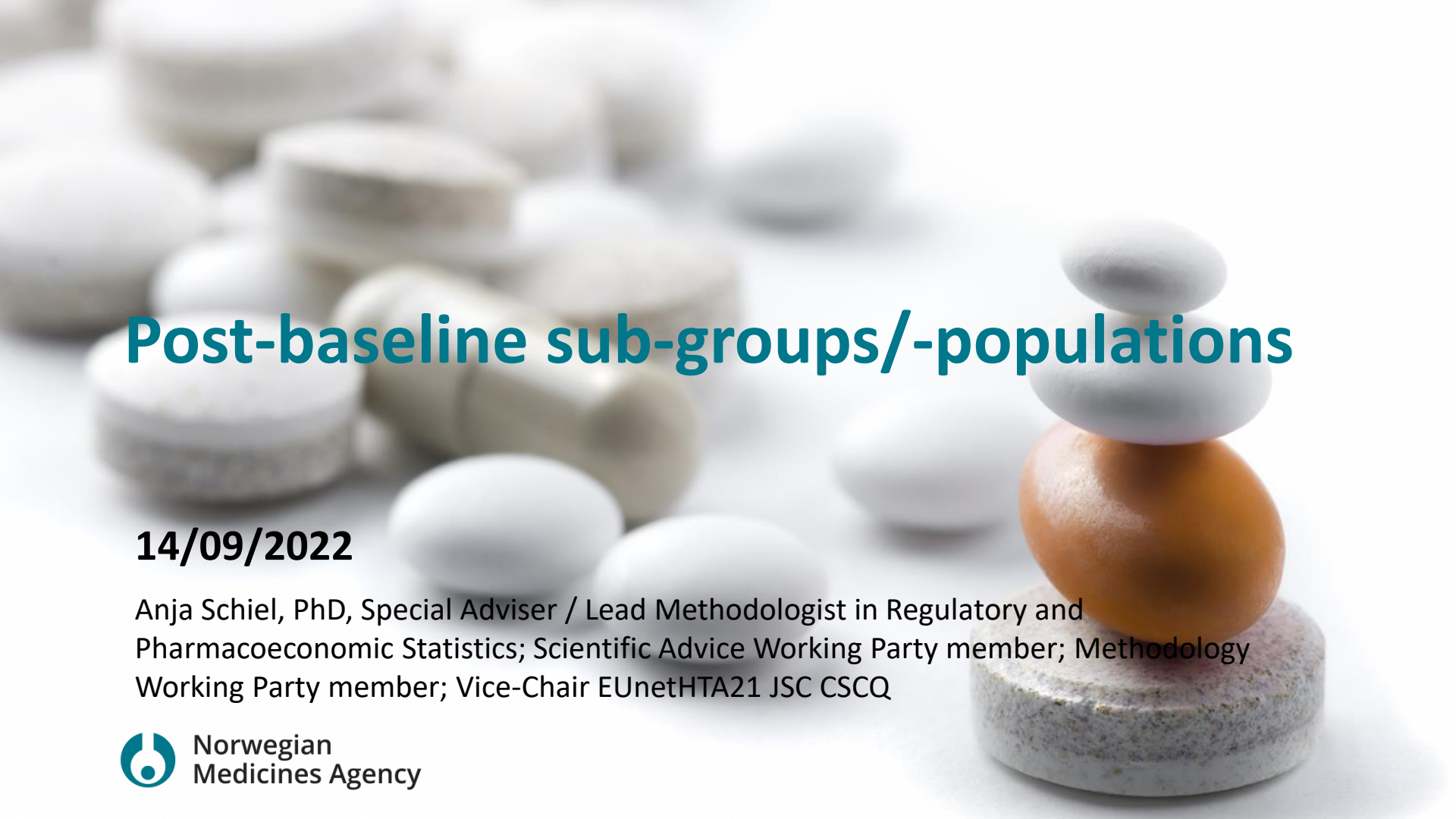


Post-baseline sub-groups/-populations

14/09/2022

Anja Schiel, PhD, Special Adviser / Lead Methodologist in Regulatory and Pharmacoeconomic Statistics; Scientific Advice Working Party member; Methodology Working Party member; Vice-Chair EUnetHTA21 JSC CSCQ

Norwegian Medicines Agency

Disclaimer

The views expressed are those of the presenter and should not be understood or quoted as being made on behalf of the Norwegian Medicines Agency (NoMA) or of the European Medicines Agency (EMA) or its scientific committees or reflecting the position of the EMA.

Sub-groups or sub-populations

- Splitting hairs or do we need clarity?

- For my presentation I will stick to

  - Sub-groups are baseline defined and protected by randomisation
  - Sub-populations might be defined by randomisation but can also be defined by post-randomisation events/outcomes/changes

Building on the EMA GL

- In this guideline the term ‘subgroup’ will be used to refer to a subset of the clinical trial population defined by one or more intrinsic and extrinsic factors (see ICH-E5) of the patients under investigation, usually measured at baseline.
- The term ‘sub-population’ will be used to refer to a subset of the patient population described in the targeted therapeutic indication.
- **Post-baseline covariates may be affected by treatment received and will not usually be appropriate to define subgroups for the investigation of a treatment effect.**

Building on the EMA GL

- The possibility of false positive findings is often quoted as a reason to ignore or dismiss differential effects in a subgroup and its complement. Critically, this would mean not investigating the underlying hypothesis that effects across different subgroups are consistent with the overall outcome of the trial.
- When assessing results from a clinical trial there is the **additional risk that the balance afforded by randomisation is not fully preserved when looking into subgroups**, such that findings in one of multiple subgroups are more likely to be driven by baseline-imbalance in covariates between treatment groups than by an effect of treatment.

A warning but what if we really need this?

Trials

BioMed Central

Methodology

Open Access

Outcome based subgroup analysis: a neglected concern

Karim F Hirji[1] and Morten W Fagerland*[2]

Address: [1]Department of Epidemiology and Biostatistics, Muhimbili University of Health and Allied Sciences, Dar es Salaam, Tanzania and [2]Ullevål Department of Research Administration, Oslo University Hospital, Norway

Email: Karim F Hirji - kfhirji@aol.com; Morten W Fagerland* - morten.fagerland@medisin.uio.no

* Corresponding author

Published: 20 May 2009

Trials 2009, 10:33 doi:10.1186/1745-6215-10-33

This article is available from: http://www.trialsjournal.com/content/10/1/33

Received: 9 December 2008

Accepted: 20 May 2009

A warning but what if we really need this?

- A subgroup of clinical trial subjects identified by baseline characteristics is a **proper subgroup** while a subgroup determined by post randomization events or measures is an **improper subgroup**.
- ***The analysis of improper subgroups thereby not only flourishes in numerous disguised ways but also does so without a corresponding awareness of its pitfalls.***

Randomisation and causality

- Randomisation based inference allows one to judge evidence against the null of there being no treatment effect in the specific patients recruited into the trial.
- Is the population model based inference for trials appropriate when the notion that the patients recruited are a random sample does not align at all with the reality of how patients are recruited in most trials?

➢ This directly ties into the issue of generalisability, contextualisation and external validity.

Is randomisation just an illusion?

- We assume that the patients randomised to the two treatments are random samples from infinite populations of patients taking the two treatments.
- If that is not true (inclusion/exclusion come to mind) then the power of randomisation might be overemphasised?

➢ If the exercise is to predict ‘probabilities to experience an event’ or ‘treatment effect’ for future patients then how does randomisation in itself help us?

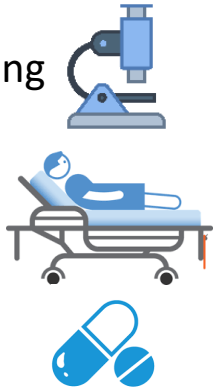
Who should decide to formulate ‘the question’ to begin with?

The difference in when and how we ask the question

- Clinical trial = Regulator

Efficacy (**B/R**)

- Does it work in experimental setting
- Population selected
- Placebo or a selected comparator

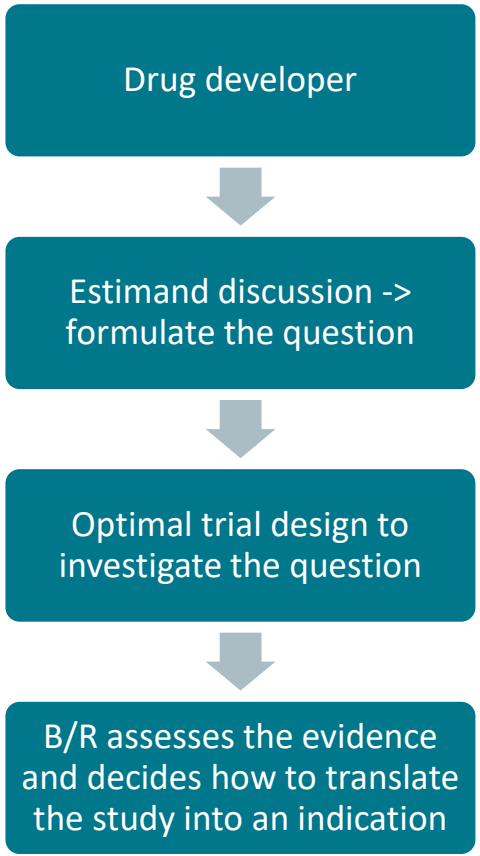


- Real world = HTA

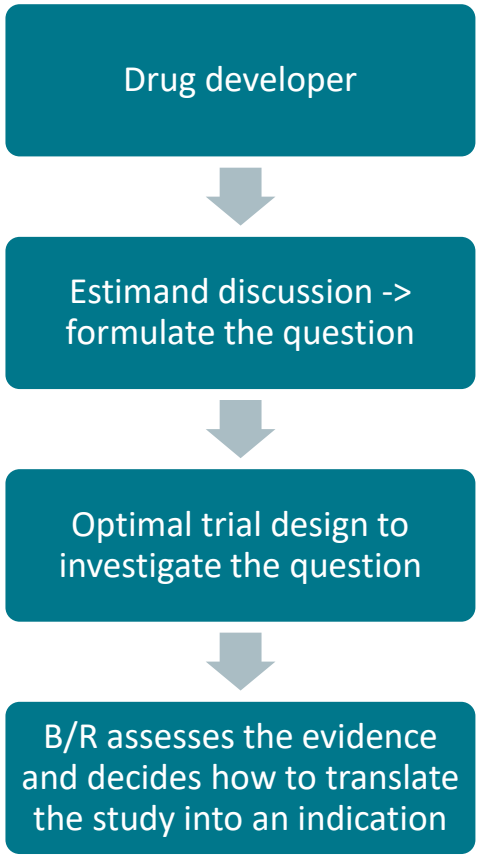
Relative Effectiveness (**C/E**)

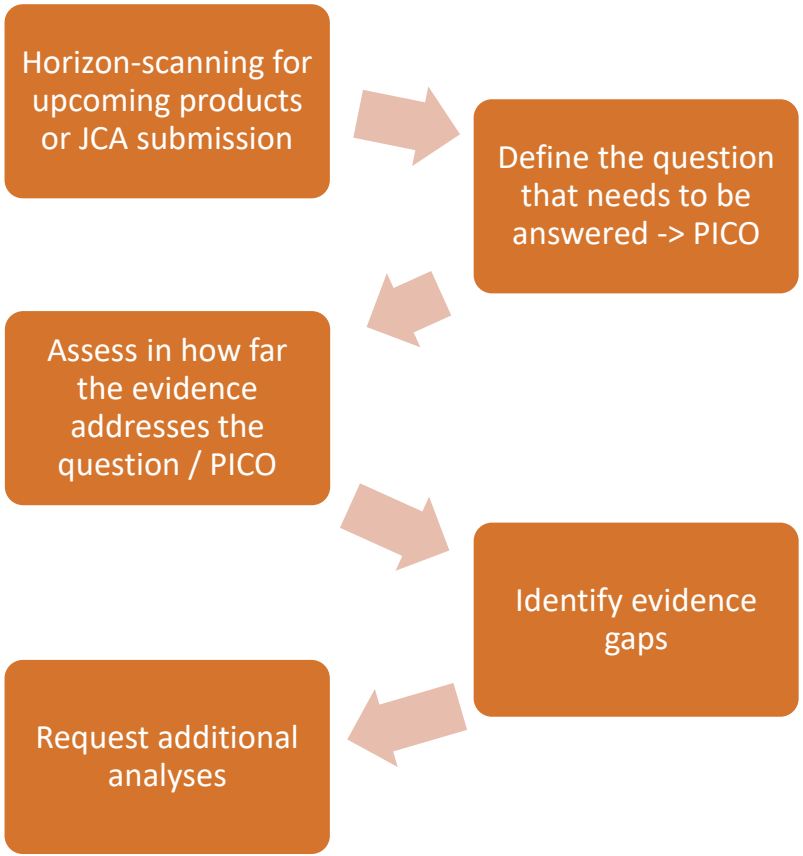
- How does it work in clinical practice
- Patients as they come
- Many alternative treatments

Process of defining the questions



Process of defining the questions





The difference in when and how we ask the question

- Drug Developer

Formulate the question about (B/R)

- Does it work in experimental setting
- Population selected
- Placebo or a selected comparator

- HTA

Relative Effectiveness (**C/E**)

- How does it work in clinical practice
- Patients as they come
- Many alternative treatments

Formulate the question the evidence has to support

Pareto improvement in social welfare

- Can the ‘gainer’ of some policy change compensate the ‘losers’ of the change?
- Within each product assessment we need to fully understand the winners and losers (example would be a responder analysis)
- This can include assessing baseline defined sub-groups but it can also include what HTAs call sub-populations (and these can be outcome defined or by specific criteria defined in the PICO).

For good or for bad

- Responder analysis
- Sub-populations defined by available/reimbursement of comparators
- Availability/clinical practice regarding subsequent therapies
- Treatment switch (especially if non-protocol/non-event driven)
- Start/stop/be on concomitant therapies not baseline defined

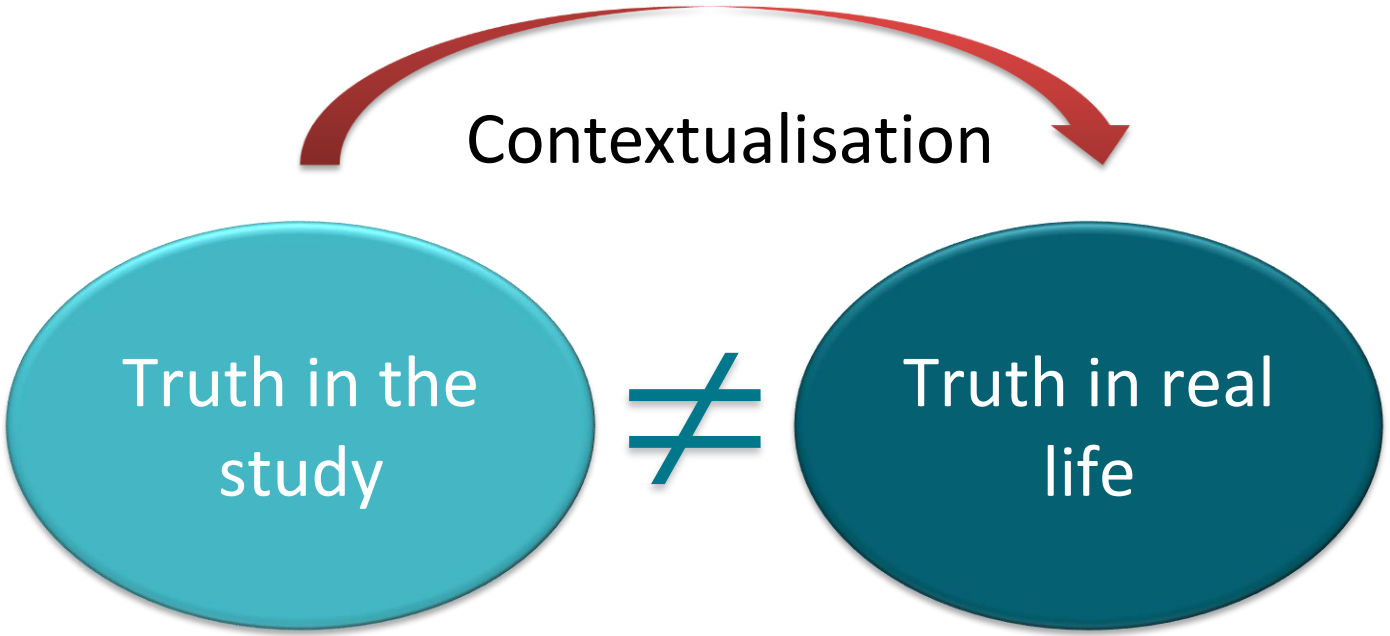
For good or for bad

- Responder analysis
- Sub-populations defined by available/reimbursement of comparators
- Availability/clinical practice regarding subsequent therapies
- Treatment switch (especially if non-protocol/non-event driven)
- Start/stop/be on concomitant therapies not baseline defined
- The analysis of improper subgroups thereby not only flourishes in numerous disguised ways but also does so **without a corresponding awareness of its pitfalls.**

Help is needed with the pitfalls

- The questions are legitimate because HTAs need to answer **additional** questions (for which the trial was not designed).
- They will request analyses that can be problematic (or some might even consider ‘stupid’)
- Statisticians need to not just point out **why not to do** a specific analysis but rather explain the **dangers involved**, support avoiding **wrong interpretations** and provided **better alternatives**.
- The PICO is predictable, hence such analyses can be discussed/encouraged/discouraged early on?

Internal versus External validity



Internal Validity

External Validity

Contextualisation

Truth in the study ≠ Truth in real life

Internal Validity

External Validity

Follow us

@legemiddelinfo

legemiddelverket

noma.no

Norwegian Medicines Agency