Statistical Analysis of Biomarker Data



Ontario Clinical Oncology Group Escarpment Cancer Research Institute Department of Oncology, McMaster University Level II Investigator, Ontario Institute for Cancer Research



10 August 2017

Conflicts

None





Fig 2. Forest plots of hazard ratios (bevacizumab plus chemotherapy *v* placebo plus chemotherapy) for (A) overall survival and (B) progression-free survival by biomarker (dichotomized by median value or human epithelial growth factor receptor **VHERV states**) (HER2-positive samples had an immunohistochemistry score of 3+ or were fluorescent in situ hybridization positive. BL, baseline; epidermal growth factor receptor; NRP1, neuropilin-1; VEGFA, vascular endothelial growth factor-A; VEGFR1, VEGFR1, VEGFR2, VEGFR2, VEGF receptor-2.

Published in: Eric Van Cutsem; Sanne de Haas; Yoon-Koo Kang; Atsushi Ohtsu; Niall C. Tebbutt; Jian Ming Xu; Wei Peng Yong; Bernd Langer; Paul Delmar; Stefan J. Scherer; Manish A. Shah; JCO 2012, 30, 2119-2127. DOI: 10.1200/JCO.2011.39.9824 Copyright © 2012

I have a new biomarker!!!

- 1. I am going to get a NEJM publication!
- 2. I am going to save thousands of lives!
- 3. Only select patients will get treatment (saving thousands of people from unnecessary treatment)!
- 4. I will patent this and get rich!





Remember

• 1 in 20 tests 'significant by chance alone' if no difference

- Is yours a chance observation, or real?
- Need to validate what is involved?



Replication (Ki-67)

- Most often over-looked detail
- Will you get the same result if repeated?
- Take multiple samples from the same patient over short/moderate time period
- May expect change from treatment
- Coefficient of Variation = σ/μ *100%



Plausibility

Does it make biological sense?

• Is it clinically important (and not just statistically significant)?

• Pearson ρ – correlation coefficient assuming normal data. Spearman ρ for ranked data

• -1 (perfect negative relationship) to +1 (perfect **McMaster** positive relationship). 0=no relationship University











Jürgen Hetzel et al. Arterioscler Thromb Vasc Biol.

2005;25:1804-1809

Copyright © American Heart Association, Inc. All rights reserved.

Regression

- Is marker prognostic for outcome?
- Linear (continuous outcome)
- Logistic (categorical outcome)
- Cox (time-to-event outcome)
- Univariable: Is there an association?



Confounding

• One variable affects both covariate and outcome causing a spurious relationship



Interaction

• The effect of two or more variables is not simply additive



Multivariable Regression

- Does it provide *new* information
- Adjust for known factors
- How best to pick factors?



Factor Selection

- Automated (forward/backward/stepwise)
- Dataset specific. May omit important factors / include spurious factors
- Clinical judgement (good)
- Full model all possible factors. Significance => potentially *new* information.



Concordance

• Measure of discrimination ability

• Probability randomly selected person with outcome will have higher score than randomly selected person without outcome

• Ability of model to distinguish between those with and without outcome

• Nguyen & Kattan (2011) Eur Urol: c-stati Master University should improve >0.015 for new biomarker HEALTH SCIENCES

Calibration

• Agreement between predicted and observed outcomes



Re-Classification Plot





Validation

• Bootstrap

• Randomly sample population from dataset with replacement

• Sample size=724. Randomly sample 724 patients from dataset and re-calculate. Repeat 2000 times

• Gives measure of variability of outcomes



External Validation

• Even bootstrap is taking sample from same population

• Treatment patterns different for Canadians vs Americans

• Biomarker effect may be different

• Generalizability



Prognostic vs Predictive Marker

• Predictive marker can only be assessed using RCT data

• Treatment effect is different for patients depending on marker status

• Treatment effect only measurable if some patients receive treatment / control

• Prognostic marker more common, but us Manaster less valuable (e.g. age, sex) HEALTH SCIENCES

REMARK Guidelines

 McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM. Reporting recommendations for tumour MARKer prognostic studies (REMARK) Br J Cancer. 2005;93:387–391. doi: 10.1038/sj.bjc.6602678

- What should be *reported* in publication
- Similar to CONSORT guidelines

