

Level of aggregation for optimal epidemiological analysis: the case of time to surgery and unnecessary removal of the normal appendix

S W Wen, K Demissie, D August, G G Rhoads

Abstract

Study objective—To illustrate the concept of “individualised fallacy”, the result of improper interpretation and inference about aggregate level associations on the basis of associations at the individual level, in epidemiology.

Design—Cohort study.

Setting—Canadian province of Ontario.

Patients—All patients who underwent primary appendectomy in 175 Ontario hospitals from 1989 to 1992. The association between rate of normal appendix removal and time to surgery was analysed at two levels: (1) at individual patient level, in which, for each patient, the exact number of days to surgery was derived, and (2) at hospital level, in which hospital specific proportions of time to surgery was calculated.

Main results—Measured at individual level, compared with patients who had an operation on the same day of admission, the odds ratio was 2.41 (95% confidence intervals 2.28, 2.56) for patients who had an operation > 1 day after admission. Measured at hospital level, each 10% increase in the proportion of patients who had an operation > 1 day after admission resulted in a 15% reduction in the odds of normal appendix removal (odds ratio 0.85, 95% confidence intervals 0.82, 0.88)

Conclusions—In this case study, hospital level measure correctly predicted a reduction in the rate of normal appendix removal by delaying surgery, whereas individual level measure biased the direction of the relation to the opposite. This example illustrates that bias in across level inference can occur either at individual or ecological level. The preferred level of analysis is the one that minimises confounding; often, it must be selected on the basis of a priori knowledge of the subject area.

(J Epidemiol Community Health 2001;55:198–203)

Measurement of health outcomes and determinants may occur at various levels of aggregation (patient, physician, institution).^{1–5} Data collected from individual patients can also be aggregated at group level and form the basis for group level analysis. Individual and group level variables may capture different domains of the same construct or may represent a different construct. Across level inference is a phenomenon whereby inference about one level of

analysis is made on the basis of associations observed at a different level.^{2–7} The most commonly discussed across level inference bias in epidemiology is “ecologic fallacy”, which is the result of improper interpretation and inference about individual level associations based on associations at the aggregate level.^{8–11} Theoretically, bias in across level inference can occur in either direction. Recently, the notion of “individualised fallacy”, the results of improper interpretation and inference about aggregate level associations on the basis of associations observed at the individual level, has gained some attention in epidemiology.^{12–14} However, concrete examples illustrating this phenomenon are lacking in the epidemiological literature.

In a study of risk factors for unnecessary normal appendix removal, based on a hospital discharge data from the Canadian province of Ontario, we hypothesised that the degree of aggressiveness towards performing surgery for a diagnosis of appendicitis is an important determinant of the rate of unnecessary normal appendix removal. Specifically, we expected the rate to be higher among patients who received immediate surgery after admission to a hospital, and the rate to be lower among patients whose surgery was delayed until the diagnosis become clearer.

In a preliminary analysis of the association between time to surgery and rate of normal appendix removal, when analysis was performed at patient level, we found the proportion of normal appendix removal was not lower but instead substantially higher among patients who waited longer in hospital. This finding was reversed when hospital level analysis was performed instead, with a higher proportion of normal appendix removal among hospitals that perform surgery on average shortly after the patient is admitted in comparison with those hospitals that have an average longer time to surgery.

The aim of this paper was to examine the reasons for this paradox and to emphasise the need for external criteria in selecting the optimal level for analysis.

Methods

In the Province of Ontario, records of all acute care hospital separations (discharges, transfers and in-hospital deaths) have been maintained on computer tapes by the Hospital Medical Records Institute (now called the Canadian Institute for Health Information). Discharge summaries, operative notes, and pathology

Bureau of
Reproductive and
Child Health, Ottawa,
Ontario, Canada
S W Wen

UMDNJ-Robert Wood
Johnson Medical
School, USA
K Demissie
D August
G G Rhoads

UMDNJ-School of
Public Health,
Piscataway, USA
K Demissie
G G Rhoads

Correspondence to:
Dr Wen, Bureau of
Reproductive and Child
Health, Centre for Healthy
Human Development,
Tunney's Pasture A-L
0701D, Ottawa, Ontario,
Canada K1A 0L2
(Shi_Wu_Wen_@hc-sc.gc.ca)

Accepted for publication
31 October 2000

reports are routinely coded by trained coders. Diagnoses codes follow the International Classification of Diseases, Ninth Revision (ICD-9)¹⁵ and procedure codes follow the Canadian Classification of Diagnostic, Therapeutic and Surgical Procedures.¹⁶

Records were selected for all Ontario residents with full demographic data who underwent a primary appendectomy for fiscal years 1989 to 1992 (1 April 1989 to 31 March 1990). Procedure and diagnosis codes were used to identify patients whose appendix was removed after the diagnosis of appendicitis. If a patient has a primary discharge diagnosis code for acute appendicitis (ICD-9: 540.0; 540.1; 540.9; 541) and at the same time a primary procedure code of 59.0 (for appendectomy), then he or she is considered to have had appendectomy for the diagnosis of acute appendicitis. Exploratory laparotomy or incidental appendectomy were not included in this definition. The diagnosis of appendicitis was based on pathology reports or surgical notes, or both. Unnecessary appendectomy is defined as primary appendectomy for which surgical and pathology reports indicate a normal appendix after surgery. Patients with subacute, chronic or relapsing appendicitis for whom appendectomy was performed were excluded from the analysis because of ambiguity about diagnostic criteria.¹⁷

Time to surgery (preoperative length of in-hospital stay) was calculated by subtracting the date of admission from the date of surgery. The variable time to surgery was validated by comparing its value with the postoperative length of stay (from date of surgery to date of discharge) and total length of stay (from date of admission to date of discharge). Variables used for the validation were recorded separately in the dataset. If the time to surgery was less than zero, or if the sum of time to surgery and postoperative length of stay was not equal to the total length of stay, time to surgery was assigned a missing value. According to this rule, 23 cases received a missing value for this variable.

The dependent variable of interest in this analysis was normal appendix removal and the independent (explanatory) variable was time to surgery. The explanatory variable (time to surgery) was measured at two levels: (1) at individual patient level, in which, for each patient, the exact number of days to surgery was derived, and (2) at hospital level, in which hospital specific proportions of time to surgery was calculated. We assumed the patient level variable reflects the ambiguity in clinical presentation while the hospital level variable is likely to measure the practice style and practice guidelines of the specific hospital. We were aware of the intra-hospital variations in surgical style among surgeons within the same hospital. However, on average, hospitals with larger proportion of delay to surgery would be more conservative.

The patient's age, sex, and admission status were considered in the analysis as appropriate. A comorbidity index for each patient was calculated using Deyo's method. This method

adapts the Charlson's index of comorbidity for ICD-9-cm based diagnoses.^{18,19} In the study of treatment outcomes, one potential confounding factor is the difference in secondary or pre-existing diseases among individuals and groups. The comorbidity index is designed to reflect these coexisting conditions, so that some adjustments can be made when comparing outcomes. In our study sample, comorbidity index related conditions were identified from up to seven additional discharge diagnoses (second to eighth diagnosis); the first diagnosis was used to define appendicitis status. Ontario discharge abstracts permit standard calculation of this comorbidity index in nearly all cases.²⁰ Information on hospital bedsize and teaching status (presence or absence of house staff) was taken from the Canadian Hospital Directory, 1991–1992.²¹

Univariate and multiple logistic regression analysis was performed to examine the association between various hospital and patient factors with the rate of normal appendix removal. To assess the differential effect of time to surgery measured at the individual level compared with that measured at the hospital level, four logistic regression models were used: model including time to surgery at the individual level alone, model including time to surgery at the hospital level alone, model including simultaneous time to surgery at individual level and hospital level, and model including nested time to surgery at the individual level and the hospital level. To facilitate the interpretation of the results from the nested logistic regression model, dummy variables were created for the four hospital levels (<15% versus 15% to <18% versus 18% to <25% versus $\geq 25\%$) of prolonged time to surgery. As an initial analysis revealed no interaction between hospital and individual measures of time to surgery, and as the hospital effect of $\geq 15\%$ overlapped in large extent, a dichotomised hospital measure of time to surgery (<15% versus $\geq 15\%$) was used in the final nested logistic regression model. Regression analyses were repeated after excluding patients admitted to low volume hospitals (less than 20 primary appendectomy cases during 1989–1992), or excluding patients admitted electively. Statistical Analysis System (SAS) software was used to perform the analysis.

Results

During the four years of study, 35 891 primary appendectomies with acute appendicitis (positive) and 7170 primary appendectomies with diagnoses unrelated to the appendix (negative) were performed in 175 Ontario hospitals. The overall rate of normal appendix removal was 16.7%.

All of the patient factors studied were related to the rate of normal appendix removal, with substantially increased rates in those patients admitted electively, or those of female or older patients, or those with comorbidity (table 1). The strongest predictor of normal appendix removal was the admission status. More than half of the patients admitted electively had a normal organ removed, compared with 13.8%

Table 1 Hospital and patient factors related to rate of normal appendix removal

Factors	Number of primary appendectomy performed	Number with a removal of normal appendix	% With a removal of normal appendix	p Value
Primary appendectomy volume of the admitted hospital				
<70	1 227	153	12.5	<0.001
70–184	5 244	936	17.9	
185–350	9 335	1519	16.3	
>350	27 255	4562	16.7	
Bed size of the admitted hospital				
<50	3 045	456	15.0	<0.05
50–119	8 238	1331	16.2	
120–199	13 892	2365	17.0	
≥200	17 884	3017	16.9	
Teaching status of the admitted hospital				
Teaching	17 021	2924	17.2	<0.05
Non-teaching	26 038	4245	16.3	
Sex				
Male	23 262	2445	10.5	<0.001
Female	19 799	4725	23.9	
Age (y)				
0–14	8 728	968	11.1	<0.001
15–29	17 610	2833	16.1	
30–44	9 721	1750	18.0	
≥45	7 002	1619	23.1	
Admission status				
Emergent	32 110	4444	13.8	<0.001
Urgent	8 462	1405	16.6	
Elective	2 489	1321	53.1	
Comorbidity				
No	40 992	6566	16.0	<0.001
Yes	2 036	600	29.5	
Time to surgery (at individual level)				
0 day	32 752	4023	12.3	<0.001
≥1 day	10 265	3126	30.5	
Time to surgery (at hospital level; measured by percentage of ≥1 day to surgery)				
<15%	8 276	1502	18.2	<0.001
15% to <18%	13 100	2004	15.3	
18% to <25%	14 439	2452	17.0	
≥25%	7 246	1212	16.7	

and 16.6% for those admitted emergently or urgently, respectively (table 1). The hospital primary appendectomy volume, bed size, and teaching status, on the other hand, had only a modest effect on the rate of normal appendix removal (table 1). The effect of time to surgery on the rate of normal appendix removal was dependent on how it was measured; at the individual patient level, the rate in those who had an operation > 1 day after admission was more than double those who had an operation on the same day of admission. At the hospital level, the rate decreased as the percentage of patients who had an operation > 1 day increased (table 1).

At either patient or hospital level analysis, adjustment for important patient and hospital factors did not change the results (table 2). The direction of the relation between time to surgery and rate of normal appendix removal depended on whether it was measured at the individual patient level or hospital level. Measured at the individual level, those patients who had an operation > 1 day after admission had a twofold higher rate of normal appendix removal than those who had an operation on the same day of admission. Measured at the hospital level, each 10% increase in the proportion of patients who had an operation > 1 day after admission resulted in an 8%

Table 2 Odds ratio obtained from multiple logistic regression analysis for the rate of normal appendix removal in 43 061 patients undergoing primary appendectomy in Ontario, 1989/90–1992/93

Independent variables	Model I: Time to surgery analysed at individual level only	Model II: Time to surgery analysed at hospital level only	Model III: Time to surgery analysed at both individual and hospital level
Time to surgery (individual level) - ≥1 day v 0 day	2.33 (2.20, 2.46)	NA	2.41 (2.28, 2.56)
Time to surgery (hospital level) - each 10% increase in ≥1 day of the admitted hospital	NA	0.92 (0.89, 0.96)	0.85 (0.82, 0.88)
Primary appendectomy volume - 50 cases increase	1.01 (1.01, 1.02)	1.01 (1.00, 1.01)	1.01 (1.00, 1.01)
Bed size - 30 bed increase	0.99 (0.98, 0.99)	1.00 (0.99, 1.00)	1.00 (0.99, 1.00)
Treated in teaching hospital - no=0; yes=1	1.03 (0.97, 1.09)	1.05 (0.99, 1.11)	1.04 (0.98, 1.10)
Sex - male=0; female=1	2.46 (2.33, 2.60)	2.56 (2.42, 2.70)	2.46 (2.33, 2.60)
Age -10 year increase	1.09 (1.07, 1.10)	1.11 (1.10, 1.13)	1.09 (1.07, 1.10)
Admission status - emergency or urgent=0; elective=1	4.37 (3.99, 4.78)	5.76 (5.28, 6.29)	4.43 (4.05, 4.85)
Comorbidity - none=0; any=1	1.71 (1.52, 1.91)	1.83 (1.63, 2.04)	1.68 (1.50, 1.88)

Numbers in parentheses represent 95% confidence intervals; NA: not applicable.

Table 3 Odds ratio obtained from multiple logistic regression analysis, nesting hospital and individual patient level measures of time to surgery

Hospital level (% in ≥ 1 day to surgery of the admitted hospital)	Patient level	
	0 day	≥ 1 day
<15	1 (...)*	2.74 (2.39, 3.13)
≥ 15	0.78 (0.72, 0.85)	1.80 (1.65, 1.96)

*Reference group. Numbers in parentheses represent 95% confidence intervals.

reduction in the rate of normal appendix removal (table 2). When the two measures of time to surgery were analysed simultaneously, the effect for individual level measure increased to 2.41 while the effect for hospital level measure decreased to 0.85 (table 2). Nested logistic regression analysis obtained similar results: given the hospital level of time to surgery, those patients who had an operation ≥ 1 day after admission had a higher rate of normal appendix removal than those who had an operation on the same day of admission, and, given patient's time to surgery, hospitals with $\geq 15\%$ of patients who had an operation ≥ 1 day after admission had a substantially lower rate of normal appendix removal than the hospitals with $<15\%$ of patients who had an operation ≥ 1 day after admission (table 3).

Results obtained from analysis after excluding patients admitted to low volume hospitals or excluding patients admitted electively were similar (data not shown).

Discussion

Data at different levels of aggregation are common in epidemiology. For example, in studies of clinical progress, predictors can be measured at individual patient level such as age and sex, or at physician level such as specialty and attitude, or at institution level such as teaching status and policy. The distinction of the level of measurement is not clear under certain circumstances. Firebaugh⁴ suggested that variables be divided into two broad groups: those defined at individual levels and those defined at the aggregate level. He referred to the latter as macro-properties. Macro-properties may be aggregates of individual level data (for example, proportion of patients with certain features in a clinic or institute) or may be non-aggregated variables (for example, the presence of a policy). Often this concept becomes ill defined when the macro-properties are aggregates of individual level data. For example, physician's attitude or practice style is seldom directly recorded in a medical chart and other clinical records. On the other hand, indirect measures of practice style can be derived from patients record through a summary index measure. However, this derived index at institution level may fundamentally differ from the original index measured at patient level, because the former measures practice style while the latter measures patient characteristics.

In the case study presented in this paper, we found an opposite effect of time to surgery when analysis was performed at patient and hospital levels. At the patient level analysis, the odds of normal appendix removal was in-

KEY POINTS

- Measurement of health outcomes and determinants may occur at various levels of aggregation (patient, physician, institution).
- Across level inference is a phenomenon whereby inference about one level of analysis is made on the basis of associations observed at a different level.
- The most commonly discussed across level inference bias in epidemiology is "ecological fallacy".
- In this study, we used a concrete example to illustrate "individualised fallacy", the result of improper interpretation and inference about aggregate level associations on the basis of associations at the individual level.
- The study case emphasises that bias in across level inference can occur at either individual or ecological level in epidemiological studies with data measured at different levels of aggregation, and the preferred level of analysis is the one that minimises confounding; often, it must be selected on the basis of a priori knowledge of the subject area.

creased almost twofold in those patients who had surgery at ≥ 1 day after admission in comparison with those who had an operation on the same day of admission. On the other hand, hospital level analysis revealed an 8% reduction in the odds of normal appendix removal with each 10% increase in the hospital specific proportion of patients who had an operation at ≥ 1 day after admission. This paradox may be the result of two forces that delay surgery in patients with suspected appendicitis: (1) the ambiguity in the clinical picture and the subsequent difficulty in making a rapid and accurate diagnosis,²² which was influential at patient level, and (2) the more conservative attitude towards surgical treatment of patients with suspected appendicitis,²²⁻²⁴ which was the subject of our hypothesis. The latter effect was confounded by the former at the individual level of analysis, but was shown more clearly at the aggregate level where such confounding was minimised. Because referral bias is not large for appendicitis cases,²⁵ the hospital specific proportion of patients who had an operation ≥ 1 day after admission may be a good indicator of the conservative attitude towards surgical treatment and watchful waiting in hospitals where the patient is admitted. The argument that there might be two components causing delayed surgery was strengthened by the fact that, when the two measures of time to surgery were analysed simultaneously, the magnitude of effect was substantially increased in its own direction (odds ratios changed from 2.33 to 2.41 and from 0.92 to 0.85 for measures at the individual level and hospital level, respectively), indicating that the "independent" individual effect and "independent" hospital effect became more demonstrable when they were adjusted each other.

Actually, all patient factors related to an increased rate of normal appendix removal, such as being admitted electively, being female, older age, or with coexisting diseases, have somehow reflected the ambiguity in the clinical picture and the difficulty in diagnosis and therefore the delay in surgery.²² The position that watchful waiting for patients with no clear cut symptoms and signs of appendicitis could reduce normal organ removal is also consistent with previous observation^{23 24} and common wisdom. We realise that there may be substantial intra-hospital variations in surgical style among surgeons within the same hospital. The effect of surgical style would be more evident if we could calculate surgeon specific proportion of delayed surgery.

It may be useful to elaborate a little further on the issue of clinical management of patients with suspected appendicitis. Diagnostic accuracy may be improved through use of diagnostic algorithms,²⁶ more reliance on watchful waiting,²³ or conservative attitude to surgical exploration in uncertain cases, or all three.²⁴ Reliance on watchful waiting may delay laparotomy and risk organ perforation. Previous studies have found that the perforation rate increased with diagnostic accuracy,^{27 28} and postoperative mortality was higher and length of in-hospital stay longer in perforated cases than non-perforated cases.²⁸⁻³⁰ Velanovich and Satava suggested a more aggressive surgical approach to patients with suspected appendicitis, despite more frequent removal of normal appendixes, because an increased diagnostic accuracy (presumably resulted from a delay in surgery) has been related to adverse postoperative outcomes such as perforation.²⁷ On the other hand, Andersson and colleagues suggested that as perforating and non-perforating appendicitis seem to be separate entities, and because self resolution is common in appendicitis, more expectant management for suspected appendicitis is needed to reduce the toll from normal appendix removal.³¹ In a large population-based study, we also observed a positive relation between diagnostic accuracy and perforation, but diagnostic accuracy was unrelated to either in-hospital death rate or length of in-hospital stay.¹⁷ How can one interpret these seemingly contradictory findings? We hypothesised that there are two categories of perforation. The first is represented by the person who arrives for medical care with obvious peritonitis and signs of generalised infection. The appendix in these people may well have perforated some time before medical attention was sought. It is not surprising that peritonitis will lead to delayed discharge and even death in these people. The other category is represented by the person with abdominal pain and a less clear cut clinical picture who is watched closely in hospital. Some of these people may have appendicitis that resolves spontaneously; others may have a different diagnosis. However, when observation and investigations add support to the diagnosis of acute appendicitis, surgery can be undertaken promptly. This type of watchful waiting, with deliberately delayed surgery, could well lead to both higher

accuracy and more perforation, but with modern antibiotics, the clinical consequences of “controlled perforation” would be small. The fact that in this study, admission status was the strongest predictor of normal appendix removal, with a fourfold to fivefold increase in patients admitted electively as compared with patients admitted emergently or urgently provides further support for the safety of delaying surgery in certain cases. Appendicitis is an acute disease that should normally be admitted emergently or urgently. Two types of patients might be admitted electively: those with a slow onset and no clear cut symptoms or signs at the very beginning (in which case the appendix might be truly “normal”), and those with an acute onset that was somehow self resolved before seeking hospital care (in which case the appendix might become “normal” while waiting at home or during an office visit).

Analysis of data collected at both individual and ecological levels can be performed using two approaches: aggregating all the individual level data (dependent and independent variables) so that the entire analysis can be performed at the ecological level, or limiting the dependent variable at the level of individual, but attaching ecological covariates to each individual. In a previous study we performed two parallel analyses: aggregating all the individual level data and performing the analysis at the level of hospital, and individualising all hospital variables and performing the analysis at the level of individual, but attaching ecological covariates to each person.¹⁷ That study is much similar to the current one: same determinant (delay in surgery), same covariables (for example, age, sex, coexisting diseases), and same population. The only difference between that study and the current one is the outcome: mortality and complications for the previous study, and normal appendix removal for the current one. The results obtained from the two parallel analyses in the previous study were in the same direction and similar magnitude. However, by forcing all individual variables into hospital specific means or proportions, the power of the individual level variables in predicting outcomes was greatly reduced. Many of the clinically meaningful and strong outcome predictors, such as age, sex, and coexisting disorders, became non-significant in the models aggregating all individual level variables into hospital level variables.¹⁷ As a result, we elected to conduct the current analysis at patient level, and attaching hospital covariables to each individual.

When there is an obvious and natural order to the hierarchy of variables, multilevel analysis is often a preferred option. Without such a priori knowledge, however, the results obtained from multilevel analysis could be misleading. For this study, we have difficulty to construct a multilevel regression model, as we were not sure how the two sets of variables interact to each other. Despite the fact that we were unable to use a more sophisticated modelling such as multilevel analysis, our study results seemed valid. The estimated effects for important prognostic variables (for example, being

admitted electively, being female, older age, or with coexisting diseases) were in the right direction and of the expected magnitude. The regression models using different combinations of time to surgery yielded similar results, demonstrating the robust of the observed associations.

The empirical evidence presented in this paper argues strongly that analysis at the individual level is not always preferable, and that the selection of an appropriate level of aggregation should be guided by external information about the likelihood of confounding. As described by Bidwell and Kasard, issues of measurement cannot be considered apart from issues of theory and conceptualisation.³² When considering a specific research question the investigator should first conceptualise the probable sources of confounding that may operate at different levels of analysis. The choice of an appropriate level will minimise confounding but will also be influenced by the extent to which suspected confounding can be measured and controlled. In this regard, the earlier works of Entwisle and colleagues in sociology are exemplary.³³⁻³⁵

The authors wish to thank the help from the following persons: Mr Marc-Erick Theriault for computer programming, and Dr Catherine McCourt for her support for this study. Endorsement by the sponsoring agencies is not implied.

Funding: none.

Conflicts of interest: none.

- 1 Von Korff M, Koepsell T, Curry S, *et al.* Multi-level analysis in epidemiologic research on health behaviors and outcomes. *Am J Epidemiol* 1992;135:1077-82.
- 2 Firebaugh G. A rule for inferring individual-level relationships from aggregate data. *Am Sociol Rev* 1978;43:557-72.
- 3 Hannan MT. *Aggregation and disaggregation in sociology*. Lexington, MA: Lexington Books, 1971.
- 4 Firebaugh G. Assessing group effects: a comparison of two methods. *Sociology Methods Research* 1979;7:384-95.
- 5 Haney W. Units and levels of analysis in large-scale evaluation. In: Roberts KH, Burstein L, eds. *New directions for methodology of social and behavioral science: issues in aggregation*. San Francisco, CA: Jossey-Bass, 1980:1-16.
- 6 Hauser RM. Context and consex: a cautionary tale. *American Journal of Sociology* 1970;75:645-64.
- 7 Glick W. Problems in cross-level inferences. In: Roberts KH, Burstein L, eds. *New directions for methodology of social and behavioral science: issues in aggregation*. San Francisco, CA: Jossey-Bass, 1980:17-20.
- 8 Greenland S, Morgenstern H. Ecological bias, confounding, and effect modification. *Int J Epidemiol* 1989;18:269-74.
- 9 Piantadosi S, Byar DP, Green SB. The ecological fallacy. *Am J Epidemiol* 1988;127:893-904.
- 10 Richardson S, Stucker I, Hemon D. Comparison of relative risks obtained in ecological and individual studies: some methodological considerations. *Int J Epidemiol* 1987;16:111-20.
- 11 Greenland S, Robins J. Invited commentary: Ecologic studies—biases, misconceptions, and counterexamples. *Am J Epidemiol* 1994;139:747-60.
- 12 Susser M. The logic in ecological: I. The logic of analysis. *Am J Public Health* 1994;84:825-9.
- 13 Schwartz S. The fallacy of the ecological fallacy: the potential misuse of a concept and the consequences. *Am J Public Health* 1994;84:819-24.
- 14 Koopman JS, Longini IM. The ecological effects of individual exposures and nonlinear disease dynamics in populations. *Am J Public Health* 1994;84:836-42.
- 15 The International Classification of Diseases. 9th revision. *Clinical modification*. Volumes 1-3. Salt Lake City: Medicode Publications, 1995.
- 16 Statistics Canada. *Canadian classification of diagnostic, therapeutic, and surgical procedures*. Ottawa: Minister of supply and services Canada, 1986.
- 17 Wen SW, Naylor CD. Diagnostic accuracy and short-term surgical outcomes in suspected acute appendicitis. *Can Med Assoc J* 1995;152:1617-26.
- 18 Deyo RA, Cherkov DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative database. *J Clin Epidemiol* 1992;45:613-19.
- 19 Charlson ME, Pompei P, Ales KL, *et al.* A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chron Dis* 1987;40:373-83.
- 20 Chen E, Naylor CD. Variation in hospital length of stay for acute myocardial infarction in Ontario, Canada. *Med Care* 1994;32:420-35.
- 21 *Canadian Hospital Directory, 1991-1992*. Volume 39. Ottawa: Canadian Hospital Association, 1992.
- 22 Pass HI, Hardy JD. The appendix. In: Hardy JD, ed. *Hardy's textbook of surgery*. 2nd ed. Philadelphia: J B Lippincott, 1988:574-81.
- 23 White JJ, Santillana M, Haller JA. Intensive in-hospital observation: a safe way to decrease unnecessary appendectomy. *Am J Surg* 1975;41:793-8.
- 24 Thomson HJ, Jones PF. Active observation in acute abdominal pain. *Am J Surg* 1986;152:522-5.
- 25 Wen SW, Naylor CD. Hospital-specific and regional-level utilization information on appendectomy. In: Naylor CD, Anderson G, Goel V, eds. *Patterns of health care in Ontario*. Ottawa: CMA Publications Group, 1994:143-50.
- 26 Adams ID, Chan M, Clifford PC, *et al.* Computer-aided diagnosis of acute abdominal pain: a multi-centre study. *BMJ* 1986;293:800-4.
- 27 Velanovich V, Satava R. Balancing the normal appendectomy rate with the perforated appendicitis rate: implications for quality assurance. *Am J Surg* 1992;58:264-9.
- 28 Ricci MA, Trevisani MF, Beck WC. Acute appendicitis. A 5-year review. *Am J Surg* 1991;57:301-5.
- 29 Ryden CI, Grunditz T, Janzon L. Acute appendicitis in patients above and below 60 years of age. Incidence rate and clinical course. *Acta Chirurgica Scandinavica* 1983;149:165-70.
- 30 Luckmann R. Incidence and case fatality rates for acute appendicitis in California. *Am J Epidemiol* 1989;129:905-18.
- 31 Andersson R, Hugander A, Thukin A, *et al.* Indications for operations in suspected appendicitis and incidence of perforation. *BMJ* 1994;304:107-10.
- 32 Bidwell CE, Kasarda JD. Problems of multi-level measurement: the case of school and schooling. In: Roberts KH, Burstein L, eds. *New directions for methodology of social and behavioral science: issues in aggregation*. San Francisco, CA: Jossey-Bass, 1980:53-64.
- 33 Entwisle B, Mason WM. Multi-level effects of socioeconomic development and family planning programs on children ever born. *American Journal of Sociology* 1985;91:616-49.
- 34 Entwisle B, Mason WM, Hermalin AI. The multi-level dependence of contraceptive use of socioeconomic development and family planning program strength. *Demography* 1986;23:199-216.
- 35 Entwisle B, Hermalin AI, Kamnuansilpa P, *et al.* A multi-level model of family planning availability and contraceptive use in rural Thailand. *Demography* 1984;21:559-74.