Requirements for Biomarker Study proposals

Ancillary Study proposals to 1) develop a biomarker prediction model, 2) validate a model, 3) combination of model development and validation should be evaluated for the following items.

1. Focused aim and hypothesis

2. Preliminary data

3. Using the TRIPOD statement for guidance, provide:
   a. Review the TRIPOD statement for multivariable prediction models for diagnosis and prognosis (attached) to develop a **Research Plan for each Phase** that follows the terminology and suggestions.
   b. Include a schematic representation of the study design analogous to the CONSORT diagram in the SomaLogic JAMA 2016 paper and/or Figure 1 of the TRIPOD statement.
   c. Provide the classification of your planned prediction and validation models as presented in Figure 3 of the TRIPOD Statement, with an explanation.
   d. Please complete the 22-item checklist as presented in the TRIPOD statement, adding explanations for each items, if needed.
   e. The analysis component of the Research Plan should follow the proposed study design and be patterned after the format used in the JAMA 2016 SomeLogic paper.
   f. Specify and justify the number of each patient-study-visit samples with power calculations that will be used for:
      i. Model development
      ii. Model validation

4. Agreement to collaborate on the development of a subcontract, supplement, or another financial mechanism with the NIDDK NASH CRN to cover costs to the JHU DCC for sample selection, dataset preparation, consultation and validation analyses.

5. Agreement to collaborate on the development of a formal, written Authorship and Data Sharing Agreement between NASH CRN sponsor, the NASH CRN liaison, to address the following on behalf of the NASH CRN investigators:
   a. Collaboration on and authorship of primary papers arising from this work.
   b. Acknowledgment of and credit to the NIDDK NASH CRN investigators on all publications and public presentations of results from the samples provided.
   c. Agree to provide the DCC with models in enough detail so that the DCC can independently calculate the clinical prediction scores. The biomarker selection algorithms need to be defined by AS investigators, but the DCC will not attempt to replicate the selections, unless asked to do so.
   d. Agree to work with NASH CRN liaison and the NIDDK to develop and sign a Data Sharing Agreement that recognizes that ownership of the data resulting from this Ancillary Study will be jointly shared between the AS investigators and the NASH CRN investigators.
   e. Agree to keep all communications, details and plans of the study confidential unless prior approval is obtained from the NASH CRN Steering Committee prior to release of information.
   f. Agree that the NASH CRN clinical data is not released outside those investigators working on the study without prior authorization from the NASH CRN Steering Committee and the NIDDK.

6. If the Ancillary Study is to be funded by Industry, then either an agreement (e.g., CTA) with NIDDK or a formal contract must be signed between the NASH CRN liaison Investigator's site and the company. This should follow the NASH CRN Industry Collaboration template.

Annals of Internal Medicine RESEARCH AND REPORTING METHODS

Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD Statement

Gary S. Collins, PhD; Johannes B. Reitsma, MD, PhD; Douglas G. Altman, DSc; and Karel G.M. Moons, PhD

Prediction models are developed to aid health care providers in estimating the probability or risk that a specific disease or condition is present (diagnostic models) or that a specific event will occur in the future (prognostic models), to inform their decision making. However, the overwhelming evidence shows that the quality of reporting of prediction model studies is poor. Only with full and clear reporting of information on all aspects of a prediction model can risk of bias and potential usefulness of prediction models be adequately assessed. The Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) Initiative developed a set of recommendations for the reporting of studies developing, validating, or updating a prediction model, whether for diagnostic or prognostic purposes. This article describes how the TRIPOD Statement was developed. An extensive list of items based on a review of the literature was created, which was reduced after a Web-based survey and revised during a 3-day meeting in June 2011 with methodologists, health care professionals, and journal editors. The list was refined during several meetings of the steering group and in e-mail discussions with the wider group of TRIPOD contributors. The resulting TRIPOD Statement is a checklist of 22 items, deemed essential for transparent reporting of a prediction model study. The TRIPOD Statement aims to improve the transparency of the reporting of a prediction model study regardless of the study methods used. The TRIPOD Statement is best used in conjunction with the TRIPOD explanation and elaboration document. To aid the editorial process and readers of prediction model studies, it is recommended that authors include a completed checklist in their submission (also available at www.tripod-statement.org).

Ann Intern Med. 2015;162:55-63. doi:10.7326/M14-0697 **www.annals.org**

For author affiliations, see end of text.

For contributors to the TRIPOD Statement, see the **Appendix** (available at www.annals.org).

Editors' Note: In order to encourage dissemination of the TRIPOD Statement, this article is freely accessible on the Annals of Internal Medicine *Web site (www.annals.org) and will be also published in* BJOG, British Journal of Cancer, British Journal of Surgery, BMC Medicine, British Medical Journal, Circulation, Diabetic Medicine, European Journal of Clinical Investigation, European Urology, *and* Journal of Clinical Epidemiology. *The authors jointly hold the copyright of this article. An accompanying explanation and elaboration article, titled Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and Elaboration, is published as an online-only supplement at http://annals.org/article.aspx?doi=10.7326/M14-0698; the American College of Physicians holds the copyright to the online-only supplement.*

In medicine, patients with their care providers are confronted with making numerous decisions on the basis of an estimated risk or probability that a specific disease or condition is present (diagnostic setting) or a specific event will occur in the future (prognostic setting) (**Figure 1**). In the diagnostic setting, the probability that a particular disease is present can be used, for example, to inform the referral of patients for further testing, initiate treatment directly, or reassure patients that a serious cause for their symptoms is unlikely. In the prognostic setting, predictions can be used for planning lifestyle or therapeutic decisions based on the risk for developing a particular outcome or state of health within a specific period (1, 2). Such estimates of risk can also be used to risk-stratify participants in therapeutic clinical trials (3, 4).

In both the diagnostic and prognostic setting, estimates of probabilities are rarely based on a single predictor (5). Doctors naturally integrate several patient characteristics and symptoms (predictors, test results) to make a prediction (see **Figure 2** for differences in common terminology between diagnostic and prognostic studies). Prediction is therefore inherently multivariable. Prediction models (also commonly called "prognostic models," "risk scores," or "prediction rules" [6]) are tools that combine multiple predictors by assigning relative weights to each predictor to obtain a risk or probability (1, 2). Well-known prediction models include the Framingham Risk Score (7), Ottawa Ankle Rules (8), EuroScore (9), Nottingham Prognostic Index (10), and the Simplified Acute Physiology Score (11).

Prediction Model Studies

Prediction model studies can be broadly categorized as model development (12), model validation (with or without updating) (13) or a combination of both (**Figure 3**). Model development studies aim to derive a prediction model by selecting the relevant predictors and combining them statistically into a multivariable model. Logistic and Cox regression are most frequently used for short-term (for example, disease absent vs. present, 30-day mortality) and long-term (for example, 10-year risk) outcomes, respectively (12–14). Studies may also focus on quantifying the incremental

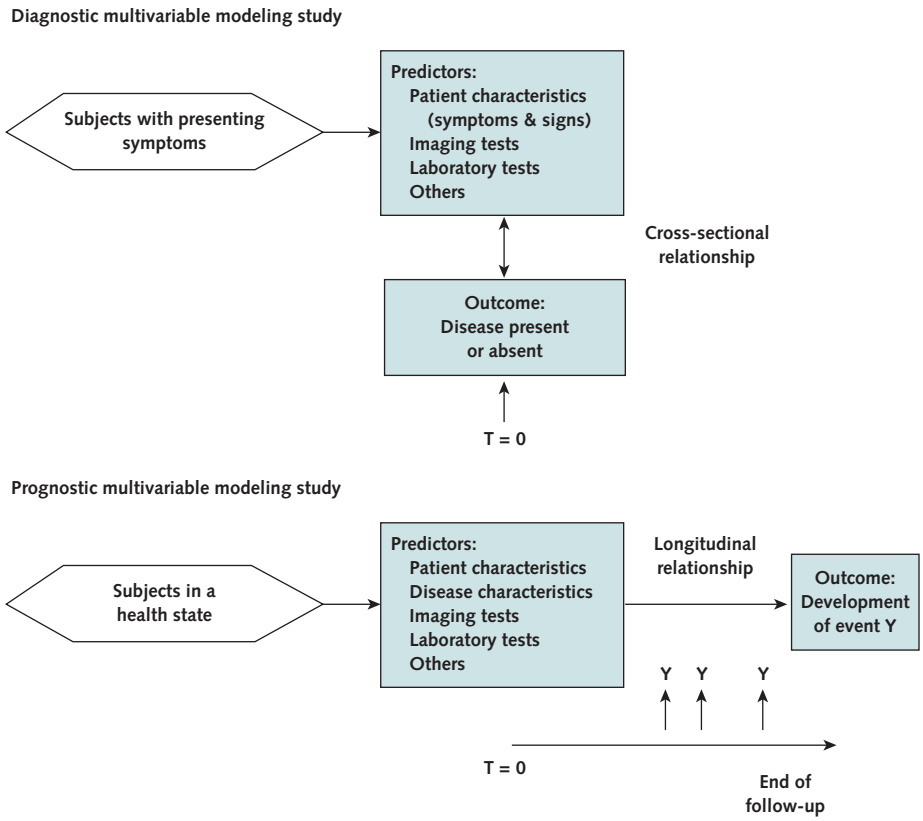
See also:

Editorial comment 73

Web-Only

Related Explanation and Elaboration article

Figure 1. Schematic representation of diagnostic and prognostic prediction modeling studies.



The nature of the prediction in diagnosis is estimating the probability that a specific outcome or disease is present (or absent) within an individual, at this point in time—that is, the moment of prediction (T = 0). In prognosis, the prediction is about whether an individual will experience a specific event or outcome within a certain time period. In other words, in diagnostic prediction the interest is in principle a cross-sectional relationship, whereas prognostic prediction involves a longitudinal relationship. Nevertheless, in diagnostic modeling studies, for logistical reasons, a time window between predictor (index test) measurement and the reference standard is often necessary. Ideally, this interval should be as short as possible and without starting any treatment within this period.

or added predictive value of a specific predictor (for example, newly discovered) to a prediction model (18).

Quantifying the predictive ability of a model on the same data from which the model was developed (often referred to as apparent performance) will tend to give an optimistic estimate of performance, owing to overfitting (too few outcome events relative to the number of candidate predictors) and the use of predictor selection strategies (19). Studies developing new prediction models should therefore always include some form of internal validation to quantify any optimism in the predictive performance (for example, calibration and discrimination) of the developed model. Internal validation techniques use only the original study sample and include such methods as bootstrapping or cross-validation. Internal validation is a necessary part of model development (2). Overfitting, optimism, and miscalibration may also be addressed and accounted for during the model development by applying shrinkage (for example, heuristic or based on bootstrapping techniques) or penalization procedures (for example, ridge regression or lasso) (20).

After developing a prediction model, it is strongly recommended to evaluate the performance of the model in other participant data than was used for the model development. Such external validation requires that for each individual in the new data set, outcome predictions are made using the original model (that is, the published regression formula) and compared with the observed outcomes (13, 14). External validation may use participant data collected by the same investigators, typically using the same predictor and outcome definitions and measurements, but sampled from a later period (temporal or narrow validation); by other investigators in another hospital or country, sometimes using different definitions and measurements (geographic or broad validation); in similar participants but from an intentionally different setting (for example, model developed in secondary care and assessed in similar participants but selected from primary care); or even in other types of participants (for example, model developed in adults and assessed in children, or developed for predicting fatal events and assessed for predicting nonfatal events) (13, 15, 17, 21, 22). In case of poor performance, the model can be updated or adjusted on the basis of the validation data set (13).

REPORTING OF MULTIVARIABLE PREDICTION MODEL STUDIES

Studies developing or validating a multivariable prediction model share specific challenges for researchers (6). Several reviews have evaluated the quality of published reports that describe the development or validation prediction models (23–28). For example, Mallett and colleagues (26) examined 47 reports published in 2005 presenting new prediction models in cancer. Reporting was found to be poor, with insufficient information described in all aspects of model development, from descriptions of patient data to statistical modeling methods. Collins and colleagues (24) evaluated the methodological conduct and reporting of 39 reports published before May 2011 describing the development of models to predict prevalent or incident type 2 diabetes. Reporting was also found to be generally poor, with key details on which predictors were examined, the handling and reporting of missing data, and model-building strategy often poorly described. Bouwmeester and colleagues (23) evaluated 71 reports, published in 2008 in 6 high-impact general medical journals, and likewise observed an overwhelmingly poor level of reporting. These and other reviews provide a clear picture that, across different disease areas and different journals, there is a generally poor level of reporting of prediction model studies (6, 23–27, 29). Furthermore, these reviews have shown that serious deficiencies in the statistical methods, use of small data sets, inappropriate handling of missing data, and lack of validation are common (6, 23–27, 29). Such deficiencies ultimately lead to prediction models that are not or should not be used. It is therefore not surprising, and fortunate, that very few prediction models, relative to the large number of models published, are widely implemented or used in clinical practice (6).

Prediction models in medicine have proliferated in recent years. Health care providers and policy makers are increasingly recommending the use of prediction models within clinical practice guidelines to inform decision making at various stages in the clinical pathway (30, 31). It is a general requirement of reporting of research that other researchers can, if required, replicate all the steps taken and obtain the same results (32). It is therefore essential that key details of how a prediction model was developed and validated be clearly reported to enable synthesis and critical appraisal of all relevant information (14, 33–36).

REPORTING GUIDELINES FOR PREDICTION MODEL STUDIES: THE TRIPOD STATEMENT

We describe the development of the TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis) Statement, a guideline specifically designed for the reporting of studies developing or validating a multivariable prediction model, whether for diagnostic or prognostic purposes. TRIPOD is not intended for multivariable modeling in etiologic studies or for studies investigating single prognostic factors (37). Furthermore, TRIPOD is also not intended for impact studies that quantify the impact of using a prediction model on participant or doctors' behavior and management, participant health outcomes, or cost-effectiveness of care, compared with not using the model (13, 38).

Reporting guidelines for observational (the STrengthening the Reporting of OBservational studies

Figure 2. Similarities and differences between diagnostic and prognostic prediction models.

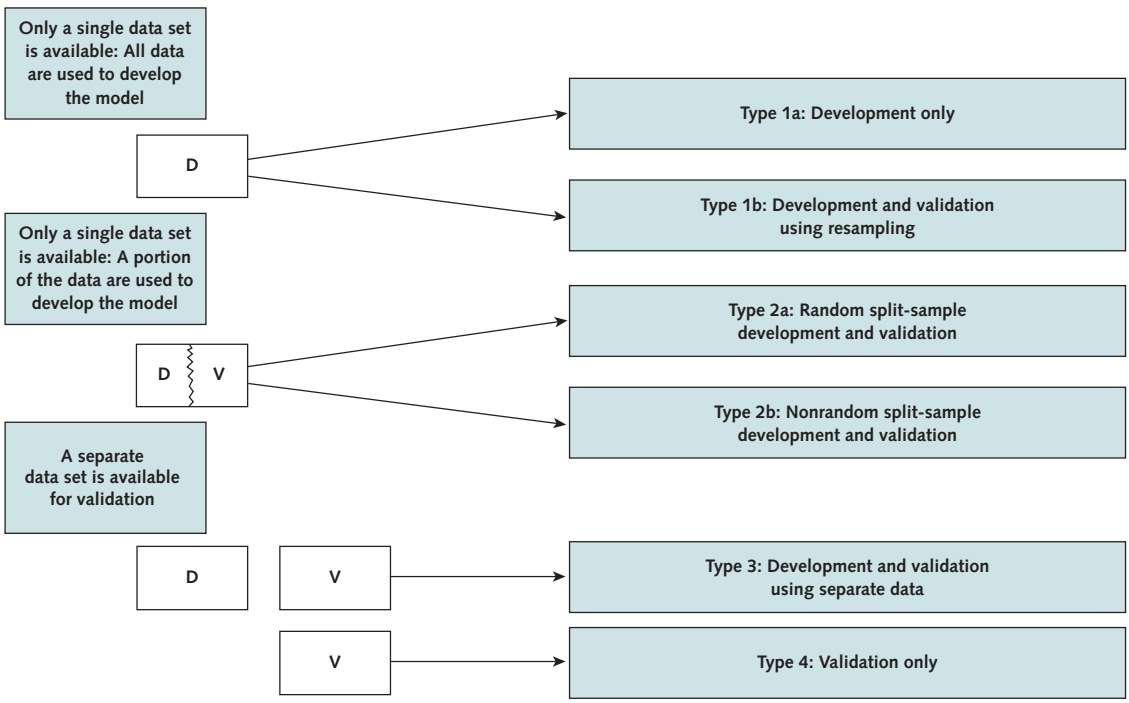
Despite the different nature (timing) of the prediction, there are many similarities between diagnostic and prognostic prediction models, including:

- **Type of outcome is often binary: either disease of interest present versus absent (in diagnosis) or the future occurrence of an event yes or no (in prognosis).**
- **The key interest is to generate the probability of the outcome being present or occurring for an individual, given the values of 2 or more predictors, with the purpose of informing patients and guiding clinical decision making.**
- **The same challenges as when developing a multivariable prediction model, such as selection of the predictors, model-building strategies, and handling of continuous predictors and the danger of overfitting.**
- **The same measures for assessing model performance.**

Different terms for similar features between diagnostic and prognostic modeling studies are summarized below.

Diagnostic Prediction Modeling Study		Prognostic Prediction Modeling Study
	Explanatory variables, predictors, covariates (X variables)	
Diagnostic tests or index tests		Prognostic factors or indicators
	Outcome (Y variable)	
Target disease/disorder (presence vs. absence)		Event (future occurrence: yes or no)
Reference standard and disease verification		Event definition and event measurement
	Missing outcomes	
Partial verification		Loss to follow-up and censoring

Figure 3. Types of prediction model studies covered by the TRIPOD Statement.



Analysis Type	Description
Type 1a	Development of a prediction model where predictive performance is then directly evaluated using exactly the same data (apparent performance).
Type 1b	Development of a prediction model using the entire data set, but then using resampling (e.g., bootstrapping or cross-validation) techniques to evaluate the performance and optimism of the developed model. Resampling techniques, generally referred to as "internal validation", are recommended as a prerequisite for prediction model development, particularly if data are limited (6, 14, 15).
Type 2a	The data are randomly split into 2 groups: one to develop the prediction model and one to evaluate its predictive performance. This design is generally not recommended or better than type 1b, particularly in case of limited data, because it leads to lack of power during model development and validation (14, 15, 16).
Type 2b	The data are nonrandomly split (e.g., by location or time) into 2 groups: one to develop the prediction model and one to evaluate its predictive performance. Type 2b is a stronger design for evaluating model performance than type 2a because it allows for nonrandom variation between the 2 data sets (6, 13, 17).
Type 3	Development of a prediction model using 1 data set and an evaluation of its performance on separate data (e.g., from a different study).
Type 4	The evaluation of the predictive performance of an existing (published) prediction model on separate data (13).

Types 3 and 4 are commonly referred to as "external validation studies." Arguably type 2b is as well, although it may be considered an intermediary between internal and external validation.

D = development data; V = validation data.

in Epidemiology [STROBE]) (39), tumor marker (REporting recommendations for tumour MARKer prognostic studies [REMARK]) (37), diagnostic accuracy (STAndards for the Reporting of Diagnostic accuracy studies [STARD]) (40), and genetic risk prediction (Genetic RIsk Prediction Studies [GRIPS]) (41) studies all contain many items that are relevant to studies developing or validating prediction models. However, none of these guidelines are entirely appropriate for prediction model studies. The 2 guidelines most closely related to prediction models are REMARK and GRIPS. However, the focus of the REMARK checklist is primarily on prognostic factors and not prediction models, whereas the GRIPS statement is aimed at risk prediction using genetic risk factors and the specific methodological issues around handling large numbers of genetic variants.

To address a broader range of studies, we developed the TRIPOD guideline: Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis. TRIPOD explicitly covers the development and validation of prediction models for both diagnosis and prognosis, for all medical domains and all

types of predictors. TRIPOD also places much more emphasis on validation studies and the reporting requirements for such studies. The reporting of studies evaluating the incremental value of specific predictors, beyond established predictors or even beyond existing prediction models (18, 42), also fits entirely within the remit of TRIPOD (see the accompanying explanation and elaboration document [43], available at www.annals.org).

DEVELOPING THE TRIPOD STATEMENT

We convened a 3-day meeting with an international group of prediction model researchers, including statisticians, epidemiologists, methodologists, health care professionals, and journal editors (from *Annals of Internal Medicine*, *BMJ*, *Journal of Clinical Epidemiology,* and *PLoS Medicine*) to develop recommendations for the TRIPOD Statement.

We followed published guidance for developing reporting guidelines (44) and established a steering committee (Drs. Collins, Reitsma, Altman, and Moons) to organize and coordinate the development of TRIPOD. We conducted a systematic search of MEDLINE, EMBASE, PsychINFO, and Web of Science to identify any published articles making recommendations on reporting of multivariable prediction models (or aspects of developing or validating a prediction model), reviews of published reports of multivariable prediction models that evaluated methodological conduct or reporting and reviews of methodological conduct and reporting of multivariable models in general. From these studies, a list of 129 possible checklist items was generated. The steering committee then merged related items to create a list of 76 candidate items.

Twenty-five experts with a specific interest in prediction models were invited by e-mail to participate in the Web-based survey and to rate the importance of the 76 candidate checklist items. Respondents (24 of 27) included methodologists, health care professionals, and journal editors. (In addition to the 25 meeting participants, the survey was also completed by 2 statistical editors from *Annals of Internal Medicine*.)

The results of the survey were presented at a 3-day meeting in June 2011, in Oxford, United Kingdom; it was attended by 24 of the 25 invited participants (22 of whom had participated in the survey). During the 3-day meeting, each of the 76 candidate checklist items was discussed in turn, and a consensus was reached on whether to retain, merge with another item, or omit the item. Meeting participants were also asked to suggest additional items. After the meeting, the checklist was revised by the steering committee during numerous face-to-face meetings, and circulated to the participants to ensure it reflected the discussions. While making revisions, conscious efforts were made to harmonize our recommendations with other reporting guidelines, and where possible we chose the same or similar wording for items (37, 39, 41, 45, 46).

TRIPOD COMPONENTS

The TRIPOD Statement is a checklist of 22 items that we consider essential for good reporting of studies developing or validating multivariable prediction models (**Table**). The items relate to the title and abstract (items 1 and 2), background and objectives (item 3), methods (items 4 through 12), results (items 13 through 17), discussion (items 18 through 20), and other information (items 21 and 22). The TRIPOD Statement covers studies that report solely development (12, 15), both development and external validation, and solely external validation (with or without updating), of a prediction model (14) (**Figure 3**). Therefore, some items are relevant only for studies reporting the development of a prediction model (items 10a, 10b, 14, and 15), and others apply only to studies reporting the (external) validation of a prediction model (items 10c, 10e, 12, 13c, 17, and 19a). All other items are relevant to all types of prediction model development and validation studies. Items relevant only to the development of a prediction model are denoted by *D*, items relating solely to validation of a prediction model are denoted by *V*, whereas items relating to both types of study are denoted *D;V*.

The recommendations within TRIPOD are guidelines only for reporting research and do not prescribe how to develop or validate a prediction model. Furthermore, the checklist is not a quality assessment tool to gauge the quality of a multivariable prediction model.

An ever-increasing number of studies are evaluating the incremental value of specific predictors, beyond established predictors or even beyond existing prediction models (18, 42). The reporting of these studies fits entirely within the remit of TRIPOD (see accompanying explanation and elaboration document [43]).

THE TRIPOD EXPLANATION AND ELABORATION DOCUMENT

In addition to the TRIPOD Statement, we produced a supporting explanation and elaboration document (43) in a similar style to those for other reporting guidelines (47–49). Each checklist item is explained and accompanied by examples of good reporting from published articles. In addition, because many such studies are methodologically weak, we also summarize the qualities of good (and the limitations of less good) studies, regardless of reporting (43). A comprehensive evidence base from existing systematic reviews of prediction models was used to support and justify the rationale for including and illustrating each checklist item. The development of the explanation and elaboration document was completed after several face-to-face meetings, teleconferences, and iterations among the authors. Additional revisions were made after sharing the document with the whole TRIPOD group before final approval.

Role of the Funding Source

There was no explicit funding for the development of this checklist and guidance document. The consen-

Table. Checklist of Items to Include When Reporting a Study Developing or Validating a Multivariable Prediction Model for Diagnosis or Prognosis*

Section/Topic	Item	Development or Validation?	Checklist Item	Page
Title and abstract				
Title	1	D;V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	
Abstract	2	D;V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	
Introduction				
Background and objectives	3a	D;V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	
	3b	D;V	Specify the objectives, including whether the study describes the development or validation of the model, or both.	
Methods				
Source of data	4a	D;V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation datasets, if applicable.	
	4b	D;V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	
Participants	5a	D;V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	
	5b	D;V	Describe eligibility criteria for participants.	
	5c	D;V	Give details of treatments received, if relevant.	
Outcome	6a	D;V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	
	6b	D;V	Report any actions to blind assessment of the outcome to be predicted.	
Predictors	7a	D;V	Clearly define all predictors used in developing the multivariable prediction model, including how and when they were measured.	
	7b	D;V	Report any actions to blind assessment of predictors for the outcome and other predictors.	
Sample size	8	D;V	Explain how the study size was arrived at.	
Missing data	9	D;V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	
	10c	V	For validation, describe how the predictions were calculated.	
	10d	D;V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	
Risk groups	11	D;V	Provide details on how risk groups were created, if done.	
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	
Results				
Participants	13a	D;V	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	
	13b	D;V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors, and outcome).	
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	
Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	
	15b	D	Explain how to use the prediction model.	
Model performance	16	D;V	Report performance measures (with CIs) for the prediction model.	
Model updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	
Discussion				
Limitations	18	D;V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	
Interpretation	19a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	
	19b	D;V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	
Implications	20	D;V	Discuss the potential clinical use of the model and implications for future research.	

Continued on following page

Table–Continued

Section/Topic	Item	Development or Validation?	Checklist Item	Page
Other information				
Supplementary information	21	D;V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and datasets.	
Funding	22	D;V	Give the source of funding and the role of the funders for the present study.	

* Items relevant only to the development of a prediction model are denoted by *D*, items relating solely to a validation of a prediction model are denoted by *V*, and items relating to both are denoted *D;V*. We recommend using the TRIPOD checklist in conjunction with the TRIPOD explanation and elaboration document.

sus meeting in June 2011 was partially funded by a National Institute for Health Research Senior Investigator Award held by Dr. Altman, Cancer Research UK, and the Netherlands Organization for Scientific Research. Drs. Collins and Altman are funded in part by the Medical Research Council. Dr. Altman is a member of the Medical Research Council Prognosis Research Strategy (PROGRESS) Partnership. The funding sources had no role in the study design, data collection, analysis, preparation of the manuscript, or decision to submit the manuscript for publication.

DISCUSSION

Many reviews have showed that the quality of reporting in published articles describing the development or validation of multivariable prediction models in medicine is poor (23–27, 29). In the absence of detailed and transparent reporting of the key study details, it is difficult for the scientific and health care community to objectively judge the strengths and weaknesses of a prediction model study (34, 50, 51). The explicit aim of this checklist is to improve the quality of reporting of published prediction model studies. The TRIPOD guideline has been developed to support authors in writing reports describing the development, validation or updating of prediction models, aid editors and peer reviewers in reviewing manuscripts submitted for publication, and help readers in critically appraising published reports.

The TRIPOD Statement does not prescribe how studies developing, validating, or updating prediction models should be undertaken, nor should it be used as a tool for explicitly assessing quality or quantifying risk of bias in such studies (52). There is, however, an implicit expectation that authors have an appropriate study design and conducted certain analyses to ensure all aspects of model development and validation are reported. The accompanying explanation and elaboration document describes aspects of good practice for such studies, as well as highlighting some inappropriate approaches that should be avoided (43).

TRIPOD encourages complete and transparent reporting reflecting study design and conduct. It is a minimum set of information that authors should report to inform the reader about how the study was carried out. We are not suggesting a standardized structure of reporting, rather that authors should ensure that they address all the checklist items somewhere in their article with sufficient detail and clarity.

We encourage researchers to develop a study protocol, especially for model development studies, and even register their study in registers that accommodate observational studies (such as ClinicalTrials.gov) (53, 54). The importance of also publishing protocols for developing or validating prediction models, certainly when conducting a prospective study, is slowly being acknowledged (55, 56). Authors can also include the study protocol when submitting their article for peer review, so that readers can know the rationale for including individuals into the study or whether all of the analyses were prespecified.

To help the editorial process; peer reviewers; and, ultimately, readers, we recommend submitting the checklist as an additional file with the report, indicating the pages where information for each item is reported. The TRIPOD reporting template for the checklist can be downloaded from www.tripod-statement.org.

Announcements and information relating to TRIPOD will be broadcast on the TRIPOD Twitter address (@TRIPODStatement). The Enhancing the QUAlity and Transparency Of health Research (EQUATOR) Network (www.equator-network.org) will help disseminate and promote the TRIPOD Statement.

Methodological issues in developing, validating, and updating prediction models evolve. TRIPOD will be periodically reappraised, and if necessary modified to reflect comments, criticisms, and any new evidence. We therefore encourage readers to make suggestions for future updates so that ultimately, the quality of prediction model studies will improve.

From Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, Botnar Research Centre, University of Oxford, Oxford, United Kingdom, and Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, the Netherlands.

Grant Support: There was no explicit funding for the development of this checklist and guidance document. The consensus meeting in June 2011 was partially funded by a National Institute for Health Research Senior Investigator Award held by Dr. Altman, Cancer Research UK (grant C5529), and the Netherlands Organization for Scientific Research (ZONMW 918.10.615 and 91208004). Drs. Collins and Altman are funded in part by the Medical Research Council (grant G1100513). Dr. Altman is a member of the Medical Research Council Prognosis Research Strategy (PROGRESS) Partnership (G0902393/99558).

Disclosures: Authors have disclosed no conflicts of interest. Forms can be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M14-0697.

Requests for Single Reprints: Gary S. Collins, PhD, Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, Botnar Research Centre, University of Oxford, Oxford OX3 7LD, United Kingdom; e-mail, gary.collins@csm.ox.ac.uk.

Current author addresses and author contributions are available at www.annals.org.

References

1. **Moons KG, Royston P, Vergouwe Y, Grobbee DE, Altman DG.** Prognosis and prognostic research: what, why, and how? BMJ. 2009; 338:b375. [PMID: 19237405]
2. **Steyerberg EW.** Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. New York: Springer; 2009.
3. **Dorresteijn JA, Visseren FL, Ridker PM, Wassink AM, Paynter NP, Steyerberg EW, et al.** Estimating treatment effects for individual patients based on the results of randomised clinical trials. BMJ. 2011; 343:d5888. [PMID: 21968126]
4. **Hayward RA, Kent DM, Vijan S, Hofer TP.** Multivariable risk prediction can greatly enhance the statistical power of clinical trial subgroup analysis. BMC Med Res Methodol. 2006;6:18. [PMID: 16613605]
5. **Riley RD, Hayden JA, Steyerberg EW, Moons KG, Abrams K, Kyzas PA, et al; PROGRESS Group.** Prognosis Research Strategy (PROGRESS) 2: prognostic factor research. PLoS Med. 2013;10: e1001380. [PMID: 23393429]
6. **Steyerberg EW, Moons KG, van der Windt DA, Hayden JA, Perel P, Schroter S, et al; PROGRESS Group.** Prognosis Research Strategy (PROGRESS) 3: prognostic model research. PLoS Med. 2013;10: e1001381. [PMID: 23393430]
7. **Anderson KM, Odell PM, Wilson PW, Kannel WB.** Cardiovascular disease risk profiles. Am Heart J. 1991;121(1 Pt 2):293-8. [PMID: 1985385]
8. **Stiell IG, Greenberg GH, McKnight RD, Nair RC, McDowell I, Worthington JR.** A study to develop clinical decision rules for the use of radiography in acute ankle injuries. Ann Emerg Med. 1992;21:384-90. [PMID: 1554175]
9. **Nashef SA, Roques F, Michel P, Gauducheau E, Lemeshow S, Salamon R.** European system for cardiac operative risk evaluation (EuroSCORE). Eur J Cardiothorac Surg. 1999;16:9-13. [PMID: 10456395]
10. **Haybittle JL, Blamey RW, Elston CW, Johnson J, Doyle PJ, Campbell FC, et al.** A prognostic index in primary breast cancer. Br J Cancer. 1982;45:361-6. [PMID: 7073932]
11. **Le Gall JR, Loirat P, Alperovitch A, Glaser P, Granthil C, Mathieu D, et al.** A simplified acute physiology score for ICU patients. Crit Care Med. 1984;12:975-7. [PMID: 6499483]
12. **Royston P, Moons KG, Altman DG, Vergouwe Y.** Prognosis and prognostic research: developing a prognostic model. BMJ. 2009; 338:b604. [PMID: 19336487]
13. **Moons KG, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, et al.** Risk prediction models: II. External validation, model updating, and impact assessment. Heart. 2012;98:691-8. [PMID: 22397946]
14. **Altman DG, Vergouwe Y, Royston P, Moons KG.** Prognosis and prognostic research: validating a prognostic model. BMJ. 2009;338: b605. [PMID: 19477892]
15. **Moons KG, Kengne AP, Woodward M, Royston P, Vergouwe Y, Altman DG, et al.** Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. Heart. 2012;98:683-90. [PMID: 22397945]
16. **Steyerberg EW, Harrell FE, Borsboom GJJM, Eijkemans MJC, Vergouwe Y, Habbema JDF.** Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. J Clin Epidemiol. 2001;54:774-81. [PMID: 11470385]
17. **Justice AC, Covinsky KE, Berlin JA.** Assessing the generalizability of prognostic information. Ann Intern Med. 1999;130:515-24. [PMID: 10075620]
18. **Steyerberg EW, Pencina MJ, Lingsma HF, Kattan MW, Vickers AJ, Van Calster B.** Assessing the incremental value of diagnostic and prognostic markers: a review and illustration. Eur J Clin Invest. 2012; 42:216-28. [PMID: 21726217]
19. **Steyerberg EW, Bleeker SE, Moll HA, Grobbee DE, Moons KG.** Internal and external validation of predictive models: a simulation study of bias and precision in small samples. J Clin Epidemiol. 2003; 56:441-7. [PMID: 12812818]
20. **Steyerberg EW, Eijkemans MJ, Habbema JD.** Application of shrinkage techniques in logistic regression analysis: a case study. Statistica Neerlandica. 2001;55:76-88.
21. **Reilly BM, Evans AT.** Translating clinical research into clinical practice: impact of using prediction rules to make decisions. Ann Intern Med. 2006;144:201-9. [PMID: 16461965]
22. **Wallace E, Smith SM, Perera-Salazar R, Vaucher P, McCowan C, Collins G, et al; International Diagnostic and Prognosis Prediction (IDAPP) group.** Framework for the impact analysis and implementation of clinical prediction rules (CPRs). BMC Med Inform Decis Mak. 2011;11:62. [PMID: 21999201]
23. **Bouwmeester W, Zuithoff NP, Mallett S, Geerlings MI, Vergouwe Y, Steyerberg EW, et al.** Reporting and methods in clinical prediction research: a systematic review. PLoS Med. 2012;9:1-12. [PMID: 22629234]
24. **Collins GS, Mallett S, Omar O, Yu LM.** Developing risk prediction models for type 2 diabetes: a systematic review of methodology and reporting. BMC Med. 2011;9:103. [PMID: 21902820]
25. **Collins GS, Omar O, Shanyinde M, Yu LM.** A systematic review finds prediction models for chronic kidney were poorly reported and often developed using inappropriate methods. J Clin Epidemiol. 2013;66:268-77. [PMID: 23116690]
26. **Mallett S, Royston P, Dutton S, Waters R, Altman DG.** Reporting methods in studies developing prognostic models in cancer: a review. BMC Med. 2010;8:20. [PMID: 20353578]
27. **Laupacis A, Sekar N, Stiell IG.** Clinical prediction rules. A review and suggested modifications of methodological standards. JAMA. 1997;277:488-94. [PMID: 9020274]
28. **Collins GS, de Groot JA, Dutton S, Omar O, Shanyinde M, Tajar A, et al.** External validation of multivariable prediction models: a systematic review of methodological conduct and reporting. BMC Med Res Methodol. 2014;14:40. [PMID: 24645774]
29. **Ettema RG, Peelen LM, Schuurmans MJ, Nierich AP, Kalkman CJ, Moons KG.** Prediction models for prolonged intensive care unit stay after cardiac surgery: systematic review and validation study. Circulation. 2010;122:682-9. [PMID: 20679549]
30. **Goff DC, Lloyd-Jones DM, Bennett G, Coady S, D'Agostino RB, Gibbons R, et al.** 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. Circulation. 2014;129:S49-73. [PMID: 24222018]
31. **Rabar S, Lau R, O'Flynn N, Li L, Barry P; Guideline Development Group.** Risk assessment of fragility fractures: summary of NICE guidance. BMJ. 2012;345:e3698. [PMID: 22875946]
32. **Laine C, Goodman SN, Griswold ME, Sox HC.** Reproducible research: moving toward research the public can really trust. Ann Intern Med. 2007;146:450-3. [PMID: 17339612]
33. **Siontis GC, Tzoulaki I, Siontis KC, Ioannidis JP.** Comparisons of established risk prediction models for cardiovascular disease: systematic review. BMJ. 2012;344:e3318. [PMID: 22628003]
34. **Seel RT, Steyerberg EW, Malec JF, Sherer M, Macciocchi SN.** Developing and evaluating prediction models in rehabilitation populations. Arch Phys Med Rehabil. 2012;93(8 Suppl):S138-53. [PMID: 22840880]

35. **Collins GS, Moons KG.** Comparing risk prediction models. BMJ. 2012;344:e3186. [PMID: 22628131]
36. **Knottnerus JA.** Diagnostic prediction rules: principles, requirements and pitfalls. Prim Care. 1995;22:341-63. [PMID: 7617791]
37. **McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM; Statistics Subcommittee of the NCI-EORTC Working Group on Cancer Diagnostics.** Reporting recommendations for tumor marker prognostic studies (REMARK). J Natl Cancer Inst. 2005;97:1180-4. [PMID: 16106022]
38. **Moons KG, Altman DG, Vergouwe Y, Royston P.** Prognosis and prognostic research: application and impact of prognostic models in clinical practice. BMJ. 2009;338:b606. [PMID: 19502216]
39. **von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative.** Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ. 2007;335:806-8. [PMID: 17947786]
40. **Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al; Standards for Reporting of Diagnostic Accuracy.** Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD Initiative. Radiology. 2003;226:24-8. [PMID: 12511664]
41. **Janssens AC, Ioannidis JP, van Duijn CM, Little J, Khoury MJ; GRIPS Group.** Strengthening the reporting of genetic risk prediction studies: the GRIPS statement. Eur J Clin Invest. 2011;41:1004-9. [PMID: 21434891]
42. **Tzoulaki I, Liberopoulos G, Ioannidis JP.** Use of reclassification for assessment of improved prediction: an empirical evaluation. Int J Epidemiol. 2011;40:1094-105. [PMID: 21325392]
43. **Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, et al.** Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med. 2015;162:W1-73.
44. **Moher D, Schulz KF, Simera I, Altman DG.** Guidance for developers of health research reporting guidelines. PLoS Med. 2010;16: e1000217. [PMID: 20169112]
45. **Little J, Higgins JP, Ioannidis JP, Moher D, Gagnon F, von Elm E, et al; STrengthening the REporting of Genetic Association Studies.** STrengthening the REporting of Genetic Association Studies (STREGA): an extension of the STROBE statement. PLoS Med. 2009; 6:e22. [PMID: 19192942]
46. **Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group.** Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med. 2009;151:264-9. [PMID: 19622511]
47. **Janssens AC, Ioannidis JP, Bedrosian S, Boffetta P, Dolan SM, Dowling N, et al.** Strengthening the reporting of genetic risk prediction studies (GRIPS): explanation and elaboration. Eur J Clin Invest. 2011;41:1010-35. [PMID: 21434890]
48. **Altman DG, McShane LM, Sauerbrei W, Taube SE.** Reporting recommendations for tumor marker prognostic studies (REMARK): explanation and elaboration. BMC Med. 2012;10:51. [PMID: 22642691]
49. **Vandenbroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, et al; STROBE Initiative.** Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. Epidemiology. 2007;18:805-35. [PMID: 18049195]
50. **Collins GS, Michaëlsson K.** Fracture risk assessment: state of the art, methodologically unsound, or poorly reported? Curr Osteoporos Rep. 2012;10:199-207. [PMID: 22688862]
51. **Järvinen TL, Jokihaara J, Guy P, Alonso-Coello P, Collins GS, Michaëlsson K, et al.** Conflicts at the heart of the FRAX tool. CMAJ. 2014;186:165-7. [PMID: 24366895]
52. **Moons KG, de Groot JA, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG, et al.** Critical appraisal and data extraction for the systematic reviews of prediction modelling studies: the CHARMS checklist. PLoS Med. 2014;11:e1001744. [PMID: 25314315]
53. **Williams RJ, Tse T, Harlan WR, Zarin DA.** Registration of observational studies: is it time? CMAJ. 2010;182:1638-42. [PMID: 20643833]
54. **Hemingway H, Riley RD, Altman DG.** Ten steps towards improving prognosis research. BMJ. 2009;339:b4184. [PMID: 20042483]
55. **Canadian CT Head and C-Spine (CCC) Study Group.** Canadian C-Spine Rule study for alert and stable trauma patients: I. Background and rationale. CJEM. 2002;4:84-90. [PMID: 17612425]
56. **Canadian CT Head and C-Spine (CCC) Study Group.** Canadian C-Spine Rule study for alert and stable trauma patients: II. Study objectives and methodology. CJEM. 2002;4:185-93. [PMID: 17609004]

Current Author Addresses: Drs. Collins and Altman: Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, Botnar Research Centre, University of Oxford, Oxford OX3 7LD, United Kingdom.
Drs. Reitsma and Moons: Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, PO Box 85500, 3508 GA Utrecht, the Netherlands.

Author Contributions: Conception and design: G.S. Collins, J.B. Reitsma, D.G. Altman, K.G.M. Moons.
Analysis and interpretation of the data: G.S. Collins, D.G. Altman, K.G.M. Moons.
Drafting of the article: G.S. Collins, J.B. Reitsma, D.G. Altman, K.G.M. Moons.
Critical revision of the article for important intellectual content: G.S. Collins, J.B. Reitsma, D.G. Altman, K.G.M. Moons.
Final approval of the article: G.S. Collins, J.B. Reitsma, D.G. Altman, K.G.M. Moons.
Provision of study materials or patients: G.S. Collins, K.G.M. Moons.
Statistical expertise: G.S. Collins, J.B. Reitsma, D.G. Altman, K.G.M. Moons.
Obtaining of funding: G.S. Collins, D.G. Altman, K.G.M. Moons.
Administrative, technical, or logistic support: G.S. Collins, K.G.M. Moons.
Collection and assembly of data: G.S. Collins, D.G. Altman, K.G.M. Moons.

Appendix: Members of the TRIPOD Group

Gary Collins (University of Oxford, Oxford, United Kingdom); Douglas Altman (University of Oxford, Oxford, United Kingdom); Karel Moons (University Medical Center Utrecht, Utrecht, the Netherlands); Johannes Reitsma (University Medical Center Utrecht, Utrecht, the Netherlands); Virginia Barbour (*PLoS Medicine*, United Kingdom and Australia); Nancy Cook (Division of Preventive Medicine, Brigham & Women's Hospital, Boston, Massachusetts); Joris de Groot (University Medical Center Utrecht, Utrecht, the Netherlands); Trish Groves (*BMJ*, London, United Kingdom); Frank Harrell, Jr. (Vanderbilt University, Nashville, Tennessee); Harry Hemingway (University College London, London, United Kingdom); John Ioannidis (Stanford University, Stanford, California); Michael W. Kattan (Cleveland Clinic, Cleveland, Ohio); André Knottnerus (Maastricht University, Maastricht, the Netherlands, and *Journal of Clinical Epidemiology*); Petra Macaskill (University of Sydney, Sydney, Australia); Susan Mallett (University of Oxford, Oxford, United Kingdom); Cynthia Mulrow (*Annals of Internal Medicine*, American College of Physicians, Philadelphia, Pennsylvania); David Ransohoff (University of North Carolina at Chapel Hill, Chapel Hill, North Carolina); Richard Riley (University of Birmingham, Birmingham, United Kingdom); Peter Rothwell (University of Oxford, Oxford, United Kingdom); Patrick Royston (Medical Research Council Clinical Trials Unit at University College London, London, United Kingdom); Willi Sauerbrei (University of Freiburg, Freiburg, Germany); Ewout Steyerberg (University Medical Center Rotterdam, Rotterdam, the Netherlands); Ian Stiell (University of Ottawa, Ottawa, Ontario, Canada); Andrew Vickers (Memorial Sloan Kettering Cancer Center, New York).

Correction: Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD Statement

A recent article (1) was erroneously published with the American College of Physicians copyright symbol. The ACP does not hold copyright on this manuscript.

This has been corrected in the online version.

Reference

1. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis TRIPOD): the TRIPOD statement. Ann Intern Med. 2015;162:55-63.

Original Investigation | INNOVATIONS IN HEALTH CARE DELIVERY

Development and Validation of a Protein-Based Risk Score for Cardiovascular Outcomes Among Patients With Stable Coronary Heart Disease

Peter Ganz, MD; Bettina Heidecker, MD; Kristian Hveem, MD, PhD; Christian Jonasson, PhD; Shintaro Kato, MS; Mark R. Segal, PhD; David G. Sterling, PhD; Stephen A. Williams, MD, PhD

IMPORTANCE Precise stratification of cardiovascular risk in patients with coronary heart disease (CHD) is needed to inform treatment decisions.

OBJECTIVE To derive and validate a score to predict risk of cardiovascular outcomes among patients with CHD, using large-scale analysis of circulating proteins.

DESIGN, SETTING, AND PARTICIPANTS Prospective cohort study of participants with stable CHD. For the derivation cohort (Heart and Soul study), outpatients from San Francisco were enrolled from 2000 through 2002 and followed up through November 2011 (≤11.1 years). For the validation cohort (HUNT3, a Norwegian population-based study), participants were enrolled from 2006 through 2008 and followed up through April 2012 (5.6 years).

EXPOSURES Using modified aptamers, 1130 proteins were measured in plasma samples.

MAIN OUTCOMES AND MEASURES A 9-protein risk score was derived and validated for 4-year probability of myocardial infarction, stroke, heart failure, and all-cause death. Tests, including the C statistic, were used to assess performance of the 9-protein risk score, which was compared with the Framingham secondary event model, refit to the cohorts in this study. Within-person change in the 9-protein risk score was evaluated in the Heart and Soul study from paired samples collected 4.8 years apart.

RESULTS From the derivation cohort, 938 samples were analyzed, participants' median age at enrollment was 67.0 years, and 82% were men. From the validation cohort, 971 samples were analyzed, participants' median age at enrollment was 70.2 years, and 72% were men. In the derivation cohort, C statistics were 0.66 for refit Framingham, 0.74 for 9-protein, and 0.75 for refit Framingham plus 9-protein models. In the validation cohort, C statistics were 0.64 for refit Framingham, 0.70 for 9-protein, and 0.71 for refit Framingham plus 9-protein models. Adding the 9-protein risk score to the refit Framingham model increased the C statistic by 0.09 (95% CI, 0.06-0.12) in the derivation cohort, and in the validation cohort, the C statistic was increased by 0.05 (95% CI, 0.02-0.09). Compared with the refit Framingham model, the integrated discrimination index for the 9-protein model was 0.12 (95% CI, 0.08-0.16) in the derivation cohort and 0.08 (95% CI, 0.05-0.10) in the validation cohort. In analysis of paired samples among 139 participants with cardiovascular events after the second sample, absolute within-person annualized risk increased more for the 9-protein model (median, 1.86% [95% CI, 1.15%-2.54%]) than for the refit Framingham model (median, 1.00% [95% CI, 0.87%-1.19%]) (P = .002), while among 375 participants without cardiovascular events, both scores changed less and similarly (P = .30).

CONCLUSIONS AND RELEVANCE Among patients with stable CHD, a risk score based on 9 proteins performed better than the refit Framingham secondary event risk score in predicting cardiovascular events, but still provided only modest discriminative accuracy. Further research is needed to assess whether the score is more accurate in a lower-risk population.

JAMA. 2016;315(23):2532-2541. doi:10.1001/jama.2016.5951

Editorial page 2525

Supplemental content at jama.com

Author Affiliations: Department of Medicine, University of California-San Francisco (Ganz, Heidecker); Division of Cardiology, San Francisco General Hospital (Ganz); Division of Cardiology, University of Zurich, Zurich, Switzerland (Heidecker); HUNT Research Center, Department of Public Health, NTNU, Levanger, Norway (Hveem, Jonasson); NEC Corporation of America, Irving, Texas (Kato); Department of Epidemiology and Biostatistics, University of California-San Francisco (Segal); SomaLogic, Boulder, Colorado (Sterling, Williams).

Corresponding Author: Peter Ganz, MD, Department of Medicine, University of California-San Francisco, 1001 Potrero Ave, Room 5G1, San Francisco, CA 94110 (peter.ganz@ucsf.edu).

Coronary heart disease (CHD) remains a leading cause of mortality and morbidity.[1] Despite the importance of risk assessment,[2] considerable room for improvement remains.[3] Genetic risk factors[4,5] and candidate proteins, such as C-reactive protein, have delivered only modest advances[2] and do not adequately enable precision medicine—management based on accurately stratified personal phenotyping.

A recent scientific statement from the American Heart Association predicted that proteomics will be transformative,[6] but the proteomic characterization of cardiovascular risk phenotypes in large populations requires a high-throughput technology. In this study, such a technology was applied, based on modified aptamers as binding reagents,[7] to quantify 1130 proteins in 2 prospective cohorts of participants with stable CHD. The objectives of this study were the following: (1) to evaluate a broader range of prognostic plasma protein biomarkers than previously possible; (2) to create a multiprotein model of biomarkers for prognostic stratification; (3) to validate the performance of the model in an external cohort[8]; (4) to assess the robustness of this model and key prognostic proteins within it to typical variations in sample collection and processing[9]; (5) to determine whether inclusion of this multiprotein panel in a risk score composed of traditional risk factors improves risk prediction; and (6) to determine from analysis of paired samples collected nearly 5 years apart whether the interval change in multiprotein panel risk score is greater among participants who experience a cardiovascular event after the second sample than among participants who do not. This study focused on participants with stable CHD because they have a broad range of risk that is not adequately identified by traditional risk factors.[10,11]

Methods

Study Populations

Studies in both cohorts were approved by the appropriate institutional review boards, and all participants provided written informed consent. The derivation cohort consisted of 938 baseline plasma samples from the Heart and Soul study—a prospective cohort of patients with stable CHD from 12 clinics in the San Francisco Bay Area (enrollment, September 2000-December 2002; last follow-up, November 2011). The Heart and Soul study included participants with history of myocardial infarction (MI), angiographic evidence of at least 50% stenosis in 1 or more coronary vessels, prior evidence of inducible ischemia by stress testing, or history of coronary revascularization. Participants were excluded if they had an MI within the previous 6 months, were unable to walk 1 block, or were planning to relocate from the local area within 2 years. From this cohort, a prognostic 9-protein model was constructed and then validated on 971 samples from HUNT3, a prospective population-based cohort study from Nord-Trøndelag County in Norway (enrollment, 2006-2008; last follow-up, April 2012).[12] HUNT3 participants were included who met Heart and Soul study inclusion criteria and had not had an MI within the previous 6 months. In the Heart and Soul study, race was self-identified in a questionnaire with categories of white, black, Asian, Latino, or other.[13] HUNT3 was a racially homogeneous cohort (≥98% white).[14] The information about race was used to discern whether the racial composition of the subset of participants with paired samples was similar to that of the overall Heart and Soul population in this study.

In contrast to the more standardized sample collection in the derivation cohort (fasted samples were collected at the same time of day and centrifuged and frozen within 1 hour of collection), sample collection in the validation cohort was more representative of likely clinical practice conditions: participants did not fast, and samples were collected at random times of day and processed (≤24 hours) after blood draw.

Changes in the 9-protein risk score were assessed by using paired samples from 514 participants in the Heart and Soul study in whom second plasma samples were taken a median 4.8 years after the first; participants had no cardiovascular events between these 2 samples. The study evaluated whether the second 9-protein risk score or the change from the baseline risk score could help to differentiate those participants who had a cardiovascular event after the second sample from those who did not. A flowchart of the sample and statistical process is shown in **Figure 1** and explained further in section 1 of the Supplement.

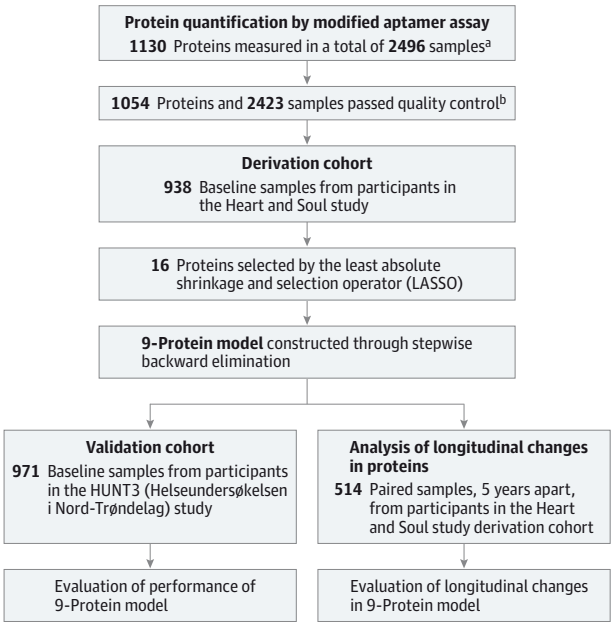
Quantification of Proteins in Human Plasma by Modified Aptamers

The method of quantification of proteins by modified aptamers has been previously described.[7,15,16] In brief, each of the 1130 individual proteins measured (eTable 1 in the Supplement) has its own binding reagent made of chemically modified DNA, referred to as modified aptamer.[7] Each sample of plasma was incubated with the mixture of modified aptamers to generate modified aptamer-protein complexes. Unbound modified aptamers and unbound or nonspecifically bound proteins were eliminated by 2 bead-based immobilization steps. After eluting the modified aptamers from the target protein, the fluorescently labeled modified aptamers were directly quantified on an Agilent hybridization array (Agilent Technologies). Calibrators were included so that the degree of fluorescence was a quantitative reflection of protein concentration. The 1054 proteins that passed quality control (eTable 1 in the Supplement) had median intraassay and interassay coefficient of variation of less than 5%. The key data processing steps, statistical modeling, and specific assessments are summarized in Figure 1.

Statistical Methods

The primary outcome in this study was defined as the first event among MI, stroke/transient ischemic attack (referred to as stroke), heart failure hospitalization, or all-cause death. Cox proportional hazards models were used to estimate the association between levels of individual proteins and risk of primary outcome. In single-variable analysis of an association of individual proteins with the primary outcome, Bonferroni-corrected significance levels were reported, adjusting for 1054 comparisons, resulting in a nominal significance level

Figure 1. Sample and Statistical Process for Evaluation of the 9-Protein Model



[a] Samples were sourced from the Heart and Soul study and the HUNT3 study.

[b] Proteins (n = 76) and samples (n = 73) that failed standard interrun and intrarun assay quality control acceptance metrics (section 1 of the Supplement) were deemed unfit for analysis.

($P < 4.74 \times 10^{-5}$). All other statistical tests were 2-sided using a nominal 5% significance level ($P < .05$). To construct the multiprotein risk model for primary outcome, the least absolute shrinkage and selection operator[17] (LASSO) was used for variable (protein) selection with the Cox model. This method penalized the sum of the absolute values of the regression coefficients leading to some coefficients shrinking to zero and thus simultaneously performed variable selection.[17-19] LASSO regularization level was chosen by cross-validation using the 1 standard error rule (section 3 in the Supplement). LASSO was used for variable selection only, with the fully parametric (Weibull) survival model as the final prognostic model. Stepwise backward elimination, starting from the set of LASSO-selected proteins, was used to remove proteins that were not significant predictors in the absence of the constraint imposed by the LASSO penalty using the Bayesian information criterion stopping criteria.

As a comparative reference for the multiprotein risk model, the variables from the Framingham secondary event risk model[20] were refit to the Heart and Soul derivation cohort (referred to as refit Framingham). This model included age, sex, total cholesterol, high-density lipoprotein cholesterol (HDL-C), diabetes, systolic blood pressure, and current smoking status.[20] The 4-year time horizon was retained, for which this risk score was originally validated.[20]

Model performance within each cohort was assessed by discrimination and calibration. For discrimination, both the C statistic[21] and discrimination slope[8] are reported. The category-free net reclassification index (NRI>0)[22] and integrated discrimination index (IDI)[8] were used to assess reclassification performance and improvement in discrimination over the refit Framingham model. Calibration performance was assessed with a calibration plot and summarized across the full range of risk scores using the Hosmer-Lemeshow statistic. Calibration-in-the-large is also reported—the difference between the observed 4-year event frequency and the mean predicted risk score. Both the refit Framingham and protein models were recalibrated (Section 4, eTables 2 and 3; eFigures 1 and 2 in the Supplement) for use in the validation cohort to enable an equal comparison and reduce the effect of miscalibration.[23,24] Distribution-free (nonparametric) 95% CIs were reported for median values and bootstrap intervals for point estimates of performance metrics when asymptotic intervals were not available.

Changes in risk score in paired samples were assessed using the Wilcoxon rank sum test comparing the within-person change for patients with and without events after their second blood sample. Within-person risk score differences were expressed, relative to the elapsed time between the 2 blood collections, and annualized. A likelihood ratio test was used to compare the fit of the augmented model and combining within-person change with the baseline proteomic risk score. All statistical computing was performed using the R Language for Statistical Computing (version 3.2.1).[25]

Results

Population Characteristics

The characteristics of the derivation and external validation cohorts are summarized in **Table 1**. There were fewer events in the validation cohort, primarily because of shorter follow-up.

Proteins Prognostic of Outcomes

At a Bonferroni significance level of 5%, corrected for 1054 comparisons, 200 proteins were associated with the primary outcome (145 positively and 55 negatively). The hazard ratios (HRs) and levels of statistical significance for these 200 prognostic proteins are listed in eTable 4 in the Supplement. In the construction of the risk model, the LASSO process selected 16 prognostic proteins, for which biological functions are listed in section 5.1 of the Supplement and HRs in the derivation and validation cohorts are shown in eFigure 3 in the Supplement. Stepwise backward elimination reduced these to the subset of 9 proteins used in the final prognostic model. The 9 proteins and their HRs are angiopoietin-2 (ANGPT2) (HR, 1.67 [95% CI,1.53-1.82]; $P < 1.00 \times 10^{-16}$), matrix metalloproteinase-12 (MMP12) (HR, 1.65 [95% CI, 1.50-1.80]; $P < 1.00 \times 10^{-16}$), chemokine (C-C motif) ligand 18 (CCL18) (HR, 1.47 [95% CI, 1.34-1.61]; $P = 1.11 \times 10^{-16}$), complement 7 (C7) (HR, 1.47 [95% CI, 1.36-1.59]; $P < 1.00 \times 10^{-16}$), α_1-antichymotrypsin complex (SERPINA3) (HR, 1.39 [95% CI, 1.28-1.51]; $P = 1.97 \times 10^{-14}$), angiopoietin-related protein 4 (ANGPTL4) (HR, 1.27 [95% CI, 1.18-1.37]; $P = 4.95 \times 10^{-11}$), troponin I (TNNI3) (HR, 1.27 [95% CI, 1.19-1.35]; $P = 1.02 \times 10^{-12}$), growth differentiation factor 11/8 (GDF8/11) (HR, 0.72 [95% CI, 0.57-0.69]; $P = 8.79E \times 10^{-9}$),

Table 1. Baseline Characteristics of the Study Cohorts

	Median (Interquartile Range)			
	Derivation Cohort (Heart and Soul)			Validation Cohort (HUNT3)
	All Participants (N = 938)	Subset With Follow-up Samples (n = 514)	Annualized Within-Person Change for Subset With Follow-up Samples[a]	All Participants (N = 971)
Follow-up, y	7.9 (3.5 to 9.0)	9.0 (8.4 to 9.9)		4.3 (3.9 to 4.9)
Age, y	67.0 (59.3 to 75.0)	66.0 (59.0 to 73.0)	1.0 (0.87 to 1.06)	70.2 (61.8 to 77.5)
Men, No. (%)	773 (82.4)	418 (81.3)		700 (72.1)
White, No. (%)	565 (60.2)	312 (60.7)		≥952 (≥98)
Black, No. (%)	151(16.1)	81(15.8)		
Asian, No. (%)	108(11.5)	64(12.5)		
Latino, No. (%)	82(8.7)	43(8.4)		
Diabetes, No. (%)	247 (26.4)	114 (22.2)		133 (13.7)
Current smoker, No. (%)	184 (19.7)	85 (16.6)		198 (21.4)
Events during follow-up period, No.	465	139		272
Time to event, y[b]	3.8 (1.7 to 6.8)	7.7 (6.5 to 8.9)[c]; 2.9 (1.7 to 4.1)[c]		2.1 (1.0 to 3.2)
BMI[d]	27.7 (24.8 to 31.2)	27.9 (25.23 to 30.9)	0.06 (−0.22 to 0.32)	28.0 (25.7 to 30.8)
HDL-C, mg/dL	43.0 (36.0 to 53.0)	44.0 (36.0 to 54.0)	0 (−1.05 to 1.28)	42.5 (38.7 to 54.1)[e]
LDL-C, mg/dL	99.0 (82.0 to 122.0)	99.0 (83.0 to 121.0)	−1.91 (−6.25 to 1.95)	[f]
Total cholesterol, mg/dL	171.0 (150.0 to 197.0)	173.0 (150.0 to 195.0)	−2.12 (−7.21 to 2.39)	174.0 (150.8 to 201.1)
Creatinine, mg/dL	1.0 (0.9 to 1.2)	1.0 (0.9 to 1.2)	0.02 (0 to 0.05)	1.0 (0.9 to 1.2)
CRP, mg/L	2.3 (1.0 to 4.9)	1.9 (0.8 to 4.0)	−0.07 (−0.36 to 0.10)	1.5 (0.7 to 3.3)
eGFR, mL/min[g]	73.9 (58.5 to 88.0)	76.4 (61.8 to 90.2)	−2.05 (−3.77 to −0.56)	68.4 (55.9 to 80.5)
Triglycerides, mg/dL	110.0 (74.0 to 167.0)	107.0 (71.0 to 161.0)	−2.04 (−9.13 to 3.44)	141.6 (106.2 to 194.7)
Systolic blood pressure, mm Hg	130 (120 to 144)	130.0 (120.0 to 140.5)	1.21 (−1.87 to 4.38)	133 (120 to 146)
Diastolic blood pressure, mm Hg	74 (68 to 80)	75 (68 to 80)	0.22 (−1.53 to 1.87)	73 (65 to 80)

Abbreviations: BMI, body mass index; CRP, C-reactive protein; eGFR, estimated glomerular filtration rate; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol.

SI conversion factors: To convert HDL-C, LDL-C, and total cholesterol from mg/dL to mmol/L, multiply by 0.0259; creatinine from mg/dL to μmol/L, multiply by 88.4; CRP from mg/L to nmol/L, multiply by 9.524; triglycerides from mg/dL to mmol/L, multiply by 0.0113.

[a] Annualized within-person change was calculated as the difference between values at baseline and paired second sample then divided by the elapsed time between the 2 clinical visits. Median collection time between baseline and paired second sample was 4.8 years.

[b] Calculation included only participants with events.

[c] First value is from the baseline sample and the second value is from the follow-up sample.

[d] BMI was calculated as weight in kilograms divided by height in meters squared.

[e] HDL-C was nonfasted.

[f] LDL-C was not available.

[g] eGFR was calculated using CKD-EPI 2009.

and α_2-antiplasmin (SERPINF2) (HR, 0.64 [95% CI, 0.59-0.71]; $P < 1.00 \times 10^{-16}$).

9-Protein Risk Score

The 9-protein risk score reflects the probability of a cardiovascular event occurring within 4-years and is given by risk score (Supplement, section 5.2):

$$\text{risk score} = 1 - e^{-e^{\left(\frac{Log(4)-PI}{0.85}\right)}},$$

where the prognostic index (PI) combines the measurements of the 9 proteins as follows:

$$\text{prognostic index} = 16.61 - 1.55 \times \text{ANGPT2} + 1.22 \times \text{GDF8/11} - 2.12 \times \text{C7} + 2.64 \times \text{SERPINF2} - 0.57 \times \text{CCL18} - 1.02 \times \text{ANGPTL4} - 1.43 \times \text{SERPINA3} - 0.72 \times \text{MMP12} - 0.59 \times \text{TNN13}.$$

Table 2 provides the estimated HRs and associated model coefficients for a Cox proportional hazards model based on the refit Framingham variables for the full duration of follow-up, with and without the addition of prognostic index from the 9-protein model. In the presence of the information from 9 proteins, most clinical variables remained as significant risk predictors except for HDL-C. Systolic blood pressure was not a significant risk predictor either in the refit Framingham model or with the addition of the 9 proteins. Adjusting the 9-protein prognostic index for the Framingham variables reduced its HR only modestly (eFigure 4 in the Supplement), suggesting that the 9 proteins contained prognostic information that was at least partly independent of traditional risk factors.

Proteomic Model Performance

Risk stratified survival curves of the 2 study populations are shown in **Figure 2**, illustrating that in both the derivation and validation cohorts, the participants had 4-year cumulative event rates of 60% to 80% in the 10th deciles and less than 10% in the first deciles. Discrimination performance

Table 2. Risk Prediction Models for Primary End Point of Myocardial Infarction, Stroke, Heart Failure, and Death[a]

	Framingham Variables Alone[b]			Framingham Variables[b] Plus 9-Protein Prognostic Index		
	HR (95% CI)	β	*P* Value	HR (95% CI)	β	*P* Value
Men	1.71 (1.26 to 2.32)	0.535	<.001	1.63 (1.20 to 2.20)	0.487	.002
Age, y	1.77 (1.58 to 1.99)	0.573	<.001	1.28 (1.13 to 1.44)	0.247	<.001
Total cholesterol, mg/dL	1.14 (1.03 to 1.26)	0.129	.01	1.20 (1.09 to 1.32)	0.178	<.001
HDL-C, mg/dL	0.88 (0.79 to 0.99)	−0.122	.03	0.95 (0.85 to 1.05)	−0.056	.28
Diabetes	1.84 (1.50 to 2.26)	0.611	<.001	1.44 (1.17 to 1.77)	0.363	<.001
Systolic blood pressure, mm Hg	1.03 (0.94 to 1.13)	0.029	.55	0.99 (0.90 to 1.08)	−0.014	.77
Current smoker	2.02 (1.58 to 2.58)	0.704	<.001	1.50 (1.16 to 1.94)	0.405	.002
9-Protein prognostic index	–	–	–	2.32 (2.08 to 2.58)	0.840	<.001

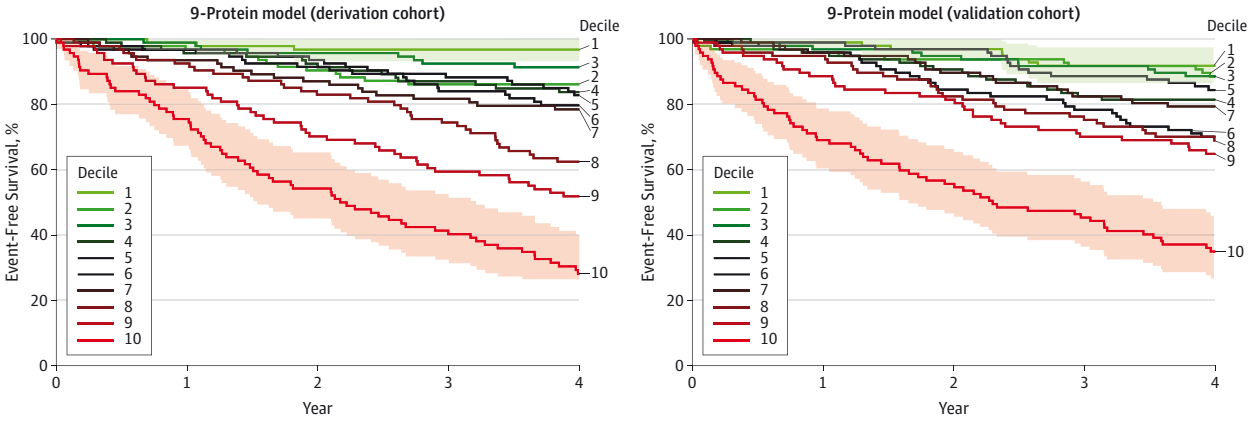
Abbreviation: HDL-C, high density lipoprotein cholesterol.

SI conversion factor: To convert HDL-C and total cholesterol from mg/dL to mmol/L, multiply by 0.0259.

[a] Continuous variables were standardized so hazard ratios reflect incremental change in hazard per 1 standard deviation change in predictor.

[b] Framingham variables were refit in the derivation cohort using a Cox proportional hazard model with and without the 9-protein prognostic index.

Figure 2. Event-Free Survival for End Points of Myocardial Infarction, Stroke, Heart Failure, and Death, Stratified by Deciles of the 9-Protein 4-Year Risk Score



No. at risk by decile

Decile					
1	94	93	90	90	86
2	94	91	85	81	79
3	94	93	89	84	82
4	93	89	85	80	74
5	94	91	86	81	76
6	94	91	88	78	73
7	93	87	81	76	73
8	94	87	78	69	57
9	94	80	66	55	48
10	94	71	51	37	25

No. at risk by decile

Decile					
1	98	97	95	90	78
2	97	93	91	89	72
3	97	94	92	89	80
4	97	93	88	80	74
5	97	96	94	86	71
6	97	93	82	76	64
7	97	93	87	80	65
8	97	92	80	73	60
9	97	86	79	68	57
10	97	67	54	44	31

The shading in the survival plots indicates 95% CI for the first and 10th deciles. Decile 1 indicates the lowest score; decile 10 indicates the highest score. Data defining the deciles of risk score are presented in Figure 3.

was assessed using the 4-year time horizon, the same as the original Framingham secondary event model.[20] **Table 3** lists the performance metrics for the refit Framingham model, the 9-protein model, and for the combination of both models. In the derivation cohort, the C statistic increased from 0.66 for the refit Framingham model to 0.74 (Δ C statistic, 0.09 [95% CI, 0.06-0.12]) for the 9-protein model alone and to 0.75 (Δ C statistic, 0.10 [95% CI, 0.08-0.12]) for the 9-protein model combined with the refit Framingham model. The discrimination slope was 0.09 (95% CI, 0.07-0.11) for the refit Framingham model, 0.21 (95% CI, 0.17-0.24) for the 9-protein model, and 0.23 (95% CI, 0.19-0.26) for the refit Framingham combined with the 9-protein model. When compared with refit Framingham, the 9-protein model had an IDI of 0.12 (95% CI, 0.08-0.16), which indicates an absolute increase of 12% in mean risk for participants with events compared with participants without events over the clinical variable model. The 9-protein model had an NRI(>0) of 0.52 (95% CI, 0.40-0.65), with event-specific components of 0.22 (95% CI, 0.11-0.36) and no event-specific components of 0.30 (95% CI, 0.22-0.36). In the validation cohort, inclusion of the 9-protein score

Table 3. Comparative Performance Metrics in Derivation and Validation Cohorts for Refit Framingham Model, 9-Protein Model, and Their Combination When Predicting Primary End Points of Myocardial Infarction, Stroke, Heart Failure, and Death

	Cohort	Refit Framingham Model	9-Protein Model	Refit Framingham Model Plus the 9-Protein Model
C statistic	Derivation	0.66 (0.63 to 0.68)	0.74 (0.72 to 0.77)	0.75 (0.73 to 0.78)
	Validation	0.64 (0.61 to 0.67)	0.70 (0.67 to 0.72)	0.71 (0.69 to 0.74)
Δ C statistic (derivation and validation)[a]	Both	0.01 (−0.01 to 0.04)	0.05 (0.03 to 0.07)	0.04 (0.02 to 0.07)
Discrimination slope	Derivation	0.09 (0.07 to 0.11)	0.21 (0.17 to 0.24)	0.23 (0.19 to 0.26)
	Validation	0.07 (0.05 to 0.08)	0.14 (0.12 to 0.17)	0.17 (0.14 to 0.20)
Δ Discrimination slope (derivation and validation)[a]	Both	0.02 (0 to 0.05)	0.07 (0.01 to 0.11)	0.06 (0.01 to 0.11)
Hazard Ratio (95% CI)				
Quintile[b]	Derivation	5.0 (3.60 to 6.94)	11.7 (8.08 to 16.86)	16.3 (10.69 to 24.93)
	Validation	6.6 (3.74 to 11.54)	7.6 (4.53 to 12.85)	9.8 (4.53 to 20.99)
Per standard deviation	Derivation	1.9 (1.72 to 2.15)	2.5 (2.27 to 2.73)	2.8 (2.49 to 3.05)
	Validation	1.7 (1.53 to 1.97)	2.1 (1.86 to 2.33)	2.2 (1.97 to 2.52)
Hosmer-Lemeshow[c]	Derivation	$6.8 (5.57 \times 10^{-1})$	$5.3 (7.25 \times 10^{-1})$	$3.5 (9.02 \times 10^{-1})$
	Validation	$23.5 (2.81 \times 10^{-3})$	$6.8 (5.62 \times 10^{-1})$	$9.7 (2.89 \times 10^{-1})$
Δ C statistic (refit Framingham model)	Derivation	1 [Reference]	0.09 (0.06 to 0.12)	0.10 (0.08 to 0.12)
	Validation		0.05 (0.02 to 0.09)	0.07 (0.04 to 0.09)
Integrated discrimination index[d]	Derivation	1 [Reference]	0.12 (0.08 to 0.16)	0.14 (0.10 to 0.17)
	Validation		0.08 (0.05 to 0.10)	0.10 (0.08 to 0.13)
NRI(>0)[d]	Derivation	1 [Reference]	0.52 (0.40 to 0.65)	0.72 (0.60 to 0.84)
	Validation		0.43 (0.26 to 0.57)	0.48 (0.33 to 0.62)
Event NRI[d]	Derivation	1 [Reference]	0.22 (0.11 to 0.36)	0.29 (0.19 to 0.42)
	Validation		0.08 (−0.06 to 0.22)	0.30 (0.16 to 0.44)
No-event NRI[d]	Derivation	1 [Reference]	0.30 (0.22 to 0.36)	0.43 (0.36 to 0.48)
	Validation		0.35 (0.28 to 0.41)	0.18 (0.11 to 0.24)

Abbreviation: NRI, net reclassification index.

[a] Δ C statistic and Δ discrimination slope indicate the difference in C statistic and discrimination slope either between derivation and validation or between 9-protein model and refit Framingham model.

[b] Quintile hazard ratio is the ratio of hazard for patients in the 5th (highest) quintile risk category compared with those in the first (lowest) quintile risk category.

[c] Point estimates and 95% CIs are shown for all values except Hosmer-Lemeshow calibration statistic, for which the point estimate (mean square difference between predicted and observed risk across the deciles) and associated *P* value are shown.

[d] The integrated discrimination index and category-free NRI(>0) were calculated using the refit Framingham model as the reference model with event NRI and no-event NRI indicating the fraction of participants correctly reclassified by the 9-protein model within the event and no-event groups.

with the refit Framingham model generated an NRI(>0) of 0.48 (95% CI, 0.33-0.62) (Table 3). The mean 4-year risk proteomic risk was within 2 percentage points of the observed event rate in the external validation cohort (calibration-in-the-large). Calibration performance across the full range of the 9-protein risk scores is shown in **Figure 3** (eFigure 2 [for refit Framingham model] in the Supplement); for the 9-protein model, the observed risk in each decile of the validation cohort was within 5 percentage points of the mean protein risk score. The 9-protein model was developed for the composite end points of MI, heart failure, stroke, and death. For individual end points, median 9-protein risk score in derivation for MI was 33% (95% CI, 25.6%-38.6%); for heart failure, 37% (95% CI, 31.5%-43.7%); for stroke, 24% (95% CI, 19.6%-29.7%); and for death, 30% (95% CI, 27.0%-34.0%). In the absence of any event, the median 4-year 9-protein risk score was 14.2% (95% CI, 13.5%-15.2%). Similar risk score distributions across these event types were observed in the validation cohort (eFigure 5 in the Supplement).

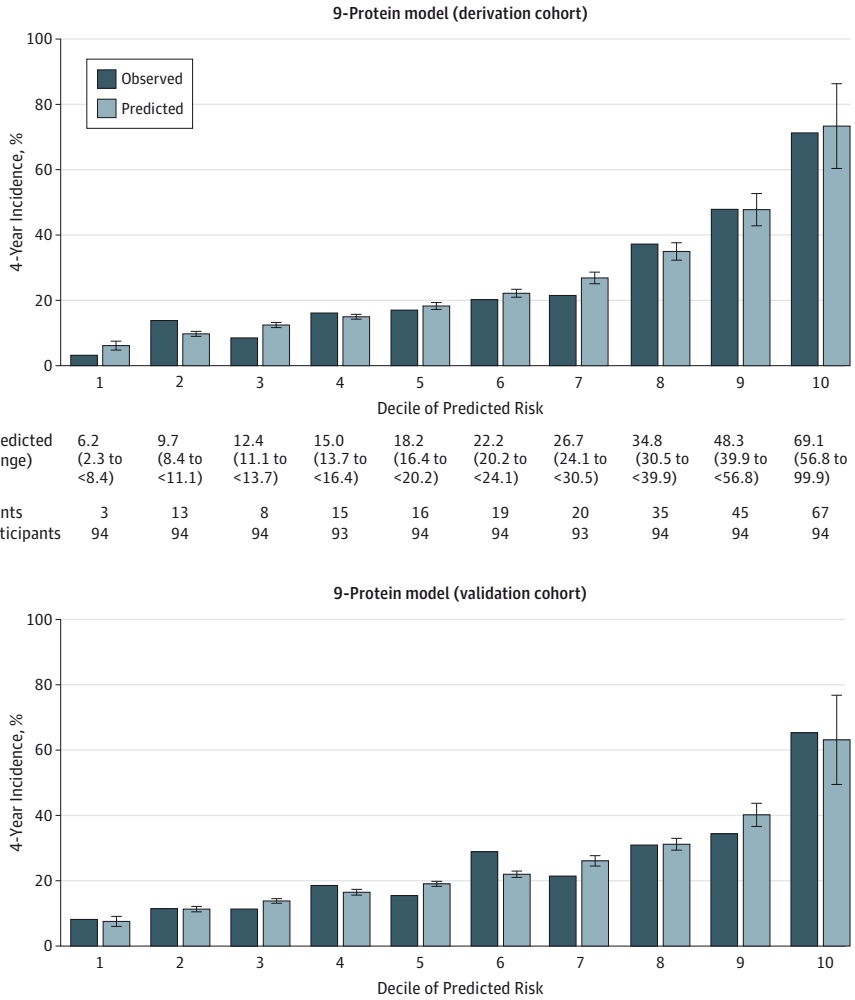
Analysis of Paired Samples

Changes in the 9-protein risk score were evaluated from paired samples from 514 participants (Heart and Soul study) in whom second plasma samples were taken a median of 4.8 years after the first, and participants were event-free between these 2 samples. The baseline characteristics of this subset of participants were similar to all Heart and Soul participants in this study (Table 1) except the time to the first event was longer because of the requisite absence of events prior to the second sample.

Among the participants with paired samples, 139 had an event (MI, heart failure, stroke, or death) after the second sample; the paired samples were taken a median of 2.8 years and 7.7 years prior to that event. The remaining 375 participants had paired samples a median of 4.3 and 9.0 years prior to completing their event-free follow-up. This analysis assessed whether the 9-protein risk score changed to a greater extent for participants approaching an event compared with participants who remained event free.

As **Figure 4** shows, 139 participants who experienced an event after the second sample had a median 9-protein risk of

Figure 3. Agreement Between Observed vs Predicted 4-Year Incidence of Myocardial Infarction, Stroke, Heart Failure, and Death With the 9-Protein Model



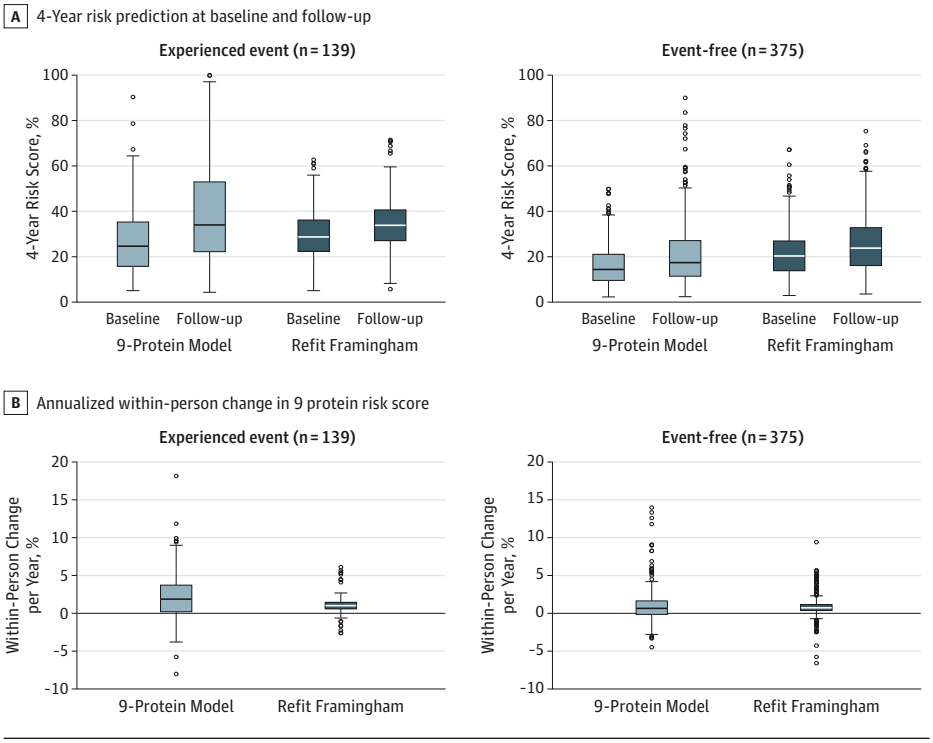
	1	2	3	4	5	6	7	8	9	10
Median predicted risk, % (range)	6.2 (2.3 to <8.4)	9.7 (8.4 to <11.1)	12.4 (11.1 to <13.7)	15.0 (13.7 to <16.4)	18.2 (16.4 to <20.2)	22.2 (20.2 to <24.1)	26.7 (24.1 to <30.5)	34.8 (30.5 to <39.9)	48.3 (39.9 to <56.8)	69.1 (56.8 to 99.9)
No. of events	3	13	8	15	16	19	20	35	45	67
No. of participants	94	94	94	93	94	94	93	94	94	94

	1	2	3	4	5	6	7	8	9	10
Median predicted risk, % (range)	7.8 (2.7 to <9.7)	11.4 (9.7 to <12.6)	13.8 (12.6 to <15.0)	16.5 (15.0 to <17.9)	19 (17.9 to <20.5)	21.9 (23.6 to <23.6)	26.0 (24.1 to <28.6)	30.7 (28.6 to <34.7)	39.5 (34.7 to <46.6)	59.2 (46.6 to 99.2)
No. of events	8	11	11	18	15	28	21	30	33	64
No. of participants	98	97	97	97	97	97	97	97	97	97

Predicted 4-year incidence for each decile is the mean predicted risk score in each decile. Error bars indicate standard deviations. For additional details, see section 4 in the Supplement.

24.6% (95% CI, 22.6%-27.7%) at baseline and 34.0% (95% CI, 29.2%-38.4%) at 4.8 years while median refit Framingham risk was 28.7% (95% CI, 26.6%-30.3%) at baseline and 33.8% (95% CI, 32.5%-36.2%) at 4.8 years. The 375 participants who were event free during the entire study had a median 9-protein risk of 14.4% (95% CI, 13.5%-16.5%) at baseline and 17.4% (95% CI, 16.0%-19.0%) at 4.8 years while median refit Framingham risk was 20.3% (95% CI, 19.0%-21.5%) at baseline and 23.8% (95% CI, 22.2%-25.8%) at 4.8 years. The absolute within-person change in the 9-protein risk was greater than for the refit Framingham model for participants with events (P = .002); median annualized within-person change was 1.86% (95% CI, 1.15%-2.54%) for the 9-protein model compared with 1.00% (95% CI, 0.87%-1.19%) for refit Framingham. Over 5 years, these annualized values represent an absolute change in risk of 9.3% for the 9-protein score and 5.0% for refit Framingham. For both risk models, these within-person changes were greater than for the event-free group (P < .001), in which the median annualized within-person change in the 9-protein risk group was 0.65% (95% CI, 0.45%-0.86%) compared with 0.72% (95% CI, 0.64%-0.80%) for refit Framingham (P = .3). The IDI for the 9-protein risk predictions at baseline, compared with 4.8 years, was 0.07 (95% CI, 0.04-0.10)—an absolute increase in mean risk of 7% for participants with events after the second sample over their baseline risk. Combining the 9-protein prognostic index at 4.8 years with the within-person change from baseline yielded an augmented model that fit slightly better (P = .03) than the 9-protein prognostic index at 4.8 years alone, although the discriminatory power was not meaningfully improved (IDI, 0.009; NRI(>0), 0.26; Δ C statistic, 0.006).

Figure 4. Changes in Risk Scores of Myocardial Infarction, Stroke, Heart Failure, and Death in Paired Samples 4.8 Years Apart



A, Predicted cardiovascular risk in paired samples among 139 participants who experienced an event after the second sample (left), and 375 participants who were event free during the entire study (right). Both panels show 4-year risk predictions at baseline and follow-up for the 9-protein and refit Framingham models.

B, Annualized within-person change in 9-protein risk was greater than for refit Framingham for participants who experienced events (P = .002) and similar to change in refit Framingham for participants who were event free (P = .30). Median absolute annualized within-person change in the 9-protein risk score was 1.86% compared with 1.00% for refit Framingham.

Key to symbols: horizontal line indicates the median, top and bottom ends of the boxes indicate the interquartile range (IQR), upper and lower error bars extend to 1.5 × the IQR, and the circles indicate data points beyond 1.5 × the IQR.

Discussion

Individualized risk assessment in patients diagnosed with apparently stable CHD is necessary because stable CHD appears to be a heterogeneous entity with a broad range of outcomes.[10,11] For stratification of cardiovascular risk using the "omics" technologies, genomics has been investigated most extensively, but genomic risk scores do not substantively improve risk discrimination over traditional risk factors.[4,5,26] Even if genomic approaches are ultimately successful, they will succeed primarily in predicting risk related to lifelong exposure and will not discern any changes in risk over time.[4,5,11,26] Compared with genomics, proteomics offers several advantages: proteins integrate both environmental and genetic influences; proteins are responsive to lifestyle and therapeutic interventions, informing of changes in risk[27,28]; and proteins are effectors of biological process and thus potential targets of therapies.[29] However, limitations in proteomic techniques have to this point hindered the implementation of these advantages.

In this study, levels of 1130 plasma proteins were measured using modified aptamers[7,15,30,31] to identify prognostic proteins that improve cardiovascular risk prediction. A prediction time horizon of 4 years was chosen—sufficiently long to implement therapeutic changes[11]—and yet not so distant that risk becomes deniable, losing its motivation. In the discovery cohort, 200 proteins were prognostic of cardiovascular events (eTable 4 in the Supplement), many of which are newly discovered biomarkers of cardiovascular risk.

An unbiased statistical approach was used to arrive at a 9-protein risk prediction model which, by itself, performed better than traditional risk factors represented by a refit Framingham secondary event model[20] and offered fair discrimination based on the C statistic (Table 3). The discrimination slope represents the separation in mean risk between participants with and without events.[8,18] The addition of the 9-protein risk score to refit Framingham offered a substantial improvement in this separation (Table 3). Admittedly, the large magnitude of the improvement in discrimination (in C statistic, discrimination slope separation, and IDI) and net reclassification by the 9-protein model (Table 3) was partly reflective of the weak performance of traditional risk factors in predicting the risk of secondary events,[32] also observed in the present study.

By including an independent external cohort in this study, best practices for validation were followed,[8] reducing the risk of translation to clinical use by verifying the predictive capacity of the key prognostic proteins and their combination in the proteomic model to less-stringent sample collection and processing that are more typical of clinical practice.[6,9,12] In applying protein-based risk assessment to patients with stable CHD, this diagnosis was found to be associated with a broad range of cardiovascular and mortality risks (Figure 2, Figure 3), suggesting that stable CHD may not represent a single homogeneous entity.

Paired samples were used to evaluate whether the proteomic risk changed over time as participants approached a cardiovascular event. The 9-protein risk score changed more than the refit Framingham model among participants approach-

ing new events. In addition, the 9-protein risk score generated at the follow-up sample was a stronger predictor of subsequent outcomes than the preceding baseline risk score. The mutability of the proteomic risk score, in relation to future events, offers a potential advantage over genetic risk prediction, which remains unchanged during lifetime. It remains unclear, however, whether the magnitude of changes in the proteomic risk score among participants with future events might lead to a change in management.

Other cardiovascular risk algorithms for stable CHD are available, including a model from the REACH (Reduction of Atherothrombosis for Continued Health) registry, which combines traditional risk factors with information about the extent of diseased vascular beds, heart failure, atrial fibrillation, medical treatments, and geographic location.[33] The REACH registry algorithm reported a C statistic for the prediction of a next cardiovascular event of 0.67 (95% CI, 0.66-0.68) and lacked external validation. The present study results could not be directly compared with the REACH model because some of the REACH variables were unavailable in its 2 cohorts.

Another cardiovascular risk prediction model used the best available candidate biomarkers for cardiovascular outcomes in the Heart and Soul cohort,[10] including high-sensitivity troponin, NT-proBNP, C-reactive protein, and urine albumin: creatinine ratio. This risk prediction model did not replicate well in external validation.[10] Genetic variants have also been associated with the risk of CHD. A recent study tested how well a genetic risk score based on 27 variants could predict recurrent CHD events in the CARE (Cholesterol and Recurrent Events) and PROVE IT-TIMI 22 (Pravastatin or Atorvastatin Evaluation and Infection Therapy-Thrombolysis in Myocardial Infarction 22) trial populations.[4] The adjusted quintile HR was 1.81 (95% CI, 1.22-2.67), a risk prediction that is appreciably smaller than proteomics yielded in the present study, with an adjusted quintile HR of 7.63 (95% CI, 4.53-12.85) in the validation set (Table 3).

Study Strengths

This study conducted a large-scale proteomic analysis of cardiovascular risk, using a high-throughput proteomic platform.[7,16,30,31] The study was conducted in 2 large well-characterized cohorts with standardized adjudication of outcome events[12,34] across 2 continents and included cross-sectional and longitudinal assessments. Specimen quality has been noted as an important reason why omics findings reported from one laboratory may not replicate in others.[9] Accordingly, the analyses in the present study were conducted across a range of specimen qualities, representative of standardized (derivation) and clinical practice conditions (validation). The findings were consistent across this range of specimen quality.

Limitations

This initial analysis of circulating proteins focused on a population of relatively high-risk individuals with established CHD. There is additional need for accurate cardiovascular risk prediction in the lower-risk general population or in even higher-risk individuals with CHD. Another limitation is that this study investigated only the sensitivity to increasing risk as represented by an approaching event; it will be important to evaluate individual medical interventions that alter risk to learn how well proteins can discern changes in risk in specific settings.

Conclusions

Among patients with stable CHD, a risk score based on 9 proteins performed better than the refit Framingham secondary event risk score in predicting cardiovascular events but still only provided modest discriminative accuracy. Further research is needed to assess whether the score is more accurate in a lower-risk population.

ARTICLE INFORMATION

Author Contributions: Dr Ganz had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Study concept and design: Ganz, Hveem, Jonasson, Kato, Williams.
Acquisition, analysis, or interpretation of data: Ganz, Heidecker, Hveem, Kato, Segal, Sterling, Williams.
Drafting of the manuscript: Ganz, Kato, Segal, Sterling, Williams.
Critical revision of the manuscript for important intellectual content: Ganz, Heidecker, Hveem, Jonasson, Kato, Segal, Williams.
Statistical analysis: Ganz, Kato, Segal, Sterling, Williams.
Administrative, technical, or material support: Ganz, Hveem, Jonasson, Williams.
Study supervision: Ganz, Segal, Williams.

Conflict of Interest Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Dr Segal reports receipt of payment from SomaLogic to support statistical analyses costs. Drs Hveem and Jonasson report affiliation with the HUNT study, which received a payment from SomaLogic for providing plasma specimens and database information. Drs Sterling and Williams are employees of SomaLogic. Dr Williams reports serving on the board of Venaxis Inc. Mr Kato is an employee of the NEC Corporation of America, which has a research contract with SomaLogic. No other disclosures are reported.

Funding/Support: The Heart and Soul cohort was supported by the Department of Veterans Affairs; the National Heart, Lung, and Blood Institute (R01 HL079235); the American Federation for Aging Research; the Robert Wood Johnson Foundation; and the Ischemia Research and Education Foundation. The HUNT3 cohort was funded by the Norwegian Ministry of Health, Norges Teknisk-Naturvitenskapelige University, the Norwegian Research Council, Central Norway Regional Health Authority, the Nord-Trøndelag County Council, and the Norwegian Institute of Public Health. Measurement of all proteins in both cohorts was funded by SomaLogic.

Role of the Funder/Sponsor: SomaLogic provided the funding for and execution of the protein assays. SomaLogic had no veto rights regarding the content of the study or the decision to submit the manuscript for publication.

Additional Contributions: The authors thank Christopher J. Bock, BA, Joy L. Boyd, BS, Kaitlin Gargiulo, BS, Amanda L. Grisco, BA, Tracy R. Keeney, BA, and Elyn E. Lytton, BS, all current or former compensated employees of SomaLogic, for performing the protein assays; Rosalynn Gill, PhD, a former compensated employee of SomaLogic, and Christina Lee, PhD, MBA, a current compensated employee of SomaLogic, for providing logistical support; Edward N. Brody, MD, PhD, Trudi Foreman, BA, Richard M. Lawn, PhD, Robert E. Mehler, MD, Rachel Ostroff, PhD, Britta Singer, PhD, all current or former compensated employees of SomaLogic, for assisting with the design of the study, sample management and interpretation of analytical data; Stephan Kraemer, PhD, a compensated employee of SomaLogic, for overseeing the execution of protein assays; Hilde Kjeldstad Berg, MSc, a compensated employee of the HUNT cohort research group, for providing logistical support; and Bjørnar Klykken, MD, a compensated employee of the HUNT cohort research group, for adjudicating outcomes.

REFERENCES

1. Go AS, Mozaffarian D, Roger VL, et al; American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics—2014 update: a report from the American Heart Association. *Circulation*. 2014;129(3):e28-e292.

2. Goff DC Jr, Lloyd-Jones DM, Bennett G, et al; American College of Cardiology/American Heart Association Task Force on Practice Guidelines. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation*. 2014;129(25)(suppl 2):S49-S73.

3. Ridker PM, Cook NR. Statins: new American guidelines for prevention of cardiovascular disease. *Lancet*. 2013;382(9907):1762-1765.

4. Mega JL, Stitziel NO, Smith JG, et al. Genetic risk, coronary heart disease events, and the clinical benefit of statin therapy: an analysis of primary and secondary prevention trials. *Lancet*. 2015;385(9984):2264-2271.

5. Paynter NP, Chasman DI, Paré G, et al. Association between a literature-based genetic risk score and cardiovascular events in women. *JAMA*. 2010;303(7):631-637.

6. Lindsey ML, Mayr M, Gomes AV, et al; American Heart Association Council on Functional Genomics and Translational Biology, Council on Cardiovascular Disease in the Young, Council on Clinical Cardiology, Council on Cardiovascular and Stroke Nursing, Council on Hypertension, and Stroke Council. Transformative impact of proteomics on cardiovascular health and disease: a scientific statement from the American Heart Association. *Circulation*. 2015;132(9):852-872.

7. Gold L, Ayers D, Bertino J, et al. Aptamer-based multiplexed proteomic technology for biomarker discovery. *PLoS One*. 2010;5(12):e15004.

8. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*. 2010;21(1):128-138.

9. McShane LM, Cavenagh MM, Lively TG, et al. Criteria for the use of omics-based predictors in clinical trials. *Nature*. 2013;502(7471):317-320.

10. Beatty AL, Ku IA, Bibbins-Domingo K, et al. Traditional Risk Factors Versus Biomarkers for Prediction of Secondary Events in Patients With Stable Coronary Heart Disease: From the Heart and Soul Study. *J Am Heart Assoc*. 2015;4(7):e001646.

11. Eagle KA, Ginsburg GS, Musunuru K, et al. Identifying patients at high risk of a cardiovascular event in the near future: current status and future directions: report of a national heart, lung, and blood institute working group. *Circulation*. 2010;121(12):1447-1454.

12. Krokstad S, Langhammer A, Hveem K, et al. Cohort profile: the HUNT study, Norway. *Int J Epidemiol*. 2013;42(4):968-977.

13. Marcus GM, Olgin JE, Whooley M, et al. Racial differences in atrial fibrillation prevalence and left atrial size. *Am J Med*. 2010;123(4):375 e1-375 e7.

14. Olson KA, Beatty AL, Heidecker B, et al. Association of growth differentiation factor 11/8, putative anti-ageing factor, with cardiovascular outcomes and overall mortality in humans: analysis of the Heart and Soul and HUNT3 cohorts. *Eur Heart J*. 2015;36(48):3426-3434.

15. Vaught JD, Bock C, Carter J, et al. Expanding the chemistry of DNA for in vitro selection. *J Am Chem Soc*. 2010;132(12):4141-4151.

16. Brody E, Gold L, Mehan M, et al. Life's simple measures: unlocking the proteome. *J Mol Biol*. 2012;422(5):595-606.

17. Tibshirani R. The lasso method for variable selection in the Cox model. *Stat Med*. 1997;16(4):385-395.

18. McClelland RL, Jorgensen NW, Budoff M, et al. 10-Year coronary heart disease risk prediction using coronary artery calcium and traditional risk factors: derivation in the MESA (Multi-Ethnic Study of Atherosclerosis) with validation in the HNR (Heinz Nixdorf Recall) study and the DHS (Dallas Heart Study). *J Am Coll Cardiol*. 2015;66(15):1643-1653.

19. Tibshirani R. Regression shrinkage and selection via the lasso: a retrospective. *J R Stat Soc B*. 2011;73(3):273-282. doi:10.1111/j.1467-9868.2011.00771.x

20. D'Agostino RB, Russell MW, Huse DM, et al. Primary and subsequent coronary risk appraisal: new results from the Framingham study. *Am Heart J*. 2000;139(2 Pt 1):272-281.

21. Heagerty PJ, Zheng Y. Survival model predictive accuracy and ROC curves. *Biometrics*. 2005;61(1):92-105.

22. Pencina MJ, D'Agostino RB Sr, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat Med*. 2011;30(1):11-21.

23. Pencina MJ, D'Agostino RB, Pencina KM, Janssens AC, Greenland P. Interpreting incremental value of markers added to risk prediction models. *Am J Epidemiol*. 2012;176(6):473-481.

24. Pepe MS, Janes H. Methods for evaluating prediction performance of biomarkers and tests, UW Biostatistics Working Paper Series (384). http://biostats.bepress.com/uwbiostat/paper384. Accessed June, 5, 2016.

25. R Core Team RFfSC. *Austria R: A Language and Environment for Statistical Computing*. Vienna, Austria: The R Foundation; 2013.

26. Ripatti S, Tikkanen E, Orho-Melander M, et al. A multilocus genetic risk score for coronary heart disease: case-control and prospective cohort analyses. *Lancet*. 2010;376(9750):1393-1400.

27. Ridker PM, Danielson E, Fonseca FA, et al; JUPITER Trial Study Group. Reduction in C-reactive protein and LDL cholesterol and cardiovascular event rates after initiation of rosuvastatin: a prospective study of the JUPITER trial. *Lancet*. 2009;373(9670):1175-1182.

28. Nissen SE, Tuzcu EM, Schoenhagen P, et al; Reversal of Atherosclerosis with Aggressive Lipid Lowering (REVERSAL) Investigators. Statin therapy, LDL cholesterol, C-reactive protein, and coronary artery disease. *N Engl J Med*. 2005;352(1):29-38.

29. Dadu RT, Ballantyne CM. Lipid lowering with PCSK9 inhibitors. *Nat Rev Cardiol*. 2014;11(10):563-575.

30. Kraemer S, Vaught JD, Bock C, et al. From SOMAmer-based biomarker discovery to diagnostic and clinical applications: a SOMAmer-based, streamlined multiplex proteomic assay. *PLoS One*. 2011;6(10):e26332.

31. Mehan MR, Ostroff R, Wilcox SK, et al. Highly multiplexed proteomic platform for biomarker discovery, diagnostics, and therapeutics. *Adv Exp Med Biol*. 2013;735:283-300.

32. Pencina MJ, D'Agostino RB Sr, Song L. Quantifying discrimination of Framingham risk functions with different survival C statistics. *Stat Med*. 2012;31(15):1543-1553.

33. Wilson PW, D'Agostino R, Sr., Bhatt DL, et al. An international model to predict recurrent cardiovascular disease. *Am J Med*. 2012;125(7):695-703 e1.

34. Whooley MA, de Jonge P, Vittinghoff E, et al. Depressive symptoms, health behaviors, and risk of cardiovascular events in patients with coronary heart disease. *JAMA*. 2008;300(20):2379-2388.

EDITORIAL

Using Aptamer-Based Technology to Probe the Plasma Proteome for Cardiovascular Disease Prediction

Marc S. Sabatine, MD, MPH

Circulating biomarkers play a major role in risk stratification of patients with cardiovascular disease. The 3 most widely used cardiovascular biomarkers—troponin, C-reactive protein, and B-type natriuretic peptide—have each been shown to predict risk of major adverse cardiovascular outcomes beyond traditional clinical factors, and for that reason, use of these biomarkers is recommended in various practice guidelines.

Related article page 2532

How then should additional biomarkers be added to the clinical armamentarium? The traditional approach has been for candidate biomarkers to emerge from preclinical or pilot clinical data and then be tested individually in progressively larger cohorts that typically involve thousands of patients. This serial approach is slow and requires the development of dedicated assays, typically immunoassays, to detect each biomarker of interest. Multiplexed immunoassays have been developed but typically are constrained by interference between antibodies, proteins, and assay diluents, leading to degradations in accuracy that limit the number of simultaneous assays.[1]

In contrast, proteomics entails the large-scale systemic analysis of proteins. Indeed, proteomics holds substantial promise because proteins more directly shape an individual's phenotype than do genotypes. Akin to genome-wide association studies, investigators can use a nontargeted proteomic experimental strategy, in which information is acquired on every analyte in a sample.[2] With only approximately 19 000 genes that encode proteins and only a subset of these proteins appearing in the circulation, assaying the human plasma proteome might seem more tractable than genomics, in which there are more than 3 billion base pairs.

The reality, however, is far more complicated. Proteins can undergo many different sorts of posttranslational modification that may generate species with different functions. Moreover, the dynamic range of plasma protein concentrations is vast[3] with differences in abundance of more than 10^{10}. Although improvements in technology have allowed discovery work to be done,[4] the requisite upstream sample processing to reduce sample complexity prior to analysis has effectively precluded the high throughput that is necessary to analyze the large numbers of samples required to detect realistic risk ratios.

In this issue of *JAMA*, Ganz and colleagues[5] use an approach that lands somewhere in between the 2 extremes.[1,3] Specifically, the investigators relied on aptamer-based technology to assay 1130 proteins in plasma samples from 2 cohorts of patients with stable coronary heart disease, including 938 samples from patients in the derivation cohort and 971 samples from patients in the validation cohort. Aptamers are small nucleic acids (either RNA or single-stranded DNA) that can form secondary and tertiary structures capable of specifically binding proteins or other cellular targets and thus can be considered the chemical equivalent of antibodies.[6] In brief, for each of the proteins of interest, a modified aptamer was created that also contains biotin, a photocleavable linker, and a fluorophore.[7] Each sample of plasma was incubated with the aptamers and then 2 immobilization steps utilizing strepavidin-coated beads were performed to eliminate unbound aptamers and unbound or noncognately bound proteins. The aptamers were eluted from the target proteins, hybridized to a microarray containing complementary single-stranded DNA probes, and then quantified using the fluorescent tags.

Using this technology, Ganz et al[5] found 9 proteins associated with adverse cardiovascular outcomes that passed the various statistical hurdles they rightly used when engaged in multiple hypothesis testing to this degree (Bonferroni correction to the α threshold, LASSO [least absolute shrinkage and selection operator], and then backward stepwise selection). The authors constructed a multiprotein risk score using levels of the 9 proteins and found a gradient of risk for adverse cardiovascular outcomes, not only in their derivation set, as to be expected, but also in a separate validation set.

How should these findings be interpreted? The authors conclude that their 9-protein risk score performed better than a refit Framingham secondary event risk score but still only provided modest discriminatory accuracy, with a C statistic of 0.70 (vs 0.64 for the clinical score). However, this modest accuracy should not necessarily dampen enthusiasm for their results. In fact, the focus on discriminatory accuracy is misplaced in studies of biomarkers for long-term prognosis. Discriminatory accuracy is critically important when assessing a test, typically dichotomized, for diagnosis. For instance, is a patient presenting with chest pain having a myocardial infarction, a pulmonary embolism, or an aortic dissection? There is a truth to be known that is immediately knowable, and ideally for clinical application, all patients with the disease would have a positive test result and all those without disease would have a negative test result.

The situation is very different when dealing with prognosis and trying to predict the risk of an adverse cardiovascular outcome many years down the road. In this case, there is usually a gradient of risk in approximate but imperfect proportion

to the level of the biomarker, and hence, the C statistic will be modest. But clinicians are typically less interested in discriminating risk between 2 patients and more interested in being able to better calibrate a particular patient's risk of adverse events.[8] Ganz et al[5] showed the ability of their 9-protein model to do this, with patients having scores in the bottom third of the distribution having an approximate observed cardiovascular event rate over 4 years of 10%, those in the middle third 20%, and those in the upper third greater than 30%. These data would be important to physicians and patients alike.

As is the case for any study, there are limitations to be considered and further questions to be answered. Analytic validation of the specificity of the aptamer-based findings would be reassuring. The outcome was a composite of myocardial infarction, stroke, hospitalization for heart failure, and all-cause mortality. Such a broad composite was likely necessary to ensure a sufficient number of events given the relatively small size of the cohorts. However, the pathobiology underlying an acute myocardial infarction, hospitalization for heart failure, and a noncardiovascular death are likely quite different. If anything, such an admixture would make it more difficult to predict the composite outcome accurately. Nonetheless, analyses with a more homogeneous set of outcomes might yield more interesting biological associations. How would the 9 proteins fare when compared with a better clinical risk score plus commonly available and validated biomarkers such as a high-sensitivity troponin assay, a natriuretic peptide, and a measure of renal function? Such analyses will be important to truly gauge the clinically relevant prognostic value of this multiprotein approach.

Although more accurate risk prediction is always welcome, clinicians more readily embrace measuring a prognostic biomarker or calculating a risk score if the results could alter therapeutic decision making.[9,10] To that end, it would be interesting to apply these arrays to samples from patients in randomized clinical trials of therapies. If a gradient of treatment benefit existed, such data would make measurement of the relevant proteins in clinical practice more compelling (which, for the current list, is impractical). Furthermore, part of the long-term value of this sort of proteomics work may come from exploring the basic pathways that underline some of the novel associations described. If some of these proteins are true risk factors, rather than simply risk markers, then they could serve as targets for future therapies. The quest for personalized or precision medicine is an important one,[11] and the work by Ganz et al[5] is a welcome step in that direction.

ARTICLE INFORMATION

Author Affiliation: TIMI Study Group, Division of Cardiovascular Medicine, Brigham and Women's Hospital, Boston, Massachusetts.

Corresponding Author: Marc S. Sabatine, MD, MPH, TIMI Study Group, Division of Cardiovascular Medicine, Brigham and Women's Hospital, 350 Longwood Ave, Boston, MA 02115 (msabatine@partners.org).

Conflict of Interest Disclosures: The author has completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Dr Sabatine reports receipt of research grant support to his institution from Abbott Laboratories, Accumetrics, Critical Diagnostics, Daiichi-Sankyo, Eisai, Genzyme, GlaxoSmithKline, Nanosphere, Roche Diagnostics, Takeda, Gilead, Poxel, and Novartis; receipt of research grant support and personal fees for consultancy services from Amgen, AstraZeneca, Bristol-Myers Squibb, Intarcia, Merck, and Sanofi-aventis; and receipt of personal fees for consultancy services only from Cubist, MyoKardia, Pfizer, Quest Diagnostics, Zeus Scientific, CVS Caremark, Alnylam, and Duke Clinical Research Institute.

REFERENCES

1. Ellington AA, Kullo IJ, Bailey KR, Klee GG. Antibody-based protein multiplex platforms: technical and operational challenges. *Clin Chem*. 2010;56(2):186-193.

2. Benson MD, Sabatine MS, Gerszten RE. The prospects for cardiovascular proteomics: a glass approaching half full. *JAMA Cardiology*. doi:10.1001/jamacardio.2016.0279

3. Anderson NL, Anderson NG. The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics*. 2002;1(11):845-867.

4. Keshishian H, Burgess MW, Gillette MA, et al. Multiplexed, quantitative workflow for sensitive biomarker discovery in plasma yields novel candidates for early myocardial injury. *Mol Cell Proteomics*. 2015;14(9):2375-2393.

5. Ganz P, Heidecker B, Hveem K, et al. Development and validation of a protein-based risk score for cardiovascular outcomes among patients with stable coronary heart disase. *JAMA*. doi:10.1001/jama.2016.5951

6. Ni X, Castanares M, Mukherjee A, Lupold SE. Nucleic acid aptamers: clinical applications and promising new horizons. *Curr Med Chem*. 2011;18(27):4206-4214.

7. Gold L, Ayers D, Bertino J, et al. Aptamer-based multiplexed proteomic technology for biomarker discovery. *PLoS One*. 2010;5(12):e15004.

8. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation*. 2007;115(7):928-935.

9. Morrow DA, Cannon CP, Rifai N, et al; TACTICS-TIMI 18 Investigators. Ability of minor elevations of troponins I and T to predict benefit from an early invasive strategy in patients with unstable angina and non-ST elevation myocardial infarction: results from a randomized trial. *JAMA*. 2001;286(19):2405-2412.

10. Bohula EA, Bonaca MP, Aylward PE, et al. Development of a TIMI risk score for stable ischemic heart disease. *Circulation*. 2015;132(suppl 3):A17338.

11. Mega JL, Sabatine MS, Antman EM. Population and personalized medicine in the modern era. *JAMA*. 2014;312(19):1969-1970.