Landsittel, Intern Med 2012, 2:3
DOI: 10.4172/2165-8048.1000e109

Internal Medicine: Open Access

Editorial Open Access

Addressing Statistical Requirements and Practical Limitations in Development of Biomarker Panels and Prognostic Models

Douglas Landsittel*

Department of Medicine, Center for Research on Health Care, University of Pittsburgh, Pittsburgh, Pennsylvania, USA

Keywords: Prediction; Prognostic models; Biomarkers; Study design; Model validation

Introduction

A key aspect of disease prevention and early detection for improved treatment, across a range of medical disciplines, is the development and implementation of prognostic models. Formulating such models, however, involves a range of practical challenges and associated statistical and epidemiological concerns. Although analytical approaches are generally well described for proper design, model development, and further analysis and validation of prognostic models, these methods and approaches are often poorly implemented. Overcoming these challenges must involve a combination of utilization of appropriate methodology (and statistical expertise) and better integration of resources across studies.

Discussion

Prognostic modeling generally refers to prediction of outcomes in the absence of treatment, whereas predictive modeling is specific to predicting treatment response [1], although both are often used synonymously [2]. Successful application of such models is critical for identifying high risk patients, ordering appropriate diagnostic tests, detecting early stage disease, and personalizing subsequent treatment regimens [3,4]. Prognostic models are typically developed and implemented using standard multivariable regression models, more complex modern regression methods, and/or associated statistical tools, such as nomograms [5], and incorporate multiple factors, such as demographics, occupational or environmental exposures, genetics, and/or other biomarkers (although they may be based on a single measure such as Prostate-Specific Antigen (PSA) for prostate cancer or CA-125 for ovarian cancer) into prediction of disease risk or other clinical outcome (e.g. survival or disease-free survival). Clinical prognostic or prediction rules have been developed (and sometimes validated) for a number of clinical applications; Toll, et al. found over 15,000 clinical prediction rules published in 2005 alone [6,7]. For cancer research, in particular, early detection and estimation of an individual's risk [8] (for breast, prostate, or other cancers) has evolved into a significant research topic with practical implications for counseling or other aspects of clinical decision making.

Despite the substantial potential for clinical utility of prognostic models, significant challenges exist in development and validation of such models, especially for those that utilize moderate or high dimensional biomarker panels. Erroneously optimistic findings often result from a range of study design flaws, such as prediction based on factors that systematically differ by case status, and thus lead to significant bias [9,10]. In terms of the progress toward validating and utilizing potential prognostic rules, the vast majority of published studies on prediction modeling focus on developing, or fitting the model, with many fewer studies assessing validation in a separate population; almost none of the published studies assessed actual impact on physician behavior or patient outcome [6]. In terms of biomarker-based prognostic rules, very few markers have been shown to have practical clinical utility [11]. These severe limitations are likely due to both 1) the practical challenges associated with and/or failure to recognize the importance of specific stages of developing and (externally) validating biomarker panels or other prognostic rules [2,10,12,13], which then leads to subsequent failure to advance promising prognostic rules through those stages of development and validation and 2) the poor design and reporting [11] of studies that are conducted. Substantial literature has been published critiquing even the most basic aspects of prognostic modeling, such as not justifying the selection of predictor variables and failing to validate or even cross-validate the model [14], with particular emphasis on erroneous findings from, and lack of reproducibility in studies of biomarker panels [15].

A positive aspect of this review is that many of the cited limitations and errors are entirely fixable (although many must be addressed prior to designing and conducting the study) through following existing guidelines for formulating and validating prognostic rules [10,11], conducting research in a reproducible manner [15], and implementing appropriate study designs for the given phase of research [12]. As an example of the latter concern, cross-sectional studies cannot be used to validate and/or show clinical utility, as validation and illustration of actual clinical utility can only follow from a prospective study with randomization to use or non-use of the given model and measurement of effectiveness outcomes (e.g. adoption within clinical practice, and morbidity and mortality outcomes). Another easily addressed but significant issue is appropriate use of key terms, such as reference to biomarkers as predictors versus surrogate measures, as the latter represents a far more stringent criterion [1]. Finally, more consistent inclusion of statistical expertise specific to those modeling and related statistical issues is necessary to optimize the value of information collected. For instance, testing the differences between areas under the Receiver Operating Characteristic (ROC) curve continues to be a commonly utilized approach for measuring the improvement in classification accuracy between two nested models, despite the fact that doing so leads to extremely low statistical power [16], and assessing the prognostic utility of new predictors is thus often better accomplished by more recently developed approaches, such as the net reclassification index or integrated discrimination improvement [17], although some debate exists concerning the resulting summary statistics and associated p-values [18].

Despite the availability of the above-mentioned fixes for many current weaknesses in development of prognostic rules, the steps

***Corresponding author:** Douglas Landsittel, Department of Medicine, Center for Research on Health Care, University of Pittsburgh, Pittsburgh, Pennsylvania, USA, Tel: 412-864-3019; Fax: 412-586-9672; E-mail: landsitteldp@upmc.edu

Received July 11, 2012; **Accepted** July 11, 2012; **Published** July 17, 2012

Citation: Landsittel D (2012) Addressing Statistical Requirements and Practical Limitations in Development of Biomarker Panels and Prognostic Models. Intern Med 2:e109. doi:10.4172/2165-8048.1000e109

of external validation and randomization, and the corresponding availability of multiple data sets from multiple sites is still critically necessary for validating models [1,2,19] and eventually incorporating them into clinical practice. Although internal validation, through methods such as cross-validation, may substantially reduce classification bias, they do not address any site-specific or data set-specific biases, such as measurement variability across different laboratories, or other less-easily identified factors. Addressing these factors requires the use of multi-site data, where validation, and subsequent predictions and inferences are truly unbiased. These challenges are further complicated by changes in technologies and associated measurements, as well as evolving clinical practices that may occur over time, and thus potentially necessitate further validation and model updating.

Conclusion

Development and utilization of biomarker panels and prognostic rules represents a highly significant and emerging topic with many applications across a wide range of medical disciplines. Although many challenges exist in properly developing, testing and validating these rules, many such issues are already highly addressable if given the needed attention to study design, statistical analysis methods, and interpretation of results specific to the given phase of biomarker or prognostic rule development. The most difficult, but also most necessary hurdle, however, is the sharing of data across studies and institutions to achieve external validation and overcome associated biases. Whether by formal consortiums and associated mechanisms [20,21], such as government-sponsored repositories [22], or through more informal collaborations between investigators, future studies must share resources across multiple data sets and sites to produce biomarker panels and prognostic rules that yield meaningful clinical utility.

References

1. Buyse M, Sargent DJ, Grothey A, Matheson A, de Gramont A (2010) Biomarkers and surrogate end points -- the challenge of statistical validation. Nat Rev Clin Oncol 7: 309-317.
2. Taylor JM, Ankerst DP, Andridge RR (2008) Validation of biomarker-based risk prediction models. Clin Cancer Res 14: 5977-5983.
3. Harrell FE Jr, Lee KL, Mark DB (1996) Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 15: 361-387.
4. Steyerberg EW (2008) Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. Springer, New York.
5. Shariat SF, Karakiewicz PI, Suardi N, Kattan MW (2008) Comparison of nomograms with other methods for predicting outcomes in prostate cancer: a critical analysis of the literature. Clin Cancer Res 14: 4400-4407.
6. Toll DB, Janssen KJ, Vergouwe Y, Moons KG (2008) Validation, updating and impact of clinical prediction rules: a review. J Clin Epidemiol 61:1085-1094.
7. Ingui BJ, Rogers MA (2001) Searching for clinical prediction rules in MEDLINE. J Am Med Inform Assoc 8: 391-397.
8. Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, et al. (1989) Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. J Natl Cancer Inst 81: 1879-1886.
9. Ransohoff DF (2005) Bias as a threat to the validity of cancer molecular-marker research. Nat Rev Cancer 5: 142-149.
10. Pepe MS, Feng Z, Janes H, Bossuyt PM, Potter JD (2008) Pivotal evaluation of the accuracy of a biomarker used for classification or prediction: standards for study design. J Natl Cancer Inst 100: 1432-1438.
11. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, et al. (2005) REporting recommendations for tumor MARKer prognostic studies (REMARK). Br J Cancer 93: 387-391.
12. Pepe MS, Etzioni R, Feng Z, Potter JD, Thompson ML, et al. (2001) Phases of biomarker development for early detection of cancer. J Natl Cancer Inst 93: 1054-1061.
13. Pepe MS (2003) The Statistical Evaluation of Medical Tests for Classification and Prediction. Oxford University Press, USA.
14. Perel P, Edwards P, Wenz R, Roberts I (2006) Systematic review of prognostic models in traumatic brain injury. BMC Med Inform Decis Mak 6: 38.
15. Tuma RS (2010) Calls for new reporting standards, quality control in microarrays. J Natl Cancer Inst 102: 1380-1381.
16. Demler OV, Pencina MJ, D'Agostino RB Sr (2012) Misuse of DeLong test to compare AUCs for nested models. Stat Med.
17. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS (2008) Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med 27: 157-172.
18. Pepe MS (2011) Problems with risk reclassification methods for evaluating prediction models. Am J Epidemiol 173: 1327-1335.
19. Bleeker SE, Moll HA, Steyerberg EW, Donders AR, Derksen-Lubsen G, et al. (2003) External validation is necessary in prediction research: a clinical example. J Clin Epidemiol 56: 826-832.
20. Wagner JA, Prince M, Wright EC, Ennis MM, Kochan J, et al. (2010) The Biomarkers Consortium:practice and pitfalls of open-source precompetitive collaboration. Clin Pharmacol Ther 87: 539-542.
21. Woosely RL, Myers RT, Goodsaid F (2010) The Critical Path Institute's approach to precompetitive sharing and advancing regulatory science. Clin Pharmacol Ther 87: 530-533.
22. Turner CF, Pan H, Silk GW, Ardini MA, Bakalov V, et al. (2011) The NIDDK Central Repository at 8 years--ambition, revision, use and impact. Database (Oxford).