Prognosis and prognostic research: what, why, and how?

Karel G M Moons,1 Patrick Royston,2 Yvonne Vergouwe,1 Diederick E Grobbee,1 Douglas G Altman3

Doctors have little specific research to draw on when predicting outcome. In this first article in a series Karel Moons and colleagues explain why research into prognosis is important and how to design such research.

Prognosis simply means foreseeing, predicting, or estimating the probability or risk of future conditions; familiar examples are weather and economic forecasts. In medicine, prognosis commonly relates to the probability or risk of an individual developing a particular state of health (an outcome) over a specific time, based on his or her clinical and non-clinical profile. Outcomes are often specific events, such as death or complications, but they may also be quantities, such as disease progression, (changes in) pain, or quality of life.

In medical textbooks, however, prognosis commonly refers to the expected course of an illness. This terminology is too general and has limited utility in practice. Doctors do not predict the course of an illness but the course of an illness in a particular individual. Prognosis may be shaped by a patient’s age, sex, history, symptoms, signs, and other test results. Moreover, prognostication in medicine is not limited to those who are ill. Healthcare professionals, especially primary care doctors, regularly predict the future in healthy individuals—for example, using the Apgar score to determine the prognosis of newborns, cardiovascular risk profiles to predict heart disease in the general population, and prenatal testing to assess the risk that a pregnant woman will give birth to a baby with Down’s syndrome.

**What is prognosis?**

**Multivariable research**

Given the variability among patients and in the aetiology, presentation, and treatment of diseases and other health states, a single predictor or variable rarely gives an adequate estimate of prognosis. Doctors—implicitly or explicitly—use multiple predictors to estimate a patient’s prognosis. Prognostic studies therefore need to use a multivariable approach in design and analysis to determine the important predictors of the studied outcomes and to provide outcome probabilities for different combinations of predictors, or to provide tools to estimate such probabilities. These tools are commonly called prognostic models, prediction models, prediction rules, or risk scores. They enable care providers to use combinations of predictor values to estimate
an absolute risk or probability that an outcome will occur in an individual. A multivariable approach also enables researchers to investigate whether specific prognostic factors or markers that are, say, more invasive or costly to measure, have worthwhile added predictive value beyond cheap or simply obtained predictors—for example, from patient history or physical examination. Nonetheless, many prognostic studies still consider a single rather than multiple predictors.  

**Use of prognostic models**

Medical prognostication and prognostic models are used in various settings and for various reasons. The main reasons are to inform individuals about the future course of their illness (or their risk of developing illness) and to guide doctors and patients in joint decisions on further treatment, if any. For example, modifications of the Framingham cardiovascular risk score are widely used in primary care to determine the indication for cholesterol lowering and antihypertensive drugs. Examples from secondary care include use of the Nottingham prognostic index to estimate the long term risk of cancer recurrence or death in breast cancer patients, the acute physiology and chronic health evaluation (APACHE) score and simplified acute physiology score (SAPS) to predict hospital mortality in critically ill patients, and models for predicting postoperative nausea and vomiting.

Another reason for prognostication and use of prognostic models is to select relevant patients for therapeutic research. For example, researchers used a previously validated prognostic model to select women with an increased risk of developing cancer for a randomised trial of tamoxifen to prevent breast cancer. Another randomised trial on the efficacy of radiotherapy after breast conserving resection used a prognostic model to select patients with a low risk of cancer recurrence.

Prognostic models are also used to compare differences in performance between hospitals. For example, the clinical risk index for babies (CRIB) was originally developed to compare performance and mortality among neonatal intensive care units. More recently Jarman et al developed a model to predict the hospital standardised mortality ratio to explain differences between English hospitals.

**Differences from aetiological research**

Although there are clear similarities in the design and analysis of prognostic and aetiological studies, predicting outcomes is not synonymous with explaining their cause. In aetiological research, the mission is to explain whether an outcome can reliably be attributed to a particular risk factor, with adjustment for other causal factors (confounders) using a multivariable approach. In prognostic research, the mission is to use multiple variables to predict, as accurately as possible, the risk of future outcomes. Although a prognostic model may be used to provide insight into causality or pathophysiology of the studied outcome, that is neither an aim nor a requirement. All variables potentially associated with the outcome, not necessarily causally, can be considered in a prognostic study. Every causal factor is a predictor—albeit sometimes a weak one—but not every predictor is a cause. Nice examples of predictive but non-causal factors used in everyday practice include skin colour in the Apgar score and tumour markers as predictors of cancer progression or recurrence. Both are surrogates for obvious causal factors that are more difficult to measure.

Furthermore, to guide prognostication in individuals, analysis and reporting of prognostic studies should focus on absolute risk estimates of outcomes given combinations of predictor values. Relative risk estimates (for example, odds ratio, risk ratio, or hazard ratio) have no direct meaning or relevance to prognostication in practice. In prediction research, relative risks are used only to obtain an absolute probability of the outcome for an individual, as we will show in our second article.

In contrast, aetiological and therapeutic studies commonly focus on relative risks—for example, the risk of an outcome in presence of a causal factor relative to the risk in its absence. Also, the calibration and discrimination of a multivariable model are highly relevant to prognostic research but meaningless in aetiological research.

**How to study prognosis?**

Building on previous guidelines, we distinguish three major steps in multivariable prognostic research that are also followed in the other articles in this series: developing the prognostic model, validating its performance in new patients, and studying its clinical impact (box). We focus here on the non-statistical characteristics of a multivariable study aimed at developing a prognostic model. The statistical aspects of developing a model are covered in our second article.

**Objective**

The main objective of a prognostic study is to determine the probability of the specified outcome with different combinations of predictors in a well defined population.

**Study sample**

The study sample includes people at risk of developing the outcome of interest, defined by the presence of a particular condition (for example, an illness, undergoing surgery, or being pregnant).

**Study design**

The best design to answer prognostic questions is a cohort study. A prospective study is preferable as it enables optimal measurement of predictors and outcome (see below). Studies using cohorts already assembled for other reasons allow longer follow-up times but usually at the expense of poorer data. Unfortunately, the prognostic literature is dominated by retrospective studies. Case-control studies are sometimes used for prognostic analysis, but they do not automatically allow estimation of absolute risks because cases and controls are often sampled from a source population of unknown size. Since investigators are free to choose the ratio of
cases and controls, the absolute outcome risks can be manipulated. An exception is a case-control study nested in a cohort of known size. Data from randomised trials of treatment can also be used to study prognosis. When the treatment is ineffective (relative risk=1.0), the intervention and comparison group can simply be combined to study baseline prognosis. If the treatment is effective the groups can be combined, but the treatment variable should then be included as a separate predictor in the multivariable model. Here treatments are studied on their independent predictive effect and not on their therapeutic or preventive effects. However, prognostic models obtained from randomised trial data may have restricted generalisability because of strict eligibility criteria for the trial, low recruitment levels, or large numbers refusing consent.

**Predictors**
Candidate predictors can be obtained from patient demographics, clinical history, physical examination, disease characteristics, test results, and, previous treatment. Prognostic studies may focus on a cohort of patients who have not (yet) received prognosis modifying treatments—that is, to study the natural course or baseline prognosis of patients with that condition. They can also examine predictors of prognosis in patients who have received treatments.

Studied predictors should be clearly defined, standardised, and reproducible to enhance generalisability and application of study results to practice. Predictors requiring subjective interpretation, such as imaging test results, are of particular concern in this context because there is a risk of studying the predictive ability of the observer rather than that of the predictors. Also, predictors should be measured using methods applicable—or potentially applicable—to daily practice. Specialised measurement techniques may yield optimistic predictions.

As discussed above, the prognostic value of treatments can also be studied, especially when randomised trials are used. However, caution is needed in including treatments as prognostic factors when data are observational. Indications for treatment and treatment administration are often not standardised in observational studies and confounding by indication could lead to bias and large variation in the (type of) administered treatments. Moreover, in many circumstances the predictive effect of treatments is small compared with that of other important prognostic variables such as age, sex, and disease stage.

Finally, of course, studies should include only predictors that will be available at the time when the model is intended to be used. If the aim is to predict a patient’s prognosis at the time of diagnosis, for example, predictors that will not be known until actual treatment has started are of little value.

**Outcome**
Preferably, prognostic studies should focus on outcomes that are relevant to patients, such as occurrence or remission of disease, death, complications, tumour growth, pain, treatment response, or quality of life. Surrogate or intermediate outcomes, such as hospital stay or physiological measurements, are unhelpful unless they have a clear causal relation to relevant patient outcomes, such as CD4 counts instead of development of AIDS or death in HIV studies. The period over which the outcome is studied and the methods of measurement should be clearly defined. Finally, outcomes should be measured without knowledge of the predictors under study to prevent bias, particularly if measurement requires observer interpretation. Blinding is not necessary when the outcome is all cause mortality. But, if the outcome is cause specific mortality, knowledge of the predictors might influence assessment of outcomes (and vice versa in retrospective studies where predictors are documented after the outcome was assessed).

**Required number of patients**
The multivariable character of prognostic research makes it difficult to estimate the required sample size. There are no straightforward methods for this. When the number of predictors is much larger than the number of outcome events, there is a risk of overestimating the predictive performance of the model. Ideally, prognostic studies require at least several hundred outcome events. Various studies have suggested that for each candidate predictor studied at least 10 events are required, although a recent study showed that this number could be lower in certain circumstances.

**Validation and application of prognostic models**
Formally developed and validated prognostic models are often used in weather forecasting and economics (with varying success), but not in medicine. There may be several reasons for this. Firstly, prognostic models are often too complex for daily use in clinical settings without computer support. The introduction of computerised patient records will clearly enhance not only the development and validation of models in research settings but also facilitate their application in routine care. Secondly, because many prognostic models have not been validated in other populations, clinicians may and perhaps should not trust probabilities provided by these models.

Finally, clinicians often do not know how to use predicted probabilities in their decision making. Validation studies are scarce, but even fewer models are tested for their ability to change clinicians’ decisions, let alone to change patient outcome. We support the view that no prediction model should be implemented in practice until, at a minimum, its performance has been validated in new individuals. The third article in this series discusses why validation studies are important and how to design and interpret them.

Validation studies are particularly important if a prediction model is to be used in individuals who were not represented in the development study—for example, when transporting a model from secondary to primary care or from adults to children, which seems a form of extrapolation rather than validation.
SUMMARY POINTS

Prognosis is estimating the risk of future outcomes in individuals based on their clinical and non-clinical characteristics.

Predicting outcomes is not synonymous with explaining their cause.

Prognostic studies require a multivariable approach to design and analysis.

The best design to address prognostic questions is a cohort study.

discuss this further in the fourth article in the series, as well as how to update existing models to other circumstances.4

We stress that prediction models are not meant to take over the job of the doctor.7 41 44 46 They are intended to help doctors make decisions by providing more objective estimates of probability as a supplement to other relevant clinical information. Furthermore, they improve understanding of the determinants of the course and outcome of patients with a particular disease.

Funding: KGM, YV, and DEG are supported by the Netherlands Organisation for Scientific Research (ZON-MW 917.46.360). PR is supported to help doctors make decisions by providing more objective estimates of probability as a supplement to other relevant clinical information. Furthermore, they improve understanding of the determinants of the course and outcome of patients with a particular disease.

Contributors: The four articles in the series were conceived and planned by DGA. KGM, PR, and YV KGM wrote the first draft of this article. All the authors contributed to subsequent revisions.

Competing interests: None declared.

Provenance and peer review: Not commissioned; externally peer reviewed.


32 Walsraven CV, Davis D, Forster AJ, Wells GA. Time-dependent bias was common in survival analyses published in leading clinical journals. J Clin Epidemiol 2004;57:672-82.


39 Christensen E. Prognostic models including the Child-Pugh, MELD and Mayo risk scores—where are we and where should we go? Hepatology 2004a:41:344-50.


