CRITICAL REVIEW

An Example of Critical Evaluation of Clinical Journal Articles: Separating the Wheat from the Chaff.

Stephen R. Tan, B.Sc.

Faculty of Medicine, Dalhousie University, Halifax, Nova Scotia, B3H 4H7

The ability to critically evaluate journal articles for relevance, applicability, and validity has become an everyday tool used by the practicing physician. It is essential for the clinician to be able to analyze the literature, determine which studies have valid methodology, compare the study population to the patient population being dealt with, and then to apply the results such that first rate health care is being given. The purpose of this paper is to review a method for critical evaluation of journal articles concerning the natural history and prognosis of disease. Three articles addressing the relationship between high serum cholesterol and the incidence of coronary heart disease (CHD) are used as examples of the application of these criteria.

INTRODUCTION

The physician today faces the challenge of continuing medical education through self-directed learning. The ability to rapidly access and interpret information has become a necessity for the practice of medicine. The sources of information are vast. Medical knowledge is expanding exponentially and is readily accessible with computerized literature searching aids, such as MEDLINE. However, not all studies regarding the natural history, prognosis and clinical course of a disease are equally valid. Studies vary in parameters such as study type, patient selection, methodology and applicability to one's own patient population.

Knowledge of the natural history and the clinical course of diseases is gained from studies on patient populations with the condition in question. Informed and appropriate decisions regarding the treatment of a patient with a specific disease are determined by knowledge of the clinical course and the prognosis.

The purpose of this paper is to revisit a series of published criteria (2) which may be used to critically evaluate journal articles for their usefulness. Three articles addressing the relationship between high serum cholesterol and the incidence of coronary heart disease (CHD) are used as examples of the application of these criteria.

Address correspondence to:

Stephen Tan, 1502-1333 South Park Street, Halifax, NS B3J 2K9

PROBLEM STATEMENT

What is the association between serum total cholesterol and the incidence of CHD?

Incidence will be defined as the number of new events occurring in the population within a defined time period. In this case, the number of new cases of CHD, including myocardial infarction (MI), comprise the incident cases.

EVALUATION OF PAPERS

The papers are evaluated according to the criteria defined previously (1,2). Specifically, the evaluation criteria include:

1. Was an inception cohort assembled?

An inception cohort is a group of people assembled at a specific time period in relation to the natural history of the disease. For studies of natural history or prognosis, this should be near the onset of the disease (1). Furthermore, this should be a uniform point along the natu-

ral history, such as when the first unambiguous symptoms or signs appear, so that patients who die or recover from the disease are included within the study group (2).

2. Was the referral pattern described and appropriate?

The entry point of the subjects into the study should be described (2). Patients in a tertiary care hospital tend to have more advanced and severe disease than patients with the same conditions typically seen in the primary care setting and the results must be generalizable to every patient population (2).

3. Were objective outcome criteria developed and used?

The specific prognostic outcomes studied should have been provided in a consistent manner throughout the study and explained in the paper (2).

4. Was the outcome assessment appropriate and carried out "blinded"?

The outcome assessment should be performed by researchers who are "blinded" to the other features of the subjects (2). For example, diagnostic suspicion bias is bias imposed by a clinician who has prior knowledge that a patient has a particular prognostic factor of presumed importance. This may affect the interpretation of results (2).

5. Was complete follow-up achieved?

All members of the inception cohort should be accounted for and their clinical status known at the end of the follow-up period (2).

6. Was adjustment for extraneous prognostic factors carried out?

The reader should consider whether there are other factors present which may alter the results of the study by influencing the natural history or clinical course of a disease.

7. Were definitions given for potentially ambiguous terms?

8. Was there any evidence of assembly bias, migration bias or measurement bias in the methodology of the study?

Assembly bias is defined as the formation of groups of patients for study that differ in ways other than the factor(s) of interest of the study (1). Migration bias is defined as the bias occurring when patients in one cohort leave their original cohort, either moving to one of the other cohorts of the study or dropping out altogether (1). The loss of greater than 20% of the inception cohort is cause for concern as a substantial portion of the cohort is missing from the final analysis. This may adversely affect the outcome analysis (2). Finally, measurement bias is defined as the bias introduced into a study when patients in one cohort stand a better chance of having the outcome of interest detected than patients in another cohort (1).

STUDY ONE

Serum cholesterol and acute myocardial infarction: a case-control study from the GISSI-2 trial by Nobili *et al.* (3)

This is a case control study that approaches the problem of establishing the relationship between total serum cholesterol and the incidence of MI. Thus, the authors waited until after-the-fact, evaluating patients admitted to hospital because of a first MI and then checking the total serum cholesterol levels. This approach introduces a confounding variable in that the MI itself may have altered the total serum cholesterol levels from the true baseline values.

People who died from an MI were excluded from the study as they did not survive to be interviewed. Thus, the authors used a survival cohort and an assembly bias was introduced to the data analysis, as the sample did not include all cases. Also, the subjects included in the case cohort were "confirmed episodes of acute myocardial infarction." However, these criteria were not specified. This is a term which should have been defined within the text of the paper.

As controls, 1,106 patients in the same hospitals for acute conditions were recruited, including hospitalization for fractures and sprains, non-traumatic orthopedic disorders, surgery and acute infections. The exclusion criteria were not stated. The authors specified that the controls had similar distributions of sex, age and hospital admission. However, a potential problem exists with using these patients: do their conditions or the treatment of their conditions affect total serum cholesterol levels? For example, does the trauma of anaesthesia and surgery alter cholesterol metabolism? Certainly these procedures would have systemic effects, most notably the drugs used for anaesthesia and the changes in diet prior to and immediately following surgery. This could bias the results of the control population. An alternate choice would have been to use patients randomly selected from the community who had not had an MI and had similar demographics.

Blood samples were taken for fasting total serum cholesterol levels "as soon as possible after admission to hospital." This is a vague description: it could mean the day of admission or several days later. A range should have been defined and followed to minimize the potential for differences due to changes in diet after admission to hospital. Furthermore, the authors did not specify if the lab technicians were "blinded" as to the original diagnosis and the baseline status of each patient when measuring the total serum cholesterol levels. This could lead to expectation bias, with the technicians unconsciously biasing the results, perhaps expect-

ing people who were just diagnosed with MI to have elevated cholesterol levels.

There was an unacceptably low follow-up achieved in this study. Total serum cholesterol values were available for only 614 cases (67.0% of the original cohort) and 792 controls (71.6% of the original cohort). The reasons for this high rate of subject loss was not clearly stated. If all of the 302 missing cases had their cholesterol in the "less than 171mg/dl" quintile, the linear increase in risk for MI with increasing cholesterol trend would have been completely lost. It is, of course, unlikely that all of the cases would fall in this quintile but this demonstrates how the distribution and conclusions might change if follow-up of all of the cases was achieved.

For outcome analysis, the authors did adjust for extraneous factors by calculating relative risk for cholesterol when other major risk factors for MI were present. However, given the problems with this study in terms of assembly of the inception cohort, lack of definitions for potentially ambiguous terms and incomplete follow-up, the outcome assessments must be taken with caution.

STUDY TWO

Is the relationship between serum cholesterol and risk of premature death from CHD continuous and graded? by Stamler *et al.* (4)

This study approaches the problem of establishing the relationship between total serum cholesterol and a first incidence of MI with a prospective natural history study. In this study, an inception cohort was formed of 361,662 men aged 35-57 years over 18 US cities. The referral pattern consisted of people recruited by screening employee, civic and church groups, house-to-house canvassing and screening of respondents to mass media publicity. Of this original group, 356,222 (98.5%) were used for the study. People excluded from the analysis included those with a history of hospitalization for MI.

For objective outcome assessment, the total serum cholesterol was measured at 14 local laboratories using an automated system of chemical analysis with standardization by the Lipid Standardization Program of the Centers for Disease Control. For the assessment of diastolic blood pressure, a standardized protocol was defined in the paper and applied by certified technicians. Smoking was assessed by a short questionnaire and the characteristics ascertained were described. Definitions for potentially ambiguous terms such as "absolute risk" were included.

In this study, death from MI was the outcome of interest. Complete follow-up is a strength of this study, as the vital status of 100% of the cohort is being obtained on an ongoing basis from the US Social Security System and the National Death Index. Assessment for the cause of death was obtained from the death certificates and coded by one of two trained nosologists. The criteria used to classify the deaths was referenced as the International Classification of Diseases.

Adjustment for extraneous prognostic factors was carried out. The mortality rates were stratified under the categories of smoker or non-smoker and diastolic blood pressure above or below 90mm Hg. Standardization for age was carried out and 95% confidence intervals were calculated for each quintile. However, it was not stated whether or not any tests of statistical significance were carried out on the comparison of the quintiles. It seems that the authors compared relative risks within each of their stratification categories to make their conclusions. Although the trend seems clear that there is an increasing death rate with increasing total serum cholesterol and that smoking and a diastolic blood pressure greater than 90mm Hg increase the risk of death from MI, a test of significance should have been carried out.

Thus, this is a well-organized and executed study. The conclusions drawn by the authors should be valid.

STUDY THREE

Ten-year mortality from cardiovascular disease in relation to cholesterol level among men with and without preexisting cardiovascular disease by Pekkanen *et al.* (5)

This study approaches the assessment of total serum cholesterol and mortality from CHD with a prospective natural history study. This study analyzed a subset of people assembled for a larger study. The details of the selection of subjects were well explained. The selection took place through a variety of steps and the exclusion criteria were stated at each step. Initially, 81,926 people were invited to participate in the larger study by sampling households, schools, businesses and one medical practice. Of these people, 60,502 accepted (73.8%) and 16,335 (19.9% of the original invitations and 27.0% of the accepted invitations) were invited to continue. For the second phase, patients were selected by falling into one of three groups: a 15% random sample of participants accepting the initial invitation, people with elevated lipid levels without lipid lowering medications or people already on lipid-lowering medications. Eighty-five percent of these people decided to continue to the second phase.

For this study, a subset of people from the final inception cohort assembled for the larger study were selected on the basis of being male and between the ages of 40-69 to form the inception cohort, a total of 2,890 people. There were 349 people excluded from further study and the exclusion criteria were well-defined, leaving a study population of 2,541 (87.9% of the original

group selected for the subset study). These subjects were stratified into categories based on whether they had a definite history of MI, abnormal exercise test but no definite MI, other manifestations of CHD but no definite MI or abnormal exercise test, no evidence of CHD, and unclassifiable. The cohort was followed for a mean of 10.1 years.

Near complete follow-up was achieved, with the current vital status of 99.6% of the study subjects known. Objective outcome criteria were used, with cause of death being the outcome of interest. Two cardiologists, who were "blinded" as to the person's identification and base-line characteristics, coded the cause of death as CHD, cardiovascular disease (CV) disease, or "other".

Adjustment for extraneous prognostic factors was carried out by statistically adjusting for different manifestations of CV disease at baseline. Once the data was collected, it wasgrouped for age, body mass index, current smoking, hypertension and level of physical activity. Descriptions of the tests used were included. Also, the results were stratified for the presence or absence of CV disease at baseline, thereby assessing if elevated total serum cholesterol was a risk factor for a second MI. This allows one to address the need for secondary prevention of MI.

Thus, the authors did a thorough analysis of the data and accounted for other risk factors for MI. This study validly followed the natural history of the association between total serum cholesterol and the incidence of MI. Also, as the study population includes people representing all stages of CV disease, the results should be applicable to the population at large.

CONCLUSION

This paper has presented a critical analysis of three studies which investigated the association between total serum cholesterol and the incidence of CHD. The results are summarized in Table 1.

All three of the papers came to the same primary conclusion that there was an increased incidence of CHD with increased total serum cholesterol levels. However, the paper by Nobili *et al.* (3) has many flaws which invalidate their conclusions. A case control study is not the ideal design for studying the natural history of a phenomenon, as the parameters may change with time.

Table 1: *Study analysis summary*

	Nobili et al. (3)	Pekkanen et al. (4)	Stamler et al. (5)
Study type	Case control	Prospective	Prospective
Representative inception cohort	Yes	Yes	Yes
Objective outcome criteria	Yes	Yes	Yes
Consistent and objective outcome assessment	No	Yes	Yes
Blinded outcome assessment	No	Yes	N/A
Complete follow-up	No	Yes	Yes
Adjustment for extraneous prognostic factors	Yes	Yes	Yes
Definition of ambiguous terms	No	Yes	Yes
Evidence for assembly bias	N/A	No	No
Evidence for migration bias	N/A	No	No
Evidence for measurement bias	Yes	No	No

This study also lacks consistent, objective, and "blinded" outcome assessment. The clinician will be unable to determine if the population studied is similar to his/her patient population. In contrast, the studies by Pekkanen *et al.* (4) and Stamler *et al.* (5) are better designed, have no evidence for bias and provide clear definitions so that the clinician may decide upon the applicability to his/her patient population.

Drawing conclusions from the studies by Pekkanen *et al.* (4) and Stamler *et al.* (5), there appears to be a well established trend of increasing incidence of first MI with increasing levels of total serum cholesterol in males. For primary prevention, total serum cholesterol levels can predict mortality due to CHD in middle-aged men (5). Higher levels of total serum cholesterol are associated with increased mortality from CHD and CV disease in men without previously known CV disease (4,5). Furthermore, the relationship between total serum cholesterol and CHD is a continuous and graded one. There is a continuous and progressive trend of higher death rates from CHD for men with increasing total serum cholesterol levels, with no threshold value to cross for increased risk of mortality from CHD (4). A 1% higher total serum cholesterol level is associated with an almost 2% higher risk of mortality from CHD (4). Also, the risk of CHD mortality at any level of total serum cholesterol is higher for cigarette smokers than non-smokers, with or without hypertension. Furthermore, a diastolic blood pressure of greater than 90mm Hg is associated with an increased risk of CHD mortality at any total serum cholesterol level (4). With respect to secondary prevention, higher levels of total serum cholesterol are associated with increasing mortality from CHD and CV disease in men with preexisting CV disease.

The ability to critically evaluate journal articles to determine which studies have valid results is a tool needed by every physician. In the exercise presented herein using publ;ished criteria (1,2), it appears that the measurement of total serum cholesterol is a relatively good indicator for the risk of CHD.

ACKNOWLEDGEMENTS

The author would like to thank Dr. Meng Hee Tan for his invaluable suggestions and guidance in the writing of this paper.

REFERENCES

1. Fletcher RH, Fletcher SW, Wagner EH. *Clinical Epidemiology: The Essentials.* 2nd ed. Baltimore: Williams & Wilkins, 1988.
2. Department of Clinical Epidemiology and Biostatistics, McMaster University Health Science Center. How to read clinical journals: III. To learn the clinical course and prognosis of disease. *Can Med Assoc J* 1981;124:869-72.
3. Nobili A, D'Avanzo B, Santoro L *et al.* Serum cholesterol and acute myocardial infarction: a case control study from theGISSI-2 trial. *Br Heart J* 1994;71:468-73.
4. Pekkanen J, Linn S, P. H, Heiss G *et al.* Ten-year mortality from cardiovascular disease in relation to cholesterol level among men with and without preexisting cardiovascular disease. *New Engl J Med* 1990;322:1700-7.
5. Stamler J, Wentworth D, Neaton JD. Is relationship between serum cholesterol and risk of premature death from coronary heart disease continuous and graded? *JAMA* 1986;256:2823-7.