How could we handle the occurrence of death when analyzing continuous endpoints? An example of PRO endpoints

On behalf of the OncoEstimand SIG
PRO Task Force
www.oncoestimand.org

Konstantina Skaltsa, *Director, Statistics and Psychometrics, Patient-Centered Solutions*

12th June 2023, PSI annual conference

Acknowledgement

PRO Task Force

- This work started by a methods review performed by Michael O'Kelly (IQVIA) and Bohdana Ratitch (Bayer) a few years ago. It has now developed incorporating estimand perspective and a pragmatic selection of methods for PRO needs
- From the PRO Task Force
  - Rachael Lawrence has kindly provided thoughts on the slides
  - Jonathan Siegel has provided thoughts on the topic in our meetings

Contents

Setting the scene

Let’s all align on an example setting

Setting

- 2-arm (active vs control) phase 3 clinical trial in a late-phase solid-tumour oncology indication
- Primary endpoint is PFS or OS
- **Change from baseline in QoL or symptoms X at Week Y** is a (key) secondary endpoint – there may be label claims, but not relevant to the discussion
- QoL is collected through a multi-item questionnaire every Z weeks
- Patients are treated until disease progression, unacceptable toxicity, investigator’s decision etc
- Death may occur in these trials **prior to Week Y** rendering the data at the timepoint of interest unobservable

PFS: Progression-Free Survival; QoL: Quality of Life; OS: Overall Survival

A couple of FDA responses from the Oncology division

FDA has major concerns regarding the statistical analyses as proposed: In general, PROs for superiority and non-inferiority may not be interpretable for efficacy due to mortality. The mixed model repeated measures (MMRM) relies on the assumption that data are missing at random (MAR). If a patient is missing due to death, the MAR assumption is likely not a reasonable assumption, which can lead to bias in the estimated treatment effect.

FDA Oncology Division 2021

We are concerned about the interpretability of Physical functioning/Global health status/QOL for efficacy due to the observed mortality on this trial. Mixed Model Repeated Measures (MMRM) relies on the assumption that data are missing at random (MAR), therefore if a patient is missing due to death, the MAR assumption is likely not a reasonable assumption. This could lead to bias in the estimated treatment effect.

FDA Oncology Division 2021

FDA: Food and Drug Administration; MAR: Missing At Random

Estimand considerations

Treatment policy: not possible

- Slide intentionally left blank, as no data can be collected post-mortem.

While-on-treatment could be an option for some limited cases

- Not a popular strategy in efficacy endpoints in registrational trials

If drug is not expected to prolong survival

- A while-alive strategy may be appropriate for estimating a treatment effect → small portion of registrational clinical trials

If there is a survival benefit

- Deaths (or timing of them) imbalanced across arms:
  - Potential concern – is there?
  - It may be desirable to include the survival benefit in the estimand

Different estimator options may be important in this discussion

Hypothetical strategies have been implicitly used for years

The MMRM has long been a standard way to estimate a treatment effect in the PRO world. Time is typically included as categorical variable and the covariance structure ideally unstructured. There is plenty of literature reporting MMRM results in oncology.

One of its claimed strengths has been “its ability to deal with the missing data”.

MMRM has been recommended by SISAQoL’s 1st publication

Main assumption
MMRM assumes patients are **still alive** and receiving the randomized treatment

Some claim this could be acceptable “if number of deaths is low”

How low is low?

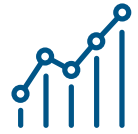
3%

5%

10%

?

Common misconception in the PRO world:
MMRM serves a **while-on-treatment** strategy



Composite strategy considers death is a poor outcome

Considerations

- Composite strategy considers death as an **unfavourable** outcome → could be argued it would be considered a rather sensible strategy by many/most
- Easily operationalized when a responder endpoint is defined – not so straightforward if an analysis on the original continuous/ordinal PRO scale is planned/desired

Numerical values

Logic

- Values after death are **assigned a numerical poor value** from the scale range (e.g. worst score)

Death ranked as a distinct category

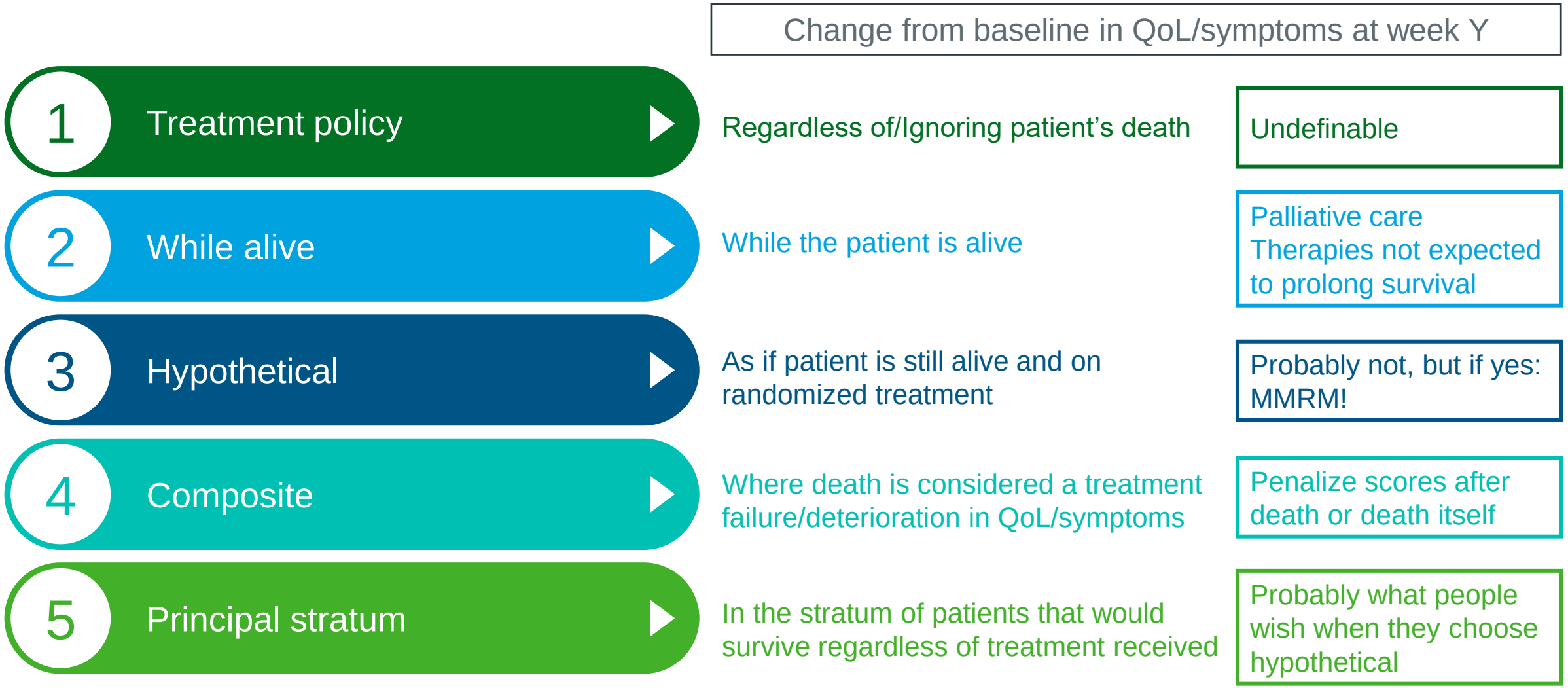
Logic

- **Qualitatively differentiate the death state** by ranking it differently to other poor PRO states, i.e. consider death as a distinct category to patients who are alive and doing very poorly

Principal stratum

- Overlooked to date, as targeting a non-ITT population which will be predicted by means of modelling, based on (potentially incomplete) confounders
  - However, may be useful when the ICE of interest is death: estimate the treatment effect **among patients who would not die**
- We argue it may be a valuable supplementary analysis to be considered together with the treatment effect estimate under the composite strategy – to help evaluate how much the composite result is driven by survival versus the PRO changes.
- Isolates death by exploring the treatment effect only in the stratum of patients who would survive
- Is the timeframe of relevance here?

So what is the treatment effect we are interested in?



How should death be dealt with when estimating the treatment effect in repeatedly collected COAs? Konstantina Skaltsa, IQVIA, 7th November 2022, ISPOR EU Conference

QoL: Quality of Life; MMRM: Mixed Model Repeated Measures

Potential estimators

Estimator options

While alive 1	Hypothetical 2	Composite 3	Principal stratum 4
• AUC • Average scores while patient alive	• MMRM • MI	• MMRM imputing poor scores • Rank-based ANCOVA • Quantile regression • Hodges-Lehmann estimator • Win ratio / win odds	• Any standard analysis estimating a treatment effect in stratum of interest (e.g. stratum of those who *would survive* irrespective of treatment)
• Well-known patient-level endpoints • Standardized and unstandardized versions	• Well-known methods	See next slide for Composite	• Multiple imputation (MI) can be used to allow appropriate uncertainty with regard to stratum.
• May not discriminate between long survival/poor HRQoL and shorter survival/great HRQoL **unstandardized AUC could be categorized as Composite*	• MAR assumption generally implausible (“if patient were still alive”)		• Inference is on stratum, not on Intention-To-Treat population • Strong assumptions when predicting belonging to stratum

ANCOVA: Analysis of Covariance; AUC: Area Under the Curve; HRQoL: Health-Related Quality of Life; MAR: Missing At Random; MI: Multiple Imputation; MMRM: Mixed Model for Repeated Measures

Some estimators targeting a composite strategy for death

Composite

MMRM imputing poor scores after patients' death
- Worst score may be appropriate for short-range scales (e.g., 0-3)
- Selection of post-mortem value for COAs challenging / Variance of outcome post-death distorted

Rank-based ANCOVA
- Based on ranks, rather than scores
- Provides p-value only, no estimate of treatment effect

Quantile regression
- Provides treatment effect estimate on original scale
- May not work if too many deaths

Hodges-Lehmann estimator
- Provides treatment effect estimate on original scale

Win ratio / win odds
- Based on ranks / Provides interpretable treatment effect
- Treatment effect is not on original scale, therefore harder to communicate to clinicians/patients

See also PSI webinar (28APR2022): Novel composite estimands and their analysis. Available at: https://www.psiweb.org/events/past-psi-events/2022/04/28/default-calendar/psi-webinar-novel-composite-estimands-and-their-analysis

Discussion

Summary thoughts

Hypothetical

- May still be acceptable if number of deaths “low”
- Simulations may reveal what “low” can be
- Strongly recommended to be accompanied by supplementary analyses, e.g. composite

Principal stratum

- Consider as useful supplement to a composite strategy
- Pushbacks on assumptions shouldn’t be an excuse – MMRM makes a lot of (implausible) assumptions as well

Composite

- General consensus that death is a poor outcome
- Penalization options vary leading to varying population-summaries - some unfamiliar to stakeholders that receive these results

While on treatment / alive

- Reserved for a few cases where treatment is not intended to affect survival

Q:

What are your experiences dealing with death in PRO data?

What are your thoughts on the **principal stratum** strategy?

What is the place of **hypothetical strategy** when dealing with death?

Is death a poor outcome? Should we make separate considerations for symptoms or functioning or QoL?