

Appraisal of an article on prognosis

Marc Lester, James P. Warner and Robert Blizzard

Prompted by a clinical question, an article on prognosis in anorexia nervosa was appraised using evidence-based guidelines. Although problems with the validity and generalisability of the study were identified, this article yielded useful information. We conclude that it is not possible to address all clinical questions using the evidence-based framework.

As evidence-based practices become more widespread, it is helpful for clinicians to acquaint themselves with the benefits and pitfalls of this process. Case conferences and journal clubs can provide a good forum to learn about the various methods of evidence-based medicine (EBM) and provide clinicians with practical experience of the process (Warner & King, 1997). This is the second article in a series based on real-life experiences of trainees using the EBM approach.

Practising EBM involves distinct stages: setting a question, undertaking a literature search and assessing the validity and applicability of the available literature. Framing the question is an important part of the process. Getting the question right will help considerably with the subsequent literature search, and improve the chance of finding the literature appropriate to the clinical scenario. In this example, we sought a paper for presentation in our journal club that answered a question on predictors of outcome in a patient with anorexia nervosa.

Predictors of outcome in a patient with anorexia nervosa

Vignette

A woman in her late teens with a four year history of anorexia nervosa, and a body mass index (BMI) of 10 was presented at our weekly case conference. She had already developed osteoporosis, confirmed on bone densitometry. She had responded poorly to treatment. This case raised issues of predicting outcome in terms of physical morbidity and mortality, in an individual with severe anorexia nervosa.

Question

In a patient with anorexia nervosa is the presence of physical complications of the illness important in predicting the outcome?

Literature search

The first task was to identify keywords to use for a literature search. A Medline search, using the medical subject heading 'anorexia nervosa', covering the years 1993–1997, identified 991 articles. A keyword search of 'prognosis' identified a daunting 29 831 articles. We combined the two sets, resulting in 26 articles. Review of these revealed little of interest as they consisted mainly of single case reports and studies with small sample sizes. We then tried a different heading 'treatment outcome', which yielded 37 906 articles. Combining these with the heading 'anorexia nervosa' produced a useful-looking 54 articles. None of these articles appeared to answer our question, however several abstracts of articles by Herzog *et al* referred to a study of 84 patients who had been followed-up over many years.

We decided to expand the search by trying the 'explode' option on Medline which increases the sensitivity of the search by including all sub-headings. This did turn up more articles under each heading but none answered our question. Since the most important physical complication in our case was 'osteoporosis' we included this in our search. When combined with 'anorexia nervosa' this produced only 11 references, but one of these appeared relevant: 'Medical findings and predictors of medical outcome in anorexia nervosa: a prospective 12 year follow-up study' (Herzog *et al*, 1997).

The abstract looked promising in terms of the number of patients ($n=84$) and the long (12 year) follow-up period, and seemed to address our question. For the sake of completeness, we searched the Embase database using the same headings. This also identified the paper by Herzog *et al*, but turned up nothing else of particular interest.

Brief outline of the article

The article reported a longitudinal follow-up of consecutively treated patients with anorexia nervosa. The particular focus was the predictive value of initial laboratory findings for long-term outcome and to describe medical comorbidity and cause of death.

Herzog *et al* followed-up all in-patients at the University Hospital in Heidelberg, who were diagnosed as having anorexia nervosa using Feighner criteria and DSM-III-R criteria, between 1 January 1971 and 31 October 1980. Of the initial 84 patients, 18 were excluded leaving a cohort of 66 patients. Baseline assessments (Time 0) included a comprehensive battery of blood tests and collection of data on length of illness. Morgan-Russell outcome criteria for each year of follow-up were assessed. The resulting mean aggregate score comprised three possible outcomes:

- (a) Good, normal menstruation and weight
- (b) Intermediate, either pathological menstrual status or deviation of body weight
- (c) Poor, amenorrhoea and reduction in body weight.

Severity of medical comorbidity was assessed by a panel of three physicians using a five-point scale. Follow-up occurred at two points. Time 1, at average 3.6 years for 44(75.9%) patients, and Time 2 at an average 11.9 years for all patients. The interval for Time 2 follow-up was from nine to 18 years.

The authors concentrated on the results at Time 2, and reported a good outcome (according to Morgan-Russell criteria) in 47% ($n=31$) of patients, intermediate in 27% ($n=18$) and poor in 14% ($n=9$). Twelve per cent of patients ($n=8$) had died, mainly of acute medical conditions such as pneumonia and arrhythmias. In the poor/deceased outcome group initial albumin and potassium were significantly lower but creatinine and uric acid were significantly higher compared with the good/intermediate group. There was a marked increase in medical comorbidity assessed by the panel from initial assessment to Time 2 follow-up, from 14% ($n=9$) across the entire group at Time 0, to 67% ($n=6$) of those in the poor outcome group and 27% ($n=13$) of the good/intermediate outcome group. The Standardised Mortality Ratio (SMR) was 9.6. This indicates that patients in the study were nearly 10 times more likely to die compared with a normal population. The authors concluded that high creatinine and uric acid predicted chronic course, especially in association with low albumin, long duration of illness and laxative misuse. The most common medical diagnoses at Time 2 were osteoporosis (14%) and chronic renal failure (5%). They suggested medical comorbidity be included in evaluation of long term outcome of anorexia nervosa patients.

Critical appraisal of an article on prognosis

The main elements of an analysis of a paper on prognosis include: Are the results valid? What

are the results? Will the results help me in caring for my patients? (Laupacis *et al*, 1994). A pocket guide to critical appraisal (Sackett *et al*, 1997) provides a clear review of this process.

Are the results valid?

Was there a representative and well-defined sample of patients at a similar point in the course of their illness? No. The study did use consecutive admissions, the diagnostic criteria were well defined, and the patients were all females. We are led to assume that these were all first admissions, although this is not clearly stated. However, the study focused only on in-patients and there was probably referral bias to a specialist unit such as this. There was a wide age range, and no information was given about the length of illness at presentation. Little information was provided on socio-demographic characteristics of the sample. Admission policies and referral patterns are likely to change over the 10-year period of recruitment of this study.

Was follow-up sufficiently long and complete? Yes. There was a long follow-up period, and all drop-outs were accounted for, although there was no mention of the interim Time 1 follow-up examination later in the article. The range of individual follow-up times, from 9–18 years, suggest that the follow-up was not planned on an individual basis. This could introduce bias.

Were objective and unbiased outcome criteria used? Yes. This is the main thrust of the study – the value of laboratory tests as objective variables. The study also used well-recognised outcome criteria in the main, with a comprehensive set of outcome measures. The study seemed almost too pathophysiological in its outcome criteria, it is difficult to measure a disease such as this in purely physical terms. However, the non-blind consensus medical diagnosis is open to question. The use of average body weight in the paper is less standard than BMI in terms of outcome of weight; average body weight may change over time, and is more age-dependent unless very specific to age and menstrual status. However this is the weight classification in the Morgan-Russell criteria which does explain its inclusion.

Was there adjustment for important prognostic factors? No. There was only a brief mention of initial weight and length of illness, in addition to the social and psychological factors. There was no mention of treatment received by the patients after their initial admission during the whole 12-year follow-up period. Another area not covered was any measure of compliance with treatment.

One would expect treatment strategies to vary over a 10-year period of subject recruitment, and this was not addressed.

What are the results?

How likely are the outcomes over time? The SMR is 9.6 and, among the survivors, many are in the poor outcome category. There is a long follow-up period here with statistical significance between outcome groups of some predictive variables. However, the absence of data from the interim Time 1 follow-up or any real indication of course of illness makes it impossible to plot outcome over time, except for start and end-point of the study. Anorexia nervosa is characterised by remitting episodes in many cases and a point follow-up is not very useful in these circumstances.

How precise are the estimates of likelihood? The authors did not consistently provide confidence intervals. The all-cause mortality for the anorexic group was 12% and we calculated the 95% confidence interval to be 6–20%. The relatively small sample size means the confidence intervals are fairly wide. This suggests the figure for the SMR may be smaller, or considerably larger than 9.6. Statistical significance was reported but we have some concerns over the statistical methods used (i.e. a mixture of parametric and non-parametric tests). Student's *t*-test requires normally distributed data and this is not certain in the variables for which it is used. The small numbers in some of the outcome groups make the use of multivariate analysis unreliable, and the large number of losses (>20%) makes interpretation of results less convincing.

Will the results help me in my patient care?

Were the study patients similar to my own? *Unclear.* Only in-patients were included in the study, and no ethnic or demographic data were given. However, our patient may well be similar in socio-demographic terms to the study sample. The patient featured in our case example had normal values for the putative predictive variables in the study.

Will the results lead directly to selecting therapy? No. This paper is aimed more at identifying predictors of outcome of anorexia nervosa, and makes some arguments for the value of initial blood tests as a prognostic tool.

Are the results useful for reassuring patients? Nothing very reassuring from the results here! The overall outcomes are similar to pre-

vious studies. A further study is needed to see if specific interventions in the high risk group are of any benefit.

Comment

This exercise highlights how a circuitous literature search is sometimes necessary to detect important articles. Even though this was essentially an outcome study, it did not appear as such in the initial search, and only came to light once the subsequent heading 'osteoporosis' was added. With hindsight, we could have used a more efficient search strategy; using the medical subject heading 'anorexia nervosa' and exploding 'cohort' would have identified this paper more quickly (Sackett *et al*, 1997), as would an author search for the name 'Herzog'.

One aspect of the journal clubs is to highlight recent research advances, but in this case, by limiting the search to the period from 1993 to 1997, we missed an often-quoted paper in this area, that of Ratnasuriya *et al* (1991). The study by Ratnasuriya, which was cited in the paper we appraised, had a longer follow-up period of 20 years and described outcome in greater detail but had smaller numbers. However, the broad outcomes stated were similar to the study by Herzog *et al*.

Although there are a number of flaws in the paper we appraised, we do have more information about prognosis of our patient with anorexia nervosa than at the start of this exercise. According to EBM guidelines, the answer to the first part of the 'Are the results valid' question is no because of the 10-year recruitment period. However the task of amassing a cohort of individuals with a first onset is prohibitive in a single centre study. Although the evidence in this paper does not satisfy strict criteria concerning disease onset, and does not answer our question fully, it is nevertheless the best available evidence.

Our question may have been somewhat ambitious. Anorexia has multiple physical complications, and any study would have to be very large in order to have sufficient power to identify which complications have an impact on prognosis. Our question was phrased in such a way to preclude looking at individual prognostic indicators. Although it was less precise than it could have been, we felt it maximised our chances of finding something.

Not all studies fit the EBM 'guidelines' easily and these studies should not be rejected out of hand. The nature of psychiatry often precludes a wholly reductionist approach and this needs to be borne in mind when appraising the literature. If clinicians take a purely mechanistic approach to EBM, then most research

would be dismissed. Instead, EBM should help to provide a framework for the clinician to use their judgement better. The process is valuable in terms of its discipline of evaluation, the encouragement of a structured method of appraising research articles, and the process of considering and questioning one's practice. The education that trainees receive in searching for articles using databases is also useful.

References

- HERZOG, W., DETER, H.-C., FIEHN, W., *et al* (1997) Medical findings and predictors of long-term physical outcome in anorexia nervosa: a prospective twelve year follow-up study. *Physiological Medicine*, **27**, 269-279.
- LAUPACIS, A. WELLS, G., RICHARDSON, W. S., *et al* (1994) Users guide to the medical literature. V. How to use an article on prognosis. *Journal of the American Medical Association*, **272**, 234-237.
- RATNASURIYA, R. H., EISLER, I., SMUKLER, G. I., *et al* (1991) Anorexia nervosa: outcome and prognostic factors after 20 years. *British Journal of Psychiatry*, **158**, 495-502.
- SACKETT, D. L., RICHARDSON, W. S., ROSENBERG, W., *et al* (1997) *Evidence-Based Medicine*. New York: Churchill Livingstone.
- WARNER, J. P. & KING, M. (1997). Evidence-based medicine and the journal club: a cross-sectional survey of participants' views. *Psychiatric Bulletin*, **21**, 532-534.

Marc Lester, *Research Registrar*; *James Warner, *Lecturer*, and Bob Blizard, *Medical Statistician*, *Royal Free Hospital School of Medicine, Rowland Hill Street, London NW3 2QG*

*Correspondence

Evidence-Based Mental Health

New Journal

EDITORS: John Geddes, Shirley Reynolds, David Streiner & Peter Szatmari

New from the BMJ Publishing Group, *Evidence-Based Mental Health* follows on from the success of its sister publication *Evidence-Based Medicine* and it is designed to meet the needs of mental health specialists worldwide. Essential reading for clinicians of all disciplines, managers and policy makers.

Evidence-Based Mental Health will:

- keep the clinician up-to-date by using scientific criteria to select and abstract the most reliable and important clinically relevant papers from an expanded range of journals
- abstract promising preclinical studies to keep the specialist informed about current developments
- cover developments in diagnosis, therapy, harm, prognosis, economic evaluation, quality improvement, causation
- provide accompanying commentaries by experienced clinicians to facilitate the integration of research and clinician experience
- provide educational and theoretical articles on development in evidence-based practice, particularly aspects relevant to the mental health clinician
- adopt a multidisciplinary approach to mental health

SUBSCRIPTION RATES: (Publication: quarterly)

Institutional: Worldwide £120, USA only \$192/Personal: Worldwide £60, USA only \$96

*Society Members: Worldwide £45, USA only \$72

*Applies to members of the Royal College of Psychiatrists, the Royal College of Nursing and the British Psychological Society

Send orders to: BMJ Publishing Group, PO Box 299, London WC1H 9TD, UK. Fax: +44 (0) 171 383 6402 Tel: +44 (0) 171 383 6270 Email: bmjsubs@dial.pipex.com. US subscribers may send payment to: BMJ Publishing Group, Box 590A, Kennebunkport, ME 04046, USA Tel: 1-800-2-FON-BMJ Fax: 1-800-2-FAX-BMJ

BMJ
Publishing
Group

Royal College
of Psychiatrists

British
Psychological
Society