Subgroups and personalisation:

Richard Grieve
richard.grieve@lshtm.ac.uk

content

- Overview of the problem
- The limitation of standard approaches to generating evidence, resting on RCTs
- Outline of the 3 talks
- Future challenges

Towards personalization, why

- Drive to better target treatment for individual patients
- This is described as stratified, precision, or personalised medicine
- Right treatment for right patient at the right time
- Clinical decision-makers intend to practice personalisation
- Interest in estimating heterogeneity of treatment effects (THE)
- Thrombolysis for acute stroke can improve or harm
- Want to find 'true' positives
- If 'miss' subgroups waste resources, minimise 'false' negatives
- See Espinoza et al, 2014

The problem - multiplicity

- Concerns about overfitting and multiple-testing
- 20 tests, will judge 1 as 'statistically significant' just by chance
- Star sign 'found' modify effect aspirin after MI (see Horton 2001)
- If we do not take account of the number subgroups tested when selecting and estimating subgroup effects
  - Estimates of differences between subgroups over-estimated…
  - uncertainty in these estimates will be under-estimated
  - Concerns about falsely identifying subgroups (minimize false positive)

The concern about false positives

Summary of problem

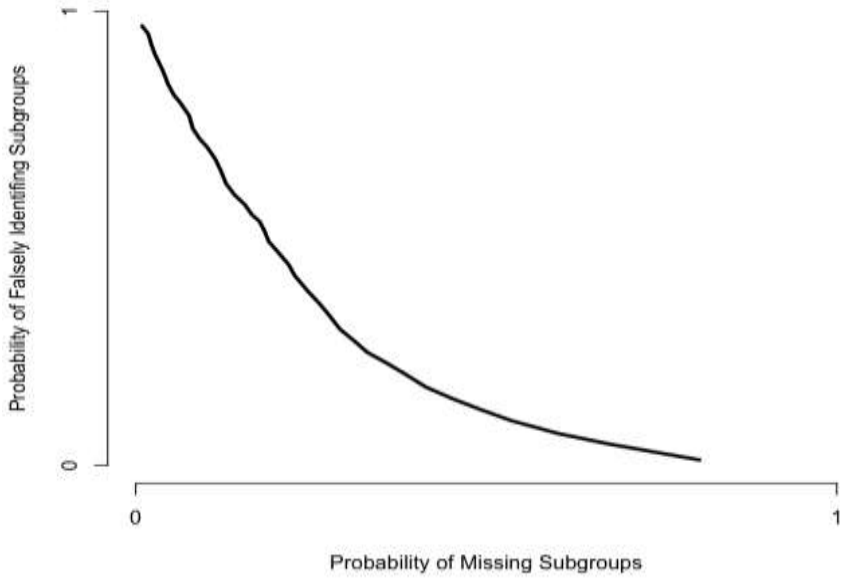
		Truth	
		Hetero	No Hetero
Claim	Hetero	✓	FALSLEY IDENTIFY SUBGROUPS
	No Hetero	MISS TRUE SUBGROUPS	✓

Approach might differ by context

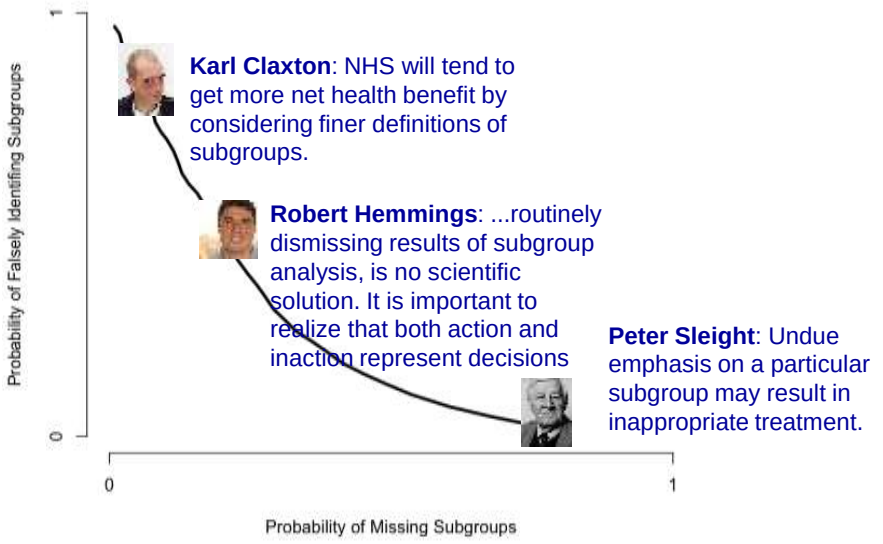
- Consider decision-making context
  - Inform regulatory decisions
  - Inform adoption and re-imbursement decisions
  - Inform clinical decision-making
  - Inform decisions regarding the conduct and design of future clinical studies

7

There are tradeoffs: where do you stand?



Are you a Karl, a Robert, or a Peter?



The poll

https://myispor.cnf.io/sessions/gjbs/#!/dashboard

Approach to subgroups.

Which one would you prefer?

1. 5% false negatives, 95% false positives [Karl]
2. 50% false negatives, 50% false positives [Robert]
3. 95% false negatives, 5% false positives [Peter]

Harnessing causal inference..
Predicting individual-level treatment effects
(Rubin 1977, Holland 1986)

- T_i is treatment indicator: 1 treatment group, 0 control
- Interested in causal relationship between T_i and Y_i
- Individual, i potential outcomes Y_{i0} and Y_{i1} under control and treated states
- Ideally observe treatment effect for each individual $\tau_i = Y_{i1} - Y_{i0}$
- BUT cannot observe both outcomes
- **Objective of methods: impute missing potential outcome**
- **Recognize each individual may have different response to treatment..**

11

RCTs inadequate

- Bradford Hill, 1960: 'do not answer the concern about the most likely outcome when this particular drug is given to a particular patient'
- Must recognize and estimate heterogeneous treatment effects

Problem 1: lack of power

- RCTs powered for main treatment effects, to give similar power for interaction effects, four-fold increase in sample size
- Subgroup interactions, 20 binary characteristics, 1 million subgroups
- Low proportion true positives, low power, high rate false discoveries

Problem 2: ignores effect modifiers not unobserved

High profile example

Comparison of ticagrelor with clopidogrel in patients with a planned invasive strategy for acute coronary syndromes (PLATO): a randomised double-blind study

Christopher P Cannon, Robert A Harrington, Stefan James, Diego Ardissino, Richard C Becker, Håkan Emanuelsson, Steen Husted, Hugo Katus, Matyas Keltai, Nardev S Khurmi, Frederic Kontny, Basil S Lewis, Philippe Gabriel Steg, Robert F Storey, Daniel Wojdyla, Lars Wallentin for the PLATelet inhibition and patient Outcomes (PLATO) investigators

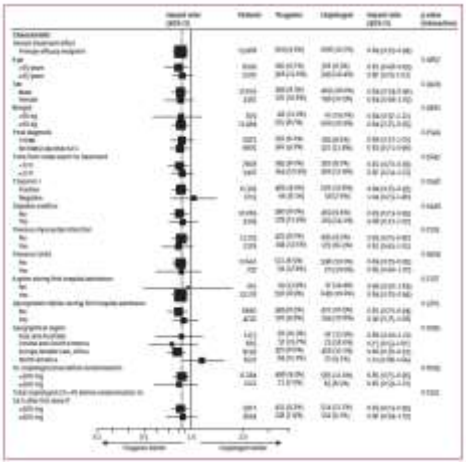
Summary

Background Variation in and irreversibility of platelet inhibition with clopidogrel has led to controversy about its optimum dose and timing of administration in patients with acute coronary syndromes. We compared ticagrelor, a more potent reversible P2Y12 inhibitor with clopidogrel in such patients.

Methods At randomisation, an invasive strategy was planned for 13 408 (72·0%) of 18 624 patients hospitalised for acute coronary syndromes (with or without ST elevation). In a double-blind, double-dummy study, patients were randomly assigned in a one-to-one ratio to ticagrelor and placebo (180 mg loading dose followed by 90 mg twice a day), or to clopidogrel and placebo (300–600 mg loading dose or continuation with maintenance dose followed by 75 mg per day) for 6–12 months. All patients were given aspirin. The primary composite endpoint was cardiovascular death, myocardial infarction, or stroke. Analyses were by intention to treat. This trial is registered with ClinicalTrials.gov, number NCT00391872.

Findings 6732 patients were assigned to ticagrelor and 6676 to clopidogrel. The primary composite endpoint occurred in fewer patients in the ticagrelor group than in the clopidogrel group (569 [event rate at 360 days 9·0%] *vs* 668 [10·7%], hazard ratio 0·84, 95% CI 0·75–0·94; p=0·0025). There was no difference between clopidogrel and ticagrelor groups in the rates of total major bleeding (691 [11·6%] *vs* 689 [11·5%], 0·99 [0·89–1·10]; p=0·8803) or severe bleeding, as defined according to the Global Use of Strategies To Open occluded coronary arteries, (198 [3·2%] *vs* 185 [2·9%], 0·91 [0·74–1·12]; p=0·3785).

Interpretation Ticagrelor seems to be a better option than clopidogrel for patients with acute coronary syndromes for whom an early invasive strategy is planned.



13

FDA Commentary

"Evidence efficacy in US population equivocal. ..PLATO not designed specifically to show evidence efficacy compared to clopidogrel in the US only.

Several potential explanatory factors explored including: compliance, statin exposure, low ticagrelor exposure, chance finding…. None satisfactorily explained the observed benefit of clopidogrel over ticagrelor in the USA.

"Although I consider the likelihood that the US/OUS was chance..
I believe evidence aspirin dose explains difference powerful further basis for approval..."

By a 7 to 1 vote FDA recommended approval

14

Warren Stephens

- Questioning current methods of evidence generation
- Highlighting the trade-offs that are made implicitly with the limitations of current data
- How to better align incentives for regulators/payers/clinicians to make more efficient (equitable) decisions

Neil Hawkins

- Vital to incorporate prior beliefs about anticipated subgroup effects
- Incorporate that into fully Bayesian approaches.
- Bayesian approach relies on ‘valid’ expert elicitation

Anirban Basu

- Heterogeneity according to unobserved and observed factors
- Instrumental variable approach can fully explore heterogeneity
- Harnesses this with large-scale RWE
- Rests on valid, continuous instruments, large data, and some parametric assumptions

Future research agendas

- Want to push out the ROC curve..
- Require adaption of appropriate methods from causal inference
- Machine learning (e.g. LASSO), try and avoid false positives.
- See also Imbens and Athey 'honest confidence intervals'


Recursive partitioning for heterogeneous causal effects

- RWE alone will not save the day
- Careful testing and evaluation of methods also essential

References

Espinoza, MA, Manca, A, Claxton, K, & Sculpher MJ. (2014). The Value of Heterogeneity for Cost-Effectiveness Subgroup Analysis: Conceptual Framework and Application. Medical Decision Making, 34(8), 951–964. https://doi.org/10.1177/0272989X14538705

Horton R. From star signs to trial guidelines (2000). Lancet 355:1033-34

Ioannidis JPA (2005) Why Most Published Research Findings Are False. PLoS Med 2(8): e124. https://doi.org/10.1371/journal.pmed.0020124

Kent D et al (2018). Personalised approaches to heterogeneous treatment effects. BMJ (in press)

Athey S; Imbens G (2016). Recursive partitioning for heterogeneous causal effects. Proc Natl Acad Sci USA 113(27):7353-60 (ISSN: 1091-6490)