in RCTs

- Using statistical methods to adjust for treatment switching in RCTs is not a new thing in HTA



NICE TA653 (2020), osimertinib for non-small-cell lung cancer
- RPSFTM used and deemed reasonable by the Appraisal Committee

NICE TA660 (2020), daralutamide for prostate cancer
- RPSFTM conducted but not used by company or Committee

NICE TA740 and 741 (2021), apalutamide for prostate cancer
- Modified RPSFTM and IPCW conducted and accepted

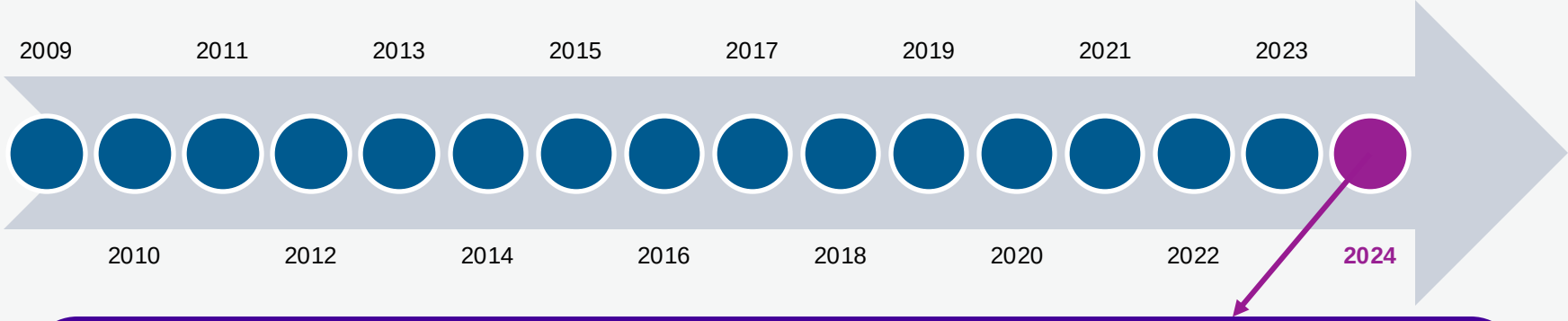
NICE TA709 (2021), pembrolizumab for colorectal cancer
- RPSFTM, IPCW and two-stage estimation conducted but not used in the economic model

NICE TA784 (2022), niraparib for ovarian cancer
- IPCW used but considered unreliable

Abbreviations: HTA, Health Technology Assessment; IPCW, Inverse probability of censoring weights; NICE, National Institute for Health and Care Excellence; RCT, Randomised Controlled Trial; RPSFTM, Rank preserving structural failure time model; TA, Technology Appraisal. Note, TA docs can be found using https://www.nice.org.uk/guidance/taxxx where "xxx" is replaced with the appraisal number

Treatment switching in RCTs

- Using statistical methods to adjust for treatment switching in RCTs is not a new thing in HTA



NICE Decision Support Unit published **TSD 24**: 'Adjusting survival time estimates in the presence of treatment switching – an update to TSD 16' https://www.sheffield.ac.uk/media/65536/download?attachment

- Clarifies which types of switching should be adjusted for (and which should not be)
- Discusses methods that can adjust for more than one type of treatment switching
- Reviews method developments – **adds more complex version of two-stage estimation, including g-estimation**
- Adds detailed reporting guidelines

Abbreviations: HTA, Health Technology Assessment; NICE, National Institute for Health and Care Excellence; RCT, Randomised Controlled Trial, TSD, Technical Support Document

Lessons learned

- TSD 24 was needed because there were problems with the way treatment switching was being dealt with in NICE appraisals
- Some adjustment analyses had been accepted, and some had been rejected. Often due to:

Abbreviations: NICE, National Institute for Health and Care Excellence; TSD, Technical Support Document

Lessons learned

- TSD 24 was needed because there were problems with the way treatment switching was being dealt with in NICE appraisals
- Some adjustment analyses had been accepted, and some had been rejected. Often due to:
  1. **Limited consideration of which types of switches to adjust for – inadequate consideration of treatment strategies**

This will also be crucial when analysing real-world data – and is a key part of the Target Trial framework

Abbreviations: NICE, National Institute for Health and Care Excellence; TSD, Technical Support Document

Lessons learned

- TSD 24 was needed because there were problems with the way treatment switching was being dealt with in NICE appraisals
- Some adjustment analyses had been accepted, and some had been rejected. Often due to:
  1. Limited consideration of which types of switches to adjust for
  2. **Lack of justification for chosen method and model specification – inadequate description of methodological assumptions and justification of variable selection**

This will also be crucial when analysing RWD – especially variable selection and justifying the 'no unmeasured confounding' assumption

Abbreviations: NICE, National Institute for Health and Care Excellence; RWD, real-world data; TSD, Technical Support Document

Lessons learned

- TSD 24 was needed because there were problems with the way treatment switching was being dealt with in NICE appraisals
- Some adjustment analyses had been accepted, and some had been rejected. Often due to:
  1. Limited consideration of which types of switches to adjust for
  2. Lack of justification for chosen method and model specification
  3. **Analyses frequently reported in inadequate detail – covariates included, missing data, model outputs (e.g. range of weights), sensitivity analysis…**

All these points will also need to be covered in an RWD setting – and the reporting guidelines in TSD 24 should help

Abbreviations: NICE, National Institute for Health and Care Excellence; RWD, real-world data; TSD, Technical Support Document

Conclusions

- It is not new to use g-methods in analyses for HTA agencies
  - Treatment switching adjustment analyses conducted over the last 15 years provide a nice example
- Provides learnings on what needs to be done for analyses that use these methods to be considered acceptable
  - Be clear about the purpose of the analysis (the Target Trial framework will help here)
  - Justify the chosen method and model specification
  - Report analyses clearly and comprehensively
- There will be additional challenges when applying these methods to real-world data
  - Data quality and consistency is less assured. Missing data is more likely
  - The lack of randomisation means all methods require 'no unmeasured confounding'
- Methods are acceptable, if applied appropriately, and data are of sufficient quality
  - Data quality likely to be the key barrier to acceptance of analyses of non-randomised studies…

Abbreviations: HTA, Health Technology Assessment

Thanks for listening!