Exercise 2: Aligning analysis with the estimand

Estimand 1

The primary objective of Dr Hesketh's trial is to compare the superiority of TD23 against placebo in routine practise for children with STRA over a 52-week treatment period.

The primary outcome is the change in Asthma Control Questionnaire (ACQ) score at week 52. The ACQ results in a numerical score ranging from 0 to 6 where 0 represents excellent asthma control and 6 represents extremely poor control. The ACQ will also be measured at weeks 12, 24.

The primary estimand Dr Hesketh has selected to align with this objective is described by the following 5 attributes:

Estimand attribute	
Treatment condition	A 52-week course of TD23 at 50mg every 2 weeks compared to a 52-week course of matching placebo at 50mg every 2 weeks, **regardless** of *any treatment discontinuation* or *use of rescue medication* or *background ICS maintenance therapy*
Population	Children with STRA, as defined by the trial inclusion and exclusion criteria.
Variable (outcome)	Change from baseline in the ACQ at week 52
Handling intercurrent events	Intercurrent event 1: a **treatment policy** strategy will be used to estimate the treatment effect, regardless of the early discontinuation of treatment for any reason (as part of treatment). Intercurrent event 2: a **treatment policy** strategy will be used to estimate the treatment effect, regardless of the use of short-acting β_2-adrenergic receptor agonist (either salbutamol or levosalbutamol) as rescue medication (as part of treatment). Intercurrent event 3: a **treatment policy** strategy will be used to estimate the treatment effect, regardless of the use of background ICS (as part of treatment).
Summary measure	Mean difference in the outcome variable between treatment conditions

For questions 1 and 2 suppose all data are complete (i.e. loss to follow-up is not an issue) in this trial. In questions 3 and 4 we will consider the additional impact of missing data.

Question 1) Your task is to discuss an analysis plan that will provide an estimate of Dr Hesketh's primary estimand.

Which of the following estimators would you suggest?

a) Censor data from time of treatment discontinuation; apply Inverse probability of censoring weighting and a weighted linear regression at the primary time point.
b) A likelihood-based repeated measures approach, such as mixed model for repeated measures. Fit to data collected prior to treatment discontinuation from all randomized patients.
c) Censor data from time of rescue medication initiation or treatment discontinuation; apply Inverse probability of censoring weighting and a weighted linear regression at the primary time point.
d) Censor data from time of treatment discontinuation or initiation of rescue medication or background ICS. Fit a likelihood-based repeated measures approach, such as mixed model for repeated measures.
e) Use all available data collected from all randomized patients and fit a likelihood-based repeated measures approach, such as a mixed model for repeated measures, or equivalently as no missing data, a linear regression at the primary time point.

Estimand 2

A supportive objective is to compare the superiority of TD23 against placebo, if rescue medications (intercurrent event 2) were not available, otherwise as used in routine practise for children with STRA over a 52-week treatment period. The following hypothetical supplementary estimand was selected to align with this objective:

Estimand attribute	
Treatment condition	A 52-week course of TD23 at 50mg every 2 weeks compared to a 52-week course of matching placebo at 50mg every 2 weeks, **regardless** of *any treatment discontinuation* or *use background ICS maintenance therapy*, **if** *rescue medications* **were not available**
Population	Children with STRA, as defined by the trial inclusion and exclusion criteria
Variable (outcome)	Change from baseline in the ACQ at week 52
Handling intercurrent events	*Intercurrent event 1:* a **treatment policy** strategy will be used to estimate the treatment effect, regardless of the early discontinuation of treatment for any reason (as part of treatment). *Intercurrent event 2:* a **hypothetical** strategy will be used to estimate what the treatment effect would have been if all patients did not have rescue medications available (no use of short-acting β_2-adrenergic receptor agonist (either salbutamol or levosalbutamol)). *Intercurrent event 3:* a **treatment policy** strategy will be used to estimate the treatment effect, regardless of the use of background ICS (as part of treatment).
Summary measure	Mean difference in the outcome variable between treatment conditions.

Question 2) Which of the following estimators would you suggest to provide the main estimate for the supplementary estimand?

a) Censor data from time of rescue medication initiation or background ICS use. Fit a likelihood-based repeated measures approach, such as mixed model for repeated measures.
b) Use a two-stage least squares instrumental variable regression to estimate the complier average causal effect (CACE), where a complier is defined as an individual who would not use rescue medication.
c) Censor data from time of rescue medication initiation; apply Inverse probability of censoring weighting and a weighted linear regression at the primary time point.
d) Censor data from time of initiation of rescue medication or treatment discontinuation or background ICS use. Fit a likelihood-based repeated measures approach, such as mixed model for repeated measures.
e) Censor data from time of initiation of rescue medication or treatment discontinuation; apply Inverse probability of censoring weighting and a weighted linear regression at the primary time point.

Advanced questions:

Question 3) We have not yet considered the additional impact of having missing data for the primary estimand. Suppose now *some* data was actually missing from patients in the TD23 group who discontinued their treatment early. Some patients in the TD23 group who discontinued treatment early continued to be observed. If you used multiple imputation to impute this missing data what assumption would this make for the missing data?

Question 4) What alternative methods of estimation could you suggest for the primary estimand to explore the impact of different missing data assumption for patients in the TD23 group who discontinued their treatment early? All missing data (from the patients in the TD23 group who discontinued their treatment early) are monotone.

Solutions

Question 1) What statistical method of estimation would you suggest to provide the main estimator for the primary estimand?

Answer: e

All patient values of the outcome variable of interest (ACQ at 52 weeks), despite the occurrence of any intercurrent events, are needed for this estimand which uses a treatment policy to handle all three relevant intercurrent events. A treatment policy strategy considers the effect of the treatment of interest, regardless of whether intercurrent event occurs. Therefore it would not be appropriate to censor data following any of the intercurrent events (as done in a, b, c or d)

A likelihood based repeated measures approach, such as a mixed model for repeated measures fitted to all collected ACQ values at week 12, 24 and 52 would be suitable. Typical models for such an analysis would include treatment group, time, treatment group crossed with time, baseline covariates (e.g. baseline value of ACQ) and any randomization stratification factors. Or, given you are told all data are complete equivalently a linear regression or an analysis of variance (ANOVA) at the primary time point of 52 weeks would be suitable (including treatment group, baseline covariates (e.g. baseline value of ACQ) and any randomization stratification factors).

Options (a) and (b) would alternatively both estimate the treatment effect in a hypothetical scenario if all treatment was adhered to (regardless of use of rescue medication or background ICS's); in both cases data is not used after treatment discontinuation. Option (a) weights the observed on-treatment data to account for the treatment discontinuation. Option (b) implicitly imputes the data censored post treatment, assuming the data follows the behaviour of the observed on-treatment data. Option (c) would estimate the treatment effect in a hypothetical scenario if all treatment was adhered to and rescue medication was not available (regardless of use of background ICS's); the remaining observed on-treatment and without rescue use data is weighted to account for the treatment discontinuation and rescue medication use. Option (d) would estimate the treatment effect in a hypothetical scenario if all treatment was adhered to and rescue medication and background ICS was not available; the data censored post treatment discontinuations, rescue medication use and ICS use is implicitly imputed, assuming it follows the behaviour of the observed on-treatment data.

Question 2) Discuss the following statistical methods and select one that will provide an estimate of Dr Hesketh's supplementary estimand.

Answer: c

Data that are observed after initiation of rescue medication are not relevant for this estimand. Observations after treatment non-adherence or use of background therapies are useful (given no initiation of rescue medication) therefore a, d and e are not suitable as they do not use data collected after either treatment discontinuation or background ICS use (or both of these events).

Option (c) appropriately censors the data from time of rescue medication initiation only; applying Inverse probability of censoring weighting based on the MAR assumption including the observed data of all patients as randomised prior to rescue medication initiation and a weighted linear regression at the primary time point (typical models for such an analysis would include treatment group, baseline covariates (e.g. baseline value of ACQ) and randomization stratification factors).

After censoring data from time of rescue medication initiation only, other potential estimators (not exhaustive) aligned with this supplementary estimand are:

- A likelihood based repeated measured approach, such as a mixed model for repeated measured fitted to the observed data of all patients as randomised prior to rescue medication initiation. This model makes the MAR assumption for missing data, which assumes that patients with missing data (those rescued) would have ACQ outcomes like those of similar patients in their treatment group who were not rescued through the time point at which data are missing.

- Multiple imputation under the MAR assumption followed by a repeated measures analysis of the longitudinal data would be suitable (or given all data will be complete post-imputation equivalently a linear regression or an analysis of variance (ANOVA) at the primary time point of 52 weeks could be applied post MI). The censored post-rescue data would be imputed to follow the distribution of the data observed in the absence of rescue medication.

Option (a) would alternatively estimate the treatment effect in a hypothetical scenario if rescue medication and background ICS were not available (regardless of any treatment discontinuation).

Option (b) would estimate the treatment effect for only the subgroup of patients who would not use rescue medication. This corresponds to a principal stratification strategy to handling rescue medication, which does not align with the estimand of interest. The hypothetical estimand of interest seeks to know the treatment effect for all children with STRA who meet the trial inclusion/exclusion criteria if rescue medication was not available.

Option (d) would alternatively estimate the treatment effect in a hypothetical scenario if rescue medication was not available and background ICS was not available and all treatment was adhered to.

Option (e) would alternatively estimate the treatment effect in a hypothetical scenario if rescue medication was not available and all treatment was adhered to (regardless of background ICS use).

Question 3) We have not yet considered the additional impact of missing data for the primary estimand. Suppose that some data was actually missing from patients in the TD23 group who discontinued their treatment early. If you used multiple imputation to impute their missing data what assumption would this make for their missing data?

Multiple imputation performed on all the observed ACQ outcome assumes that all the patients with missing data in the TD23 group who discontinued treatment early would have had ACQ outcomes like those of similar patients observed in their treatment group (including those observed who are on-treatment and others who are off-treatment) at the time point at which data are missing. This type of assumption is referred to as missing at random (MAR).

Question 4) What alternative methods of estimation could you suggest for the primary estimand to explore the impact of different missing data assumption for patients in the TD23 group who discontinued their treatment early? All missing data (from the patients in the TD23 group who discontinued their treatment early) are monotone.

Since we know the patients who were allocated to TD23 with missing data had discontinued treatment early we could alternatively employ an analysis that conditions the missingness upon the occurrence of the Intercurrent event of early treatment discontinuation, to reflect the value that would have been observed given treatment discontinuation. That is, we could use a method of estimation that assumes those patients with missing data would have ACQ outcomes like those

observed for similar patients in their treatment group who also discontinued treatment. To do so 'off-treatment' Multiple Imputation could be used. This method requires available off-treatment data in the TD23 group (i.e. data observed for some patients in the TD23 group who stopped treatment early). If none or limited data observed off treatment in the TD23 group, refence based-Multiple Imputation could be used to impute the missing data following the behaviour of individuals observed in the placebo group.

Or a delta adjustment sensitivity analysis using either Multiple Imputation or a mean score approach could be employed to explore the impact of individuals with missing data having a poorer outcome that those observed in the trial.