Evaluating Gastric Cancer Misclassification: a Potential Explanation for the Rise in Cardia Cancer Incidence

Anna Mia Ekström, Lisa B. Signorello, Lars-Erik Hansson, Reinhold Bergström, Anders Lindgren, Olof Nyreö

Background: Reports of dramatic increases in gastric cardia cancer incidence warrant concern. However, the recent introduction of a separate diagnostic code, the lack of a consensus definition of the cardia area, and the accelerating interest in cardia cancer may affect classification practices. Little is known about the magnitude of cardia cancer misclassification in large cancer registries.

Methods: In a well-defined Swedish population (1.3 million), we uniformly classified all patients with newly diagnosed gastric adenocarcinoma (from 1989 through 1994) with respect to gastric subsite, and we used this patient group as our gold standard. We then evaluated the completeness of the Swedish Cancer Registry in registering gastric adenocarcinomas against this gold standard and, further, assessed the completeness of cardia cancer registration and the rate of falsely included cases to estimate the potential impact on observed incidence trends.

Results: Our gold standard contained 1337 case subjects with gastric adenocarcinoma. Overall, the Swedish Cancer Registry was 98% complete with regard to gastric adenocarcinomas and had a 4% rate of falsely included cases. The completeness of coding cardia cancer was only 69%, and the positive predictive value for cardia cancer was 82%, with no improvement over time. Results: Our gold standard contained 1337 case subjects with gastric adenocarcinoma. Overall, the Swedish Cancer Registry was 98% complete with regard to gastric adenocarcinomas and had a 4% rate of falsely included cases. The completeness of coding cardia cancer was only 69%, and the positive predictive value for cardia cancer was 82%, with no improvement over time. Conclusions: Although overall completeness of gastric cancer registration by the Swedish Cancer Registry was excellent, accuracy in registering cardia tumors was surprisingly low. Our estimates suggest that true cardia cancer incidence could be up to 45% higher or 15% lower than that reported in the Cancer Registry. This margin of error could accommodate the observed increase in cardia cancer in Sweden. Therefore, secular trends in cardia cancer incidence should be interpreted cautiously. [J Natl Cancer Inst 1999;91:786–90]

In sharp contrast to the remarkable downward trend in distal gastric cancer incidence since the second world war (1–4), there have been several reports of increasing rates of adenocarcinoma in the gastric cardia with an annual increase of up to 4%–5% in the United States (2,5,6). Swedish Cancer Registry (hereafter the Cancer Registry) data indicated an annual 2.5% increase from 1970 through 1985 (7). Other register-based observations from Western Europe (8–10) and Australia (11) of absolute and relative increases in incidence have strengthened the suggestion that cardia cancer is epidemiologically distinct from distal gastric cancer.

Until the late 1980s, relatively little emphasis was put on the distinction between cardia and distal gastric cancer, because the possibly fundamental difference between these two cancer sites had not yet become evident. Cancer of the cardia was coded separately from other gastric cancers only after the 8th revision of the International Classification of Diseases (ICD-8). In the United States, a separate code was first introduced in 1973 (in the Surveillance, Epidemiology, and End Results Program1). England (1960), Sweden (1970), and Switzerland (1970) began to use a code for cardia cancer somewhat earlier. Most countries, however, did not start using a code for cardia cancer until the late 1970s (France and Scotland, 1975; Italy, 1976; and The Netherlands, 1978). To date, there is no consensus in the definition of the cardia area. Conse-

Affiliations of authors: A. M. Ekström, O. Nyreö, Department of Medical Epidemiology, Karolinska Institute, Stockholm, Sweden; L. B. Signorello, Department of Medical Epidemiology, Karolinska Institute, and Department of Epidemiology, Harvard School of Public Health, Boston, MA; L.-E. Hansson, Department of Medical Epidemiology, Karolinska Institute, and Department of Surgery, Mora Hospital, Sweden; R. Bergström, Department of Medical Epidemiology, Karolinska Institute, and Department of Statistics, Uppsala University, Sweden; A. Lindgren, Department of Pathology, Falu Hospital, Falun, Sweden.

Correspondence to: Anna Mia Ekström, M.D., M.P.H., Department of Medical Epidemiology, Karolinska Institute, P.O. Box 281, S-171 77 Stockholm, Sweden (e-mail: Annamia.Ekstrom@mep.ki.se).

See “Notes” following “References.”

© Oxford University Press
ently, misclassification of gastric cancer site may not be negligible.

Before designating the gastric cardia as a cancer site of rapidly increasing incidence, it is crucial to address whether the observed increase can be explained by a shift in classification. The increased focus on cardia cancer in the literature could conceivably result both in a greater completeness of registration and in a higher rate of falsely included patients. To our knowledge, there is no published information about the magnitude of misclassification of cardia cancer in the high-quality cancer registers, which are the sources of the recent alarming reports. In the absence of a concurrent gold standard registration, retrospective analyses are uninformative.

We have explored the possibility that misclassification was a major cause of the reported increase in incidence of cardia cancer by taking advantage of the comprehensive organization for the rapid ascertainment of records of gastric cancer cases. This organization was created for a population-based case–control study in Sweden over the 6-year period from 1989 through 1994 (12). Our meticulous search for new cases of cardia cancer resulted in the identification of case subjects who were uniformly characterized with regard to cancer site. These patients then were used as our gold standard against which gastric cancer registration in the same well-defined population was validated.

**SUBJECTS AND METHODS**

**Ascertainment of Case Subjects With Gastric Cancer**

We identified all patients with newly diagnosed gastric cancer occurring in five counties (two northern high-risk counties and three southern average-risk counties; total population = 1.3 million) from 1989 through 1994 from the following sources: 1) contact persons at all hospitals in the study area during the entire period under study, 2) continuous surveillance of all patients with definite or suspected gastric cancer whose specimens were evaluated by the pathology departments of the hospitals in the study, and 3) monthly double checks with the regional cancer registries for patients with gastric cancer that occurred in the period under study. Of these cancers, 22 (2%) were gastric cancer tumors but not adenocarcinomas (16 carcinoids or endocrine tumors and six other tumors). The degree of mismatch between the Cancer Registry and the gold standard incidence of gastric adenocarcinoma is shown in the Venn diagram in Fig. 1. In total, 59 patients (4%) were falsely included as having incident gastric cancer (33 esophageal tumors, four other nongastric tumors, 20 nonmalignant lesions in the stomach, one recurring cancer, and one gastric tumor found first at autopsy). After these exclusions, 1307 patients with true gastric adenocarcinoma were recorded in the Cancer Registry. Our gold standard registration identified 1337 patients with adenocarcinoma. Thus, the completeness of the Cancer Registry with respect to gastric adenocarcinoma was 1307 (98%) of 1337 patients. The positive predictive value of a gastric adenocarcinoma diag-
Subjects had a mean age of 70.4 years (95% CI 65.7–75.0). There was no indication of improved classification over time; the rate of false inclusions was 3% (24 of 724 patients) from 1989 through 1991 and 5% (59 of 1162 patients) from 1992 through 1994. The rate of missed case subjects increased to 20% (16 of 79 patients) in the second half of the study. Thus, the number of patients with true cardia cancer recorded in the Cancer Registry was 123. Hence, the completeness of the Cancer Registry with respect to cancer of the gastric cardia was 123 (69%) of 178 patients. Of the 55 patients with cardia cancer missed by the Cancer Registry, 28 patients had their tumors classified as ICD-7 code 151.9 (site unspecified) and 18 cancers had been classified as ICD-7 code 151.0 (corpus and distal stomach). The remaining nine patients were among those completely unknown to the Cancer Registry. The positive predictive value of a diagnosis of cardia cancer in the Cancer Registry was 123 (82%) of 150 patients.

Misclassification due to inclusion of patients with false cardia cancer was slightly more common in men; 78% of the case subjects falsely included as having cardia cancer were males and 75% of the true cardia cancer patients were males. The rates of missed case subjects were 28% and 41% among men and women, respectively. The falsely included patients were older than the patients with true cardia cancer; mean (95% CI) ages were 75.7 years (95% CI = 71.7–79.7) versus 67.7 years (95% CI = 65.9–69.5). The mean age among the missed case subjects was 68.6 years (95% CI = 64.9–72.2). There was no improvement in the rate of false inclusions over time. The tendency was rather the opposite; a 10% (seven of 71 patients) false inclusion rate during the first half of the study increased to 20% (16 of 79 patients) in the second half of the study.

**Fig. 1.** Venn diagram illustrating the concordance between incidence data for gastric adenocarcinoma (all sites) in five Swedish counties, registered in the Swedish Cancer Registry (hereafter the Cancer Registry) from 1989 through 1994 (box with dotted lines), and our gold standard based on a meticulous prospective search for all patients with gastric adenocarcinoma during the same period, uniformly classified with respect to histology and site (box with solid lines). There were 1366 patients with adenocarcinoma recorded in the Cancer Registry and 30 patients missed by the Cancer Registry, which gives 1307 correctly registered case subjects. The positive predictive value of the Cancer Registry was 96% (1307 of 1366 patients); the completeness of the Cancer Registry was 98% (1307 of 1337 patients).

**Fig. 2.** Concordance between incidence data for cardia cancer in five Swedish counties, registered in the Swedish Cancer Registry (hereafter the Cancer Registry) from 1989 through 1994 (large box with dotted lines), and our gold standard data (box with solid lines) for the same period. There were 150 case subjects with cardia cancer recorded in the Cancer Registry and 178 patients in our gold standard. There were 23 falsely included patients and four case subjects with tumors unclassifiable as to site (may or may not be cardia cancer) in the Cancer Registry and 55 patients missed by the Cancer Registry, which gives 123 correctly registered case subjects. The positive predictive value of the Cancer Registry was 82% (123 of 150 patients); the completeness of the Cancer Registry was 69% (123 of 178 patients).
The completeness, on the other hand, went up. The rate of missed patients decreased from 36% (34 of 95 patients) during the first part of the study (from 1989 through 1991) to 25% (21 of 83 patients) during the second part (from 1992 through 1994).

**DISCUSSION**

The Cancer Registry's completeness (98%) and positive predictive value (96%) for gastric adenocarcinoma registration overall from 1989 through 1994 should be considered satisfactory; however, the rate of misclassification of the location within the stomach was substantial. When we used the rather conservative criteria for cancer of the gastric cardia (13) in the gold standard registration, we found that no more than 69% of these cancers were recorded in the Cancer Registry as being of cardia origin (ICD-7 code 151.1). Almost one third of the true cardia tumors were reported to emanate from the corpus/fundus (ICD-7 code 151.0) or to be gastric tumors without further specifying the location (ICD-7 code 151.9). Moreover, at least 15% of the patients diagnosed with cardia cancer in the Cancer Registry did not, in fact, fulfill our topographic criteria. The deficit due to misclassification of cardia cancers as noncardia cancers was nearly balanced by a false inclusion of noncardia tumors, and consequently the reported incidence of cardia cancer in the Cancer Registry was approximately the true incidence. Our data, however, clearly demonstrate that misclassification creates a range of uncertainty around the incidence figures that needs to be taken into account when interpreting the reports of dramatically rising incidence rates.

According to our estimations of the rates of false inclusions and missed patients, the true incidence could be as much as 45% \[I_{obs}/(1/100\%)] or 15% \[I_{obs} – I_{obs}(false inclusion rate)] lower than the rates reported by the Cancer Registry. In the absence of concurrent stringent gold standard data for periods other than the period under study, we are unable to draw any firm conclusions with regard to secular trends in rates of misclassification. Because a separate ICD code for cardia cancer was introduced in 1970, the failure to use this code accordingly resulted in a high rate of missed patients in the Cancer Registry. More recently, however, a growing awareness of the specifics of cardia cancer may instead have led to increasing rates of false inclusions.

As an example of the potential effects of the misclassification, we applied the levels of misclassification observed in our study to the data on secular trends in the incidence of cardia cancer in Sweden published by Hansson et al. (7). According to that study, for men, the incidence of cancer of the gastric cardia increased from 1.9 to 3.0 cancers per 10^5 person-years (person-years = the total observed number of years added over subjects at risk) between the first period from 1970 through 1973 and the second period from 1982 through 1985. With the extreme assumption that the rates of missed patients and falsely included patients in the early period were 31% (our observed rate) and 0%, respectively, versus 0% and 15% (our observed rate of falsely included patients) in the latter period, the first incidence rate could have been as high as 2.8 \[1.9 \times (1/0.69)] per 10^5 person-years, and the second incidence could have been as low as 2.6 \[3.0 – (3.0 \times 0.15)] per 10^5 person-years. Thus, the entire increase in incidence could theoretically be explained by misclassification, particularly because the rate of missed patients in the early 1970s may have been greater than that observed from 1989 through 1994.

Our concern over the possibility of spuriously inflated secular trends due to misclassification is supported by a separate analysis of time trends of gastric cancer during the same study period. In this analysis, with meticulous case ascertainment and uniform classification of tumor site, we found no signs of a trend in incidence for cardia cancer during the 6 years of observation (unpublished results).

The misclassification observed in our study could not explain the greater than threefold increases in incidence reported from other countries (5,8). However, because the Cancer Registry compares well with most other national registers, because the prerequisites for implementation of uniform diagnostic coding may have been particularly favorable in Sweden with its centralized health care system and relatively small medical community, and because the ongoing case–control study may have stimulated more stringent coding, our data may be a conservative estimate of misclassification in other countries.

Although our criteria for cancer of the gastric cardia are widely used, there is no worldwide consensus of what characteristics should be used to define this diagnostic category. Clearly, our criteria are conservative. If we had used more liberal criteria and classified gastric tumors with their center within 2 or 3 cm of the gastroesophageal junction, the completeness of the registration in the Cancer Registry would have been even lower (as would the rate of false inclusions). Thus, the problem of misclassification cannot be resolved by applying other diagnostic criteria.

We conclude that the observed increase in adenocarcinomas of the gastric cardia may have been inflated by misclassification, possibly fueled by an increase in diagnostic awareness. It may be of considerable interest to estimate the accuracy of site classification in other countries with a reported increase of cancer in the gastric cardia.

**REFERENCES**


NOTES

1Editor’s note: SEER is a set of geographically defined, population-based central tumor registries in the United States, operated by local nonprofit organizations under contract to the National Cancer Institute (NCI). Each registry annually submits its cases to the NCI on a computer tape. These computer tapes are then edited by the NCI and made available for analysis.

Supported by Public Health Service grant R01CA50959 from the National Cancer Institute, National Institutes of Health, Department of Health and Human Services.

We thank Lotti Barlow, The Cancer Registry, Centre for Epidemiology, The National Board of Health and Welfare, Sweden, for providing us with population data.

Manuscript received October 7, 1998; revised February 22, 1999; accepted March 5, 1999.