Why must new cardiovascular risk factors be carefully re-assessed prior to clinical application?

We would like to comment on the clinical utility of biomarkers discussed in the article by Smulders et al.1 The authors stated that the ‘issue of causality [of biomarkers] is irrelevant’ but this may not be true if the ultimate goal of medical intervention is to decrease disease occurrence. In epidemiology, if a biomarker is significantly associated with an outcome in a longitudinal analysis, it is presumed that the relationship may be causal. However, not all significant biomarkers are causal factors; rather, some could be risk indicators accompanying a causal factor as in the case of microalbuminuria, troponins, and B-natriuretic protein described by Dr Wang.2 Although these non-cause risk indicators have value in predicting future events, as Dr Smulders stated, only changing the causal factors will confer any changes in disease morbidity and mortality. We agree with Dr Wang in that ‘careful assessment of evidence’ is required and ‘most current biomarkers are not ready for clinical application’.3 If we apply observational study results directly to clinical practice without further testing, the results are extrapolated beyond what the data support and may even be applied in the reverse direction.

We urge careful consideration of the relative risk, c-statistic, and Hosmer–Lemeshow statistics. A significantly high relative risk generated by a biomarker denotes that persons with high levels of the biomarker had a higher disease rate. It does not give evidence that reducing the level of this biomarker will reduce the occurrence of disease. Another test is needed to examine whether reducing the level of this biomarker will reduce the occurrence of disease. Changing the levels of highly significant biomarkers from observational studies has been proved not to change the outcome in many clinical trials.4 The discordance between these two study designs should not be interpreted as one being inferior to the other. Each provides valuable information by examining the disease–biomarker relationship from a different direction. Additionally, the relative risk is calculated by dividing the disease-proportion among those with high levels of the biomarker by the disease-proportion among those with low levels, ignoring the distribution of this biomarker among the non-diseased.5 Thus, a significantly high relative risk of a biomarker does not signify its clinical usefulness.5 The c-statistic simultaneously illustrates the distribution of a biomarker among the diseased (true positive) and the non-diseased (false positive); thus, it informs us how well this marker relates to disease status and provides more useful clinical insight than does the relative risk.5 The Hosmer–Lemeshow statistic, contrary to Dr Smulders confidence, will declare ‘no evidence of lack of fit’ if the test statistic falls within the range of 95 percentile when compared with the theoretical probability. This broad range of ‘goodness of fit’ makes this test unsuitable as a mean to assess precise model-fit. Furthermore, it does not offer any information about the model’s clinical performance.

Research evaluating new biomarkers is necessary to improve our understanding of pathophysiology. However, clinical application must be preceded by careful re-assessments utilizing other epidemiologic and statistical methods. This study was supported by the American Heart Association grant (no. 0635351N to S.-J.J.).

References
Why must new cardiovascular risk factors be carefully re-assessed prior to clinical application?: reply

Dr Sok-Ja Janket misquotes us, as we literally wrote: ‘From the perspective of risk prediction, the issue of causality is irrelevant...’. Risk stratification serves to identify high-risk patients, not to identify the variables that should be targets for treatment. In the future, optimal risk prediction models may turn out to contain variables that are not causal factors, and proven causal factors may even be excluded from simple but accurate risk models.

We agree that a biomarker that is useful for risk prediction is not necessarily an appropriate surrogate endpoint for an intervention study. For example, microalbuminuria predicts risk, but the better performance of angiotensin-converting enzyme inhibitors in reducing albuminuria does not universally translate in better risk reduction. Clearly, these points are central to the discussion, and we believe we covered them appropriately and in line with Dr Janket's views.

The limitations of relative risks and the c-statistic are discussed in our paper. The Hosmer–Lemeshow statistic is often misused as a hypothesis test. Rather, the $\chi^2$ value (and therefore the $P$-value) should better be used as a continuous parameter of overall goodness of fit.

We agree that calibration may change in risk ranges that are clinically irrelevant, and that the model’s clinical performance should be judged on correct (re-)classifications into (and out of) high-risk categories.

Yvo Smulders
UV University Medical Center
Internal Medicine
PO Box 7057
Amsterdam 1007 MB
The Netherlands
Tel: +31 20 44 44 307
Fax: +31 20 44 44 313
Email: y.smulders@vumc.nl.