# Users' Guides to the Medical Literature

## V. How to Use an Article About Prognosis

Andreas Laupacis, MD, MSc; George Wells, MSc, PhD; W. Scott Richardson, MD; Peter Tugwell, MD, MSc; for the Evidence-Based Medicine Working Group

## CLINICAL SCENARIO

You are about to see a 76-year-old retired schoolteacher for the second time. You first saw her in the clinic a month ago because of cognitive problems. Your evaluation at that time included a Standardized Mini-Mental State Examination,[1] on which she scored 18 out of a possible 30 points, and a physical examination that was normal including no focal neurological signs. You arranged investigations for the treatable causes of dementia that were negative, and you thus feel she has probable Alzheimer's disease.

The patient has lived with her son since her husband died 6 years ago. Her son thinks that she first developed significant problems with her memory about 3 years ago. However, she has become increasingly agitated and paranoid during the last year. She has refused to allow him to look after her financial affairs, despite the fact that she owns three pieces of property and isn't able to manage them herself. Her son asked you about her prognosis, and whether she is likely to die soon from the dementia. You indicated that you would discuss this with him at the second visit once the results of all the tests are available.

## SEARCH

Hoping to provide the son with the most specific information possible about his mother's prognosis, after the initial visit you searched the medical library's MEDLINE CD-ROM system via the hospital's network on the clinic computer. You entered "*Alzheimer's Disease," which yielded 3687 articles from 1990 onward. Next, you entered "prognosis," which yielded 23 004 articles; crossing the two sets yielded 27 articles. Scanning the abstracts on screen, you found several articles of potential interest, including one that seemed precisely on target: "Survival of Outpatients With Alzheimer-Type Dementia" by Walsh and colleagues.[2]

## INTRODUCTION

In this article we will suggest a framework that you can use to efficiently assess articles that deal with prognosis, using the article on patients with dementia as an example. We will follow the usual format of this series and discuss how to determine whether the results are valid, how to interpret the results, and whether the information will benefit your patients (Table).

"Prognosis" refers to the possible outcomes of a disease and the frequency with which they can be expected to occur (eg, death in a patient with dementia). Sometimes the characteristics of a particular patient can be used to more accurately predict that patient's eventual outcome (eg, a patient with dementia and behavioral problems may have a worse prognosis than someone without behavioral problems). These characteristics are called "prognostic factors." Prognostic factors can be any of several types, such as demographic (eg, age), disease-specific (eg, tumor stage), or comorbid (eg, other conditions accompanying the disease in question). They can predict any outcome, whether good (eg, cure or survival) or bad (eg, death or complication). Prognostic factors need not necessarily cause the outcomes, just be associated with them strongly enough to predict their development. In the literature, prognostic factors are usually distinguished from "risk factors," those patient characteristics associated with the development of the disease in the first place. For example, smoking is an important risk factor for the development of lung cancer, but is not as important a prognostic factor as tumor stage in someone who has lung cancer.

It is usually impossible or unethical to randomize patients to different prognostic factors. Therefore, the best study design to identify the presence of and determine the increased risk associated with a prognostic factor is a cohort study. As we described in a previous article in this series,[3] investigators conducting a cohort study follow one or more groups (cohorts) of individuals who have not yet suffered an adverse event and monitor the number of outcome events over time. An ideal cohort study consists of a well-defined sample of individuals representative of the population of interest and uses objective outcome criteria. One cohort study conducted in Framingham, Mass, in which investigators have followed a cohort of 5209 individuals since 1948, has provided clinicians with a great deal of useful information about the prognostic importance of many determinants of cardiovascular disease.[4] Since rigorous randomized trials include careful documentation of inclusion criteria and strict protocols for follow-up, patients in such trials form cohorts that can also generate information about the prognosis of a disease. However, the patients entered into the trial are often not representative of the population with the disorder.[5]

To study prognostic factors, investigators can also collect "cases" of individuals who have already suffered the outcome event and compare them with "controls" who have not. In these "case-control" studies, the investigators count the number of individuals in each group with a particular prognostic factor (eg, were the patients with dementia who died more likely to have had behavioral problems than those who did not die?). The potential for bias when selecting cases and controls, as well as the retrospective nature of data collection about prognostic factors (which often depends on the memory of the patients or their relatives or the accuracy of medical charts), limits the strength of inference clinicians can draw from case-control studies.[2] Also, case-control studies cannot provide information about the absolute risk of an event, but only about the relative risk (RR). Nevertheless, case-control studies can provide useful information and are particularly appropriate when the outcome is rare or the required duration of follow-up is long.

## ARE THE RESULTS OF THE STUDY VALID?

### Primary Guides

**Was There a Representative and Well-Defined Sample of Patients at a Similar Point in the Course of the Disease?**—This guide addresses two related issues. The first concerns how well defined the individuals in the study are, and whether they are representative of the underlying population. The authors should describe and specify their criteria for establishing that the patient has the disorder of interest (in this case, Alzheimer-type dementia) and how they selected their patient sample.

Several biases related to the assembly of the patients can distort the results of a study. For example, the sequence of referrals that leads patients from primary to tertiary centers raises the proportion of more severe or unusual cases, thus increasing the likelihood of adverse or nonfavorable outcomes. In one example of this "referral filter bias," the likelihood of a subsequent nonfebrile seizure in children with their first febrile seizure was much lower in community-based populations than in those drawn from hospitals.[6]

The second issue concerns whether the study patients are all at a similar, well-defined point in the course of their disease. Authors should provide a clear description of the stage of disease at which patients entered the study. For instance, since the duration of illness is often associated with outcome, the investigators should report the duration of illness for the sample patients. Within reason, all or most of the study patients should be at a similar point, such as survivors of a first myocardial infarction or patients newly diagnosed with lung cancer. However, the similar point in the course of disease need not be early on.

Walsh and colleagues[2] studied 126 outpatients with Alzheimer's disease who were consecutively referred to a multidisciplinary clinic for evaluation between 1980 and 1982. The diagnosis was made by consensus by a group consisting of an internist, psychiatrist, psychologist, neurologist or neuropathologist, and research nurse using the conventional *Diagnostic and Statistical Manual of Mental Disorders, Fourth Revision* criteria for dementia.[7] The tests used to exclude other causes of dementia were not described. However, given the multidisciplinary nature and expertise of the group, it seems reasonable to assume that the appropriate tests were done to exclude disorders such as hypothyroidism, depression, and space-occupying lesions of the brain that can be confused with Alzheimer's disease.

Walsh and colleagues reported survival from two different points in time: (1) referral to the clinic and (2) the point at which symptoms of memory loss were first noticed. The former is a more certain point in time, but suffers from the disadvantage that patients come to medical attention at different stages in the progression of their disease. The latter provides a more uniform starting point, but is potentially imprecise because dementia develops insidiously and the time of onset is identified retrospectively. Survival after presentation to clinic is probably more relevant for your patient's son.

**Was Follow-up Sufficiently Long and Complete?**—Since the presence of a prognostic factor often precedes the development of an outcome event by a long period, investigators must follow patients for long enough to detect the outcomes of interest. For example, recurrence in some women with early breast cancer can occur many years after initial diagnosis and treatment.[8] Patients in the dementia study were enrolled between 1980 and 1982 and followed until 1988 or their death. Thus, the follow-up was quite long, and 61% of the cohort died during this time.

Ideally, investigators will succeed in following all patients (as they did in the dementia study) but this is often not the case. Patients are not usually unavailable for follow-up for inconsequential reasons. Patients may fail to return because they have suffered exactly those events in which the investigators are interested (eg, they died or have been institutionalized). Conversely, patients who feel entirely healthy may also be less likely to return for evaluation because they feel so well. Simply put, the greater the number of patients unavailable for follow-up, the less accurate the estimate regarding the risk of the adverse outcome.

Under what circumstances does unavailability for follow-up compromise the validity of a study? It is important that you consider the relation between the proportion of patients who are unavailable and the proportion of patients who have suffered the adverse outcome of interest. The larger the number of patients whose fate is unknown relative to the number who have suffered an event, the greater the threat to the study's validity. For instance, let us assume that 30% of a particularly high-risk group (such as elderly diabetics) have suffered an adverse outcome (such as cardiovascular death) during long-term follow-up. If 5% of the patients have been lost, the true rate of patients who had died may be as high as 35%. Even if this were so, the clinical implications would not change, and the unavailability for follow-up doesn't threaten the validity of the study. However, in a much lower-risk patient sample (otherwise healthy middle-aged men, for instance) the observed event rate may be 1%. In this case, if one assumed that all 5% of the patients unavailable for follow-up had died, the event rate of 6% would have very different implications. If the number of patients unavailable for follow-up potentially jeopardizes the study's validity, you should look for the reasons for patients being unavailable, and compare the important demographic and clinical characteristics of the patients who were unavailable with the patients in whom follow-up was complete. To the extent that the reasons for disappearance are unrelated to outcome and the

patients who are unavailable are similar to those for whom information is complete, you may feel reassured. If investigators omit information about reasons for unavailability for follow-up or the characteristics of the patients who are unavailable, the strength of inference from the study results will be weaker.

### Secondary Guides

**Were Objective and Unbiased Outcome Criteria Used?**—Investigators must provide a clear and sensible definition of adverse outcomes before the study starts. Outcome events can vary from those that are objective and easily measured (death), to those that require some judgment (myocardial infarction), to those that require considerable judgment and may often be difficult to measure (eg, disability, quality of life). To minimize bias, the individual determining the outcomes should not know whether the patient had a potential prognostic factor. This is not always possible and, for unequivocal events such as death, may not be necessary. However, blinding is essential for outcomes requiring a great deal of judgment, such as transient ischemic attacks or unstable angina. In the study by Walsh and colleagues, the method and intensity of follow-up were not described. However, all patients were accounted for at the end of the study and the date of death was known for those who died.

**Was There Adjustment for Important Prognostic Factors?**—When comparing the prognosis of two groups of patients, investigators should consider whether their clinical characteristics are similar and adjust the analysis for any differences they find. The Framingham Study investigators reported that the rate of stroke in patients with atrial fibrillation and rheumatic heart disease was 41 per 1000 person years, which was very similar to the rate for patients with atrial fibrillation but no rheumatic heart disease.[9] However, patients with rheumatic heart disease were younger than those who did not have rheumatic heart disease. Once adjustments were made for the age, sex, and hypertensive status of the patients, the investigators found that the rate of stroke was sixfold greater in patients with rheumatic heart disease and atrial fibrillation than in patients with atrial fibrillation who did not have rheumatic heart disease.

Many studies of prognosis break the study group into cohorts based on suspected prognostic factors. Comparison of the pattern and frequencies of outcomes between these groups can determine the RR associated with the prognostic factor in question. For example, Pincus and colleagues[10] followed a co-

hort of patients with rheumatoid arthritis for 15 years. They separated the patients into a number of cohorts depending on their demographic characteristics, disease variables, and functional status. They found that some demographic variables (eg, age and education level) and functional status (eg, modified walking time and activities of daily living) were strongly predictive of mortality.

Since treatments can also alter patient outcomes, they should be taken into account when analyzing prognostic factors. While such treatments are not, strictly speaking, prognostic factors, the investigators should adjust for differences in treatment in the analysis. For example, in a study from Framingham that examined the prognosis of Q-wave vs non–Q-wave first myocardial infarction, the investigators adjusted for age, sex, and the presence of hypertension, angina pectoris, congestive heart failure, and cardiovascular disease prior to the infarct.[11] However, they did not take into account treatment with aspirin or β-blockers, which clinicians may have administered at the time of the infarct, and which we know have an impact on mortality.

In the study by Walsh and colleagues, no attempt was made to compare the mortality rate of the demented patients with a group of patients without dementia. However, they did evaluate the importance of 20 potential prognostic factors in their cohort. Age at symptom onset, dementia severity, wandering and falling, behavioral problems, and hearing loss all had a statistically significant relation to mortality.

### WHAT ARE THE RESULTS?

The quantitative results from studies of prognosis or risk are the number of events that occur over time. We will describe three common expressions of this relationship that provide complementary information about prognosis.

**How Large Is the Likelihood of the Outcome Event(s) in a Specified Period of Time?**—Your patient's son asked, "What are the chances that my mother will still be alive in 5 years?" You can provide a simple and direct answer in absolute terms. Five years after presentation to the clinic about one-half the patients (50%) had died. Thus, there is about a 50:50 chance that his mother will be alive in 5 years.

Your patient's son might then indicate that the only person he knows with Alzheimer's disease is a 65-year-old uncle who was diagnosed 10 years ago and is still living. He is surprised that his mother's chance of dying in the next 5 years is so high. This gives you the chance to
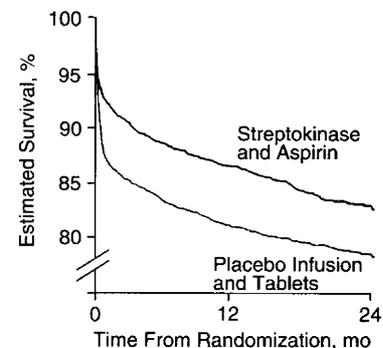


Fig 1.—Survival after myocardial infarction in patients treated with streptokinase and aspirin compared with placebo. Adapted from ISIS-2 (Second International Study of Infarct Survival) Collaborative Group. Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17 187 cases of suspected acute myocardial infarction: ISIS-2. 13:349-360, © by The Lancet Ltd, 1988.
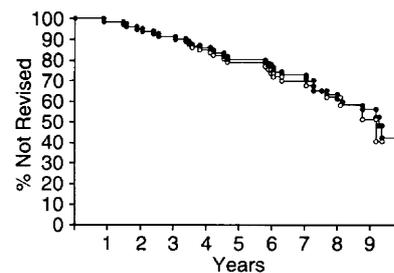


Fig 2.—Need for revision after total hip arthroplasty in two cohorts of patients in the same center (adapted from Dorey and Amstutz[14]).

discuss some of the prognostic factors for death in patients with Alzheimer's disease. As mentioned previously, the statistically significant prognostic factors for death were increasing age, dementia severity, behavioral problems, wandering and falling, and hearing loss. You explain that his mother is considerably older than his uncle was at the time of diagnosis, and that this likely explains some of the difference. It would be nice to use the prognostic factors to further refine the chance of death in his mother. Her age is almost identical to the mean age of the cohort studied by Walsh and colleagues. However, her Mini-Mental State Examination score is quite low (indicating more severe dementia), and her behavioral problems also suggest that she is at higher risk than the average patient in the study by Walsh et al. Unfortunately, no table or formula was presented that allows you to combine all of these factors and estimate a risk of mortality that is specific for your patient. However, you can feel confident in telling her son that his moth-

er's chances of dying are at least 50% during the next 5 years, and probably greater.

The son might then ask whether his mother's chances of survival could change with time. Neither the absolute nor relative expressions of results address this question. For this answer we should turn to a survival curve, a graph of the number of events over time (or conversely, the chance of being free of these events over time).[12] The events must be discrete (eg, death, stroke, recurrence of cancer), and the time at which they occur must be precisely known. In most clinical situations the chance of an outcome changes with time. Figures 1 and 2 show two survival curves, one of survival after a myocardial infarction[13] and the other the results of hip replacement surgery.[14] Note that the chance of dying after a myocardial infarction is highest shortly after the event (reflected by an initially steep slope of the curve, which then flattens), while very few hip replacements require revision until much later (this curve starts out flat and then steepens). Walsh and colleagues provided a survival curve in Fig 1 of their article that suggests that the chance of dying is more or less constant during the first 7 years after referral to the clinic for dementia.

**How Precise Are the Estimates of Likelihood?**—Even when valid, a prognostic study provides only an estimate of the true risk. After determining the size of the risk, we should next examine the precision of the estimate, which is best done with a confidence interval (CI). Walsh and colleagues found that the 95% CI for survival 5 years after presentation was approximately 39% to 58% (extrapolated from Fig 1 in their article). Note that in most survival curves, the

earlier follow-up periods usually include results from more patients than the later periods (because of unavailability for follow-up and because patients are not enrolled into the study at the same time). This means that the survival curves are more precise in the earlier periods, indicated by narrower confidence bands around the left-hand parts of the curve.

Walsh and colleagues also provided 95% CIs for the RR associated with each prognostic factor. For example, the RR associated with a behavioral problem was 1.5 with a 95% CI of 1.0 to 2.5. This means that the best estimate is that a patient with a behavioral problem is 1.5 times more likely to die than an individual without a behavioral problem. The probability that the true RR is between 1.0 (ie, no effect) and 2.5 is 95%.

## WILL THE RESULTS HELP ME IN CARING FOR MY PATIENTS?

**Were the Study Patients Similar to My Own?**—How well do the study results generalize to the patients in your practice? The authors should describe the study patients in enough detail to allow comparison with your patients. The article should list the patients' important clinical characteristics, along with the definitions used for these characteristics. The closer the match between the patient before you and those in the study, the more confident you can be in applying the study results to that patient. The characteristics of the study patients were quite similar to your patient.

**Will the Results Lead Directly to Selecting or Avoiding Therapy?**—Since there are no therapies for dementia that are routinely available and clearly effective, this guide does not directly apply to your patient. However, prognos-

tic data often provide the basis for sensible decisions about therapy. Knowing the expected clinical course of your patient's condition can help you judge whether treatment should be offered at all. For example, warfarin markedly decreases the risk of stroke in patients with nonrheumatic atrial fibrillation and is indicated for many patients with this disorder.[15] However, in one study the frequency of stroke in patients with "lone" atrial fibrillation (patients 60 years of age or younger with no associated cardiopulmonary disorders) was 1.3% over 15 years.[16] The risks of long-term warfarin therapy in this group of patients probably outweigh the benefits.

**Are the Results Useful for Reassuring or Counseling Patients?**—Even if the prognostic result does not lead you to prescribe an effective therapy, it can still be clinically useful. A valid, precise, and generalizable result of uniformly good prognosis is very helpful to the clinician when reassuring a concerned patient or relative. Some conditions, such as asymptomatic hiatal hernia or asymptomatic colonic diverticula, have such a good overall prognosis that they have been termed "nondisease."[17] On the other hand, a prognostic result of uniformly bad prognosis provides the clinician with a starting place for a discussion with the patient and family, leading to counseling about end-of-life concerns.

In your patient, information on the likelihood of death will be useful to the son and his family as they plan the future care of his mother. Of course, other prognostic information about the rate of progression of the dementing process and the need for intensive nursing care would also be useful.[18,19]

**References**

1. Molloy DW, Alemayehu E, Roberts R. Reliability of a standardized Mini-Mental State Examination compared with the traditional Mini-Mental State Examination. *Am J Psychiatry.* 1991;148:102-105.
2. Walsh JS, Welch G, Larson EB. Survival of outpatients with Alzheimer-type dementia. *Ann Intern Med.* 1990;113:429-434.
3. Levine M, Walter S, Lee H, et al, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature, IV: how to use an article about harm. *JAMA.* 1994;271:1615-1619.
4. Dawber TR, Kaunel WB, Lyell LP. An approach to longitudinal studies in a community: the Framingham Study. *Ann N Y Acad Sci.* 1963;107:539-556.
5. Bennett JC, for the Board on Health Sciences Policy of the Institute of Medicine. Inclusion of women in clinical trials: policies for population subgroups. *N Engl J Med.* 1993;329:288-292.
6. Ellenberg JH, Nelson KB. Sample selection and the natural history of disease: studies of febrile seizures. *JAMA.* 1980;243:1337-1340.
7. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders.* 4th

ed. Washington DC; 1994:139-143.
8. Early Breast Cancer Trialists' Collaborative Group. Systemic treatment of early breast cancer by hormonal, cytotoxic, or immune therapy: 133 randomised trials involving 31 000 recurrences and 24 000 deaths among 75 000 women. *Lancet.* 1992;339:1-15.
9. Wolf PA, Dawber TR, Thomas HE, Kannel WB. Epidemiologic assessment of chronic atrial fibrillation and risk of stroke: the Framingham Study. *Neurology.* 1978;28:973-977.
10. Pincus T, Brooks RH, Callahan LF. Prediction of long-term mortality in patients with rheumatoid arthritis according to simple questionnaire and joint count measures. *Ann Intern Med.* 1994;120:26-34.
11. Berger CJ, Murabito JM, Evans JC, Anderson KM, Levy D. Prognosis after first myocardial infarction: comparison of Q-wave and non–Q-wave myocardial infarction in the Framingham Heart Study. *JAMA.* 1992;228:1545-1551.
12. Katz MH, Hauck WW. Proportional hazards (Cox) regression. *J Gen Intern Med.* 1993;8:702-711.

13. ISIS-2 (Second International Study of Infarct Survival) Collaborative Group. Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17 187 cases of suspected acute myocardial infarction: ISIS-2. *Lancet.* 1988;2:349-360.
14. Dorey F, Amstutz H. The validity of survivorship analysis in total joint arthroplasty. *J Bone Joint Surg Am.* 1989;71:544-548.
15. Laupacis A, Albers G, Dunn M, Feinberg W. Antithrombotic therapy in atrial fibrillation. *Chest.* 1992;102:426S-433S.
16. Kopecky SL. The natural history of lone atrial fibrillation: a population-based study over three decades. *N Engl J Med.* 1987;317:669-674.
17. Meador CK. The art and science of nondisease. *N Engl J Med.* 1965;272:92.
18. Stern Y, Mayeux R, Hauser WA, Bush T. Predictors of disease course in patients with probable Alzheimer's disease. *Neurology.* 1987;37:1649-1653.
19. Drachman DA, O'Donnell BF, Lew RA, Swearer JM. The prognosis in Alzheimer's disease: 'how far' rather than 'how fast' best predicts the course. *Arch Neurol.* 1990;47:851-856.