

## Answering clinical questions about prognosis

### Introduction

One of your patients, a 24 year old student, has recently recovered from a first episode of functional psychosis. At a pre-discharge meeting with the patient and his relatives, they inquire about the long term outcome, in particular about the chances of a recurrence in the next few years.

“Prognosis” concerns the prediction of future events. The main form of the prognostic question is usually “what is the likelihood of this particular outcome in a patient with this disorder?” In mental health, prognostic questions may be about the medium to long term outcome of an illness—or they may be about the short term likelihood of a serious adverse event—for example, homicide or suicide. We may also ask if there are any characteristics of the patient that make it more likely that they will have a particularly good or bad outcome. These prognostic factors can be demographic (eg, age or sex), disease specific (presentation and symptomatology), or the presence of other conditions (comorbidity).<sup>1</sup> Prognostic factors are distinguished from risk factors for the disease because they are determinants of what will happen *after* development of the illness rather than determinants of whether the patient *will develop* the disorder.

Treatment is one form of prognostic factor, although a special form in that it can be manipulated to improve the likelihood of a beneficial outcome. The fact that treatment can be manipulated also means that it can be studied experimentally in randomised controlled trials. It is either impossible or unethical, however, to randomise most prognostic factors, and therefore alternative study designs are required. The aim of the alternative study design is the same as a randomised trial—that is, to provide an unbiased estimate of the effect of the prognostic factor. When randomisation is not possible, a cohort study is usually the most reliable design—although other designs such as case control studies are also used. In this article we consider some useful approaches to appraising and using an article about prognosis: we have previously covered articles about diagnosis and treatment.<sup>2–3</sup> As with most clinical questions, the best place to start is by looking for, and appraising, a systematic review of all available studies.<sup>4–6</sup>

### Critical appraisal of a prognosis article

#### STUDY SAMPLE

##### ● *Was the study sample representative?*

One of the key requirements of a prognostic study is that an unbiased sample of patients—who were representative of the target patient group—were recruited. The possibility of achieving a statistically representative sample is actually one of the advantages of a cohort study because randomised trials can rarely, if ever, recruit a truly representative sample. The main reason for this is that entry into a randomised trial involves both patient and clinician agreeing to randomisation. In a non-randomised study, there is no manipulation of treatment, and so this barrier is removed.

A secondary issue is whether the target population of patients in the study is the same as, or at least comparable to, patients in real clinical settings. In a well known study, Slater and Gilthero

followed up patients from the National Hospital for Nervous Diseases in London who had been diagnosed as having hysteria.<sup>7</sup> They found that about one half of the patients developed a clear cut neurological illness during the 10 years of follow up. These findings have been generalised to other settings, but this may not be appropriate because the National Hospital for Nervous Diseases is a tertiary referral centre. Therefore the patients will have passed through various referral filters and are likely to be different from those seen in less specialised settings. Indeed, the prognosis of patients with diagnoses of hysteria even at this institution might have changed over time due to changing diagnostic practices. Consistent with this, a partial replication of Slater and Gilthero’s study found a much lower rate of neurological illness at follow up.<sup>8</sup>

##### ● *Was the sample well defined and were the patients at a similar point in the course of their illness?*

To be clinically useful, an article needs to make clear which patients were included in the study. In our initial example, definitions of functional psychoses such as schizophrenia have changed dramatically over the past 100 years and it is only relatively recently that reliable diagnostic criteria have been used in research and practice. A meta-analysis of prognostic studies in schizophrenia illustrates this point well.<sup>9</sup> The authors of this review identified 320 studies of patients with schizophrenia in which fewer than 33% dropped out and in which there was  $\geq 1$  year of follow up. The studies had been done over a period of about 100 years and, as diagnostic fashions changed, the methods of diagnosis varied. This limits the validity of the results and of course severely reduces the clinical applicability of the results of this analysis for the modern day clinician. The key questions are “were the participants in the study reliably defined and was the diagnostic approach useful to the clinician who is reading the paper?”

The next issue concerns the stage in the course of the illness at which patients were recruited into the study. Returning to the meta-analysis referred to above, it was unclear in most studies how long the patients had been ill before inclusion in the study. It is likely that the future course of the illness will be highly influenced by the preceding course—so how can a clinician derive a useful estimate for the individual patient in front of them? To be clinically useful, the study needs to recruit patients at a uniform point in the course of the illness—this will usually be at the onset, or a very early stage, of the disorder—or at a defined point in the condition. A cohort of patients identified at an early or common point in the course of the illness is termed an *inception cohort*. In *Evidence-Based Mental Health*, we only include articles that report inception cohorts. However, a clinically useful inception cohort may be defined at change point in the clinical location of the patient. For example, it may be useful to consider the risk of suicide in all patients after discharge from hospital, and so we also include this kind of study when the study is otherwise methodologically strong and when the results are clinically useful.

## Follow up

### ● Was the follow up sufficiently long and complete?

If the clinician is interested in the long term prognosis of a disorder such as schizophrenia, then a study with 12 months of follow up will probably not be very useful. In general, follow up needs to be long enough to make it likely that a high proportion of those patients who are going to experience a particular clinically relevant event will have done so. It is also essential that most patients be followed up because the patients who are not followed up may differ systematically from those who are. It is difficult to generalise about a satisfactory follow up rate because this will depend on the reasons for the failure of some patients to be followed up. However, to ensure reliable screening by the editorial staff, *Evidence-Based Mental Health* has a requirement of at least an 80% follow up rate until the occurrence of a major study endpoint or to the end of the study.

### ● Were objective outcome criteria applied by investigators who were unaware of the baseline characteristics of the patients?

The more subjective the outcome measure, the greater the potential for bias in measurement. Many outcomes in mental health are subjective, for example, symptom levels. "Harder" outcomes such as all cause mortality or admission to hospital are likely to be less susceptible to observer bias. The potential for bias is reduced if the outcome assessors are masked to features of the study including baseline characteristics of the patients or even the purpose of or hypotheses being tested by the study.

## Results

### ● What are the results and can they be used in caring for your patient?

The most straightforward results from a prognosis study are usually a simple proportion of the number of patients who experienced a particular outcome divided by the total number of patients in the study. For example, in the study by Wiersma *et al* of the long term follow up of 82 patients with a first episode of non-affective functional psychosis, 63% of patients met DSM-III-R criteria for schizophrenia by 6 months, and 55% had relapsed within 2 years and 70% within 5 years.<sup>10</sup> It is important to remember that, even if the study has successfully avoided the systematic biases referred to above, the study is of a sample of all patients with a first episode of non-affective functional psychosis. Uncertainty remains about the true proportion because of the random error that affects all estimates derived from a sample. This uncertainty about the precision of the result can be expressed by the *confidence interval (CI)*, which is the range of values in which we can be sure the true value lies. It is conventional to use the 95% CI—this is the range of values in which we can be 95% confident that the true value lies. For example, from this study the proportion of patients who had complete remission of symptoms by 15 years was 27%—the CI was 18% to 38%, so the most likely value is 27%, but the proportion could be as low as 18% or as high as 38%.

### ● Were prognostic factors identified and how reliable are the estimates?

The identification of reliable prognostic factors for good or bad outcomes is obviously very useful for tailoring the results of a study to an individual patient. In the study by Wiersma *et al*, the authors used survival analysis to investigate whether any baseline patient characteristics were associated with outcome. Only 1 was identified: a delay in treatment was associated with longer duration of the first psychotic illness: the hazard ratio was 2.3. The hazard ratio is a measure of the risk of an outcome in the subgroup with the risk factor compared with the average risk in

the full group. Although these estimates of prognostic factors are very attractive, they need to be treated cautiously because they are analogous to subgroup analyses in randomised trials. They are much more prone to the effects of both random and systematic biases than the overall estimate because they are derived from a subgroup of the sample and therefore more uncertainty exists about their precision. In addition, they are prone to *confounding*—a situation in which the measurement of the association between a risk factor and an outcome is distorted by a third factor that is both associated with the risk factor of interest and causes or prevents the outcome. For example, a delay in the treatment of a first psychotic episode may be associated with a slow insidious onset, which might lead to a poor prognosis. The likelihood of confounding is why it is important that the factors of interest are adjusted for the effects of other important prognostic factors (when such factors are known and measurable). In general, unless prognostic factors have been adequately adjusted for confounding and revalidated in an independent sample of patients, then the clinician should be cautious about relying on them. It is usually better to rely mainly on the overall estimate of prognosis for the full cohort (with the CI). When we abstract the results of studies in *Evidence-Based Mental Health* we therefore emphasise the main results for the full cohort and where we present estimates for subgroups, we present only the measure of the relative risk rather than absolute measures. Wherever possible, we provide the relative risk or, in studies using survival analyses, the hazard ratio.

One way of using multiple predictors is to develop a *clinical prediction guide* based on multivariate statistical methods. For example, by combining a patient's clinical features it is possible to classify their risk of deep venous thrombosis as high, low, or moderate.<sup>11</sup> Reliable and validated clinical prediction guides are still uncommon in mental health. Our sister journal, *Evidence-Based Medicine*, abstracts clinical prediction guides and has developed minimal methodological criteria for such studies that include a requirement for retesting of the model in a second set of patients.<sup>11</sup> It is likely that *Evidence-Based Mental Health* will adopt a similar approach.

In future EBMH notebooks we will cover how to use the estimates of event rates from cohort studies to tailor the results of randomised trials to individual patients.

JOHN GEDDES, MD, FRCPsych  
Editor

- 1 Laupacis A, Wells G, Richardson WS, *et al*. Users' guides to the medical literature. V. How to use an article about prognosis. Evidence-based medicine working group. *JAMA* 1994;272:234–7.
- 2 Streiner D, Geddes J. Some useful concepts and terms used in articles about diagnosis [editorial]. *Evidence-Based Mental Health* 1998;1:6–7.
- 3 Szatmari P. Some useful concepts and terms used in articles about treatment [editorial]. *Evidence-Based Mental Health* 1998;1:39–40.
- 4 Glanville J, Lefebvre C. Identifying systematic reviews: key resources [editorial]. *Evidence-Based Mental Health* 2000;3:68–9.
- 5 Geddes J, Freemantle N, Streiner D, *et al*. Understanding and interpreting systematic reviews and meta-analyses. Part 1: rationale, search strategy, and describing results [editorial]. *Evidence-Based Mental Health* 1998;1:68–9.
- 6 Freemantle N, Geddes J. Understanding and interpreting systematic reviews and meta-analyses. Part 2: meta-analyses [editorial]. *Evidence-Based Mental Health* 1998;1:102–4.
- 7 Slater ET, Glithero E. A follow up of patients diagnosed as suffering from "hysteria." *J Psychosom Res* 1965;9:9–13.
- 8 Crimlisk HL, Bhatia K, Cope H, *et al*. Slater revisited: 6 year follow up study of patients with medically unexplained motor symptoms. *BMJ* 1998;316:582–6.
- 9 Hegarty JD, Baldessarini RJ, Tohen M, *et al*. One hundred years of schizophrenia: a meta-analysis of the outcome literature. *Am J Psychiatry* 1994;151:1409–16.
- 10 Wiersma D, Nienhuis FJ, Slooff CJ, *et al*. Natural course of schizophrenic disorders: a 15-year follow up of a Dutch incidence cohort. *Schizophr Bull* 1998;24:75–85.
- 11 McGinn T, Randolph A, Richardson S, *et al*. EBM notebook: clinical prediction rules [editorial]. *Evidence-Based Medicine* 1998;3:5–6.