Research Article

Cervical Intraepithelial Neoplasia Grade 3 and Adenocarcinoma In Situ: Comparison of ICD-9 Codes and Pathology Results—Kaiser Permanente, United States, 2000–2005

Diya Surie1,2, Eileen F. Dunne1, Allison L. Naleway4, Sheila Weinmann4, Nicola P. Klein5, Roger Baxter6, Kathleen Hutchins1, Julianne Gee3, and Lauri Markowitz1

Abstract

Background: Cervical intraepithelial neoplasia grade 3+ (CIN3+) and adenocarcinoma in situ incidence will be an important measure of HPV vaccine impact. Integrated healthcare delivery systems, such as Kaiser Permanente, could be used to monitor CIN3+ trends; however, limited evaluations of data from healthcare delivery systems for CIN3+ surveillance exist.

Methods: We compared CIN3+ diagnoses by ICD-9 code with CIN3+ diagnoses by pathology results among 121,211 females aged 11 to 30 years who were continuously enrolled from 2000 to 2005 in either Kaiser Permanente Northern California or Kaiser Permanente Northwest. We calculated sensitivity and positive predictive value of diagnosis by ICD-9 codes using pathology CIN3+ diagnosis as the gold standard.

Results: There were 1,090 women with at least one CIN3+ diagnosis by ICD-9 code 233.1 and 1,200 women with at least one CIN3+ diagnosis by pathology results. The sensitivity of the ICD-9 code for detecting a woman with at least one pathology diagnosis for CIN3+ was 62% (740/1,200); positive predictive value was 68% (740/1,090). Among women with at least one CIN3+ diagnosis by ICD-9 code, 679 (62%) had more than one visit with this code; whereas, among women with at least one CIN3+ diagnosis by pathology, 466 (39%) had more than one CIN3+ pathology result.

Conclusions: ICD-9 codes may underestimate the number of women with at least one CIN3+ diagnosis.

Impact: Pathology results, when available, may provide better estimates of CIN3+ incidence. Cancer Epidemiol Biomarkers Prev; 1–4. ©2013 AACR.

Introduction

Cervical cancer is the third most common cancer in women worldwide (1) and is caused by human papillomavirus (HPV; ref. 2). Two oncogenic HPV types (16 and 18) cause more than 70% of cervical cancers and a proportion of other HPV-associated cancers and cancer precursors (3). Two HPV vaccines that prevent HPV types 16 and 18 are available and recommended for routine use in females in the United States (4, 5); 1 HPV vaccine (Gardasil) is recommended for routine use in males (6). Measuring the impact of these vaccines on HPV-associated diseases is challenging because routine surveillance is limited to monitoring only for cervical, vaginal, vulvar, anal, and oropharyngeal cancers (7). Cancer morbidity and mortality are tracked by cancer registries in the United States and will be an important measure of vaccine impact (8), but it may take decades to show reductions in cervical cancer rates.

The pathology diagnosis of cervical intraepithelial neoplasia grade 3 (CIN3) is the most immediate precursor to invasive cervical cancer. HPV types 16/18 are associated with about 52% of high-grade squamous intraepithelial lesions worldwide (9), and this proportion is higher in pathology confirmed high-grade lesions such as CIN3 and adenocarcinoma in situ (AIS). Clinical trials showed that HPV vaccines had high efficacy for prevention of CIN3+ (CIN3 or AIS); assessing this outcome could be a more proximal measure of vaccine impact than cancer. However, there is no national reporting of CIN3+ diagnoses or systematic collection of these data. Integrated healthcare delivery systems, such as Kaiser Permanente (KP) have both administrative and laboratory data (including pathology) on their enrollees, which may provide useful information for monitoring vaccine impact. Few assessments have critically evaluated data available through such healthcare delivery systems for this purpose.

Two potential data sources from KP that can be used to conduct surveillance for HPV-associated diseases
include administrative databases with International Classification of Diseases—ninth revision (ICD-9) codes and laboratory databases that include histopathologic or cytologic diagnoses from biopsies or Papanicolaou smears. These sources have different purposes including claims/billing, quality control, or laboratory assessments. Understanding the strengths and limitations of ICD-9 codes compared with laboratory data for surveillance could be useful for determining the best source of data to measure vaccine impact. This evaluation compares ICD-9 codes with pathology results for the cervical cancer precursor CIN3þ.

Materials and Methods

Data on CIN3þ diagnoses were collected from 2 sites: (1) Kaiser Permanente Northern California (KPNC), which provides comprehensive medical care for over 3 million members in the San Francisco Bay area, Sacramento, and Central Valley area; and (2) Kaiser Permanente Northwest (KPNW), which provides comprehensive medical care for more than 450,000 members in Northwest Oregon and Southwest Washington. These 2 sites are part of the consortium of 10 healthcare delivery systems that collaborate with the Centers for Diseases Control and Prevention (CDC) on vaccine safety assessments through the Vaccine Safety Datalink (VSD; ref. 10). As part of the VSD, data on demographics, medical diagnosis, and health services use are collected in standardized data files. For this assessment, the participating sites’ VSD outpatient data were augmented to include relevant ICD-9 codes and pathology results.

We included 121,211 females aged 11 to 30 years who were continuously enrolled at either site from 2000 to 2005 (13,385 from KPNW and 107,826 from KPNC). We defined CIN3þ based on ICD-9 codes or pathology results. The ICD-9 code (233.1) description for CIN3þ includes the following terms: AIS of the cervix; cervical intraepithelial glandular neoplasia grade 3; CIN3; and severe dysplasia of cervix; excluding CIN2; cytologic evidence of malignancy without histologic confirmation, high-grade squamous intraepithelial lesion, and moderate dysplasia of the cervix (11). In KPNC, pathology results from cervical biopsies were reviewed and text searches for CIN3, CIS, AIS, and Glandular abnormalities/HSIL were conducted and these cases were retained in the laboratory file. In KPNW, SNOP codes were used to search for pathology results from cervical biopsies for CIN3, CIS, AIS, and Glandular abnormalities/HSIL were conducted.

Data were analyzed using SAS v.9.2. We compared CIN3þ diagnoses by ICD-9 codes and pathology results. The sensitivity and positive predictive value (PPV) of the ICD-9 code for CIN3þ were calculated. For the purposes of this evaluation, “true positives (TP)” for the sensitivity and PPV calculations were defined as women who had at least one ICD-9 code for CIN3þ, as well as a pathology result for CIN3+. “False negatives (FN)” were defined as women who had at least one pathology result for CIN3þ but did not have an ICD-9 code for CIN3+. “False positives (FP)” included women who had an ICD-9 code for CIN3þ but did not have a pathology result for CIN3þ. On the basis of these definitions, the sensitivity and PPV were calculated using the standard formulas:

Sensitivity = TP/TP + FN

PPV = TP/TP + FP

For all “false positives” and “false negatives,” we searched for the following cervical diagnoses by ICD-9 code or pathology/Papanicolaou test result: koilocytosis, atypical squamous cells of undetermined significance (ASC-US); atypical squamous cells, cannot exclude high-grade squamous intraepithelial lesion (ASC-H); Papanicolaou smear of cervix with low-grade squamous intraepithelial lesion (LSIL); Papanicolaou smear of cervix with high-grade squamous intraepithelial lesion (HSIL); CIN1; and CIN2. The ICD-9 codes included 795.01, 795.02, 795.03, 795.04, 622.11, and 622.12. We also examined the proportion of “false positives” in the years 2000 and 2005 compared with the proportion in 2001 through 2004 to assess whether the date of the pathology finding might have been outside the time frame of our analysis.

This evaluation was approved by the Institutional Review Boards at CDC and participating sites.

Results

Among the 121,211 women who were in our cohort, there were 1,090 women (1,037 in KPNC and 53 in KPNW) with at least one CIN3þ diagnosis by ICD-9 code and 1,200 women (1,122 in KPNC and 78 in KPNW) with at least one CIN3þ diagnosis by pathology (Table 1). Among the 1,090 women with CIN3þ diagnoses by ICD-9 code, 679 (62%) had more than 1 ICD-9 code for CIN3þ; among

<table>
<thead>
<tr>
<th>Site</th>
<th>ICD-9 Code</th>
<th>Pathology Results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n (%)</td>
<td>n (%)</td>
</tr>
<tr>
<td>KPNC</td>
<td>1,037 (0.96)</td>
<td>1,122 (1.04)</td>
</tr>
<tr>
<td>KPNW</td>
<td>53 (0.40)</td>
<td>78 (0.58)</td>
</tr>
<tr>
<td>Total</td>
<td>1,090 (0.90)</td>
<td>1,200 (0.99)</td>
</tr>
</tbody>
</table>

Abbreviations: CIN3þ, cervical intraepithelial neoplasia grade 3; adenocarcinoma in situ; AIS; KPNC, Kaiser Permanente Northern California; KPNW, Kaiser Permanente Northwest

aThis represents the number of continuously enrolled women at each site with at least one diagnosis of CIN3þ.
bPercentages were calculated using the number of continuously enrolled women as the (13,385 from KPNW and 107,826 from KPNC)
the 1,200 women with diagnosis by pathology, 466 (39%) had more than 1 pathology result for CIN3+ (Fig. 1).

Of the 1,200 women who had at least one pathology result for CIN3+, 740 (694 from KPNC and 46 from KPNW) also had an ICD-9 code for CIN3+. Therefore, using combined data from both sites, the overall sensitivity of the ICD-9 code for CIN3+ pathology diagnosis was 62%. Of the 1,090 women who had an ICD-9 code for CIN3+, 740 (694 from KPNC and 46 from KPNW) also had a pathology diagnosis of CIN3+, resulting in a combined PPV of the ICD-9 code for CIN3+ by pathology of 68%. Three hundred and fifty women had an ICD-9 code for CIN3+ and no laboratory diagnosis (“false positives”; Fig. 2).

Of the 428 women from KPNC with pathology results for CIN3+, but no ICD-9 code for CIN3+ (“false negatives”), 90% had received ICD-9 codes for a different cervical diagnosis. Similarly, of the 32 women at KPNW with pathology results for CIN3+, but no ICD-9 code for CIN3+, 97% had received ICD-9 codes for a different cervical diagnosis. In both sites, the most commonly used alternate cervical diagnoses were ICD-9 codes “795.0” (abnormal Papanicolaou smear of cervix and cervical HPV) and “622.1” histologically confirmed cervical dysplasia.

There were 350 women (343 women from KPNC and 7 women from KPNW) who had an ICD-9 code for CIN3+ but no CIN3+ pathology result during the 5-year period (‘false positives’). The age range of these women was 14 to 29 years (median 22 years) in KPNC and 19 to 25 (median 22 years) in KPNW. Forty percent of these women in KPNC and 57% in KPNW had more than 1 ICD-9 diagnosis for CIN3+. Most had alternate cervical diagnoses in the laboratory record (61% in KPNC and 85% in KPNW). The most common alternative cervical diagnosis was CIN2 and koilocytosis in KPNC and CIN2 and ASC-US in KPNW.

Among the 1,090 women with CIN3+ by ICD-9, the proportion with false positives had some variations by year (44.1% in 2000, 33.2% in 2001, 25.1% in 2002, 38.6% in 2003, 27.6% in 2004, and 17.3% in 2005); overall, 78 women had a diagnosis in either 2000 or 2005 and thus may have had pathology diagnoses before or after inclusion in our cohort.

Discussion

This assessment is the first to systematically compare different methods for CIN3+ case ascertainment from integrated healthcare systems. Our findings suggest that pathology results detected more CIN3+ cases compared with ICD-9 codes. Thirty eight percent of women with CIN3+ during the designated time period would have been missed if only ICD-9 codes were used.

This evaluation also provides other considerations for use of ICD-9 codes for CIN3+ surveillance. ICD-9 codes
were used repeatedly, and most of the time did not reflect unique diagnoses; this is not surprising given that ICD-9 codes typically reflect billing for outpatient visits. Furthermore, 32% of ICD-9 codes for CIN3+ were false positives based on a comparison to pathology results. Because ICD-9 codes reflect provider visits with diagnosis, and not actual pathologic diagnoses, use of these codes could lead to over or under ascertainment.

While pathology results are likely to be a better source of data on CIN3+ than ICD-9 codes, there are challenges in assessing pathology results. Although pathology results in these 2 KP systems were available electronically, many hospital systems may lack this capacity. In addition, the lack of standardized language and changing nomenclature in reporting by pathologists complicates use of these data; in contrast, ICD-9 codes have standardized descriptions, although variability exists in the ways in which these codes are used.

There were limitations to this evaluation. We detected false positives based on discrepancies with laboratory results but could not confirm that these were false positives using a secondary chart review. Also, data were limited to 2 sites and may not be similar to other populations. Our analysis included women aged less than 30 years who were in the cohort during the entire period, and this may not be representative of all women in the integrated health care delivery system. Finally, this assessment focused only on CIN3+ and may not apply to other cervical diagnoses or conditions.

Further research to determine whether similar results are found in other settings and to verify laboratory data for reporting of CIN is needed. Although ICD-9 codes consistently used over time and in large administrative data sets may provide meaningful information, discretion should be taken when using these codes for incidence calculations. Pathology results, when available, may provide useful information for surveillance of outcomes of interest, but standardized methods and reporting are necessary to provide consistent data across laboratories, organizations, and regions.

**Disclosure of Potential Conflicts of Interest**

Diya Surie's work was made possible through the CDC Experience Applied Epidemiology Fellowship, a public/private partnership supported by a grant to the CDC Foundation from External Medical Affairs, Pfizer Inc. A.L. Naleway has a commercial research grant from GlaxoSmithKline. N.P. Klein has a commercial research grant from Merck & Co. and other commercial research support from GlaxoSmithKline. R. Baxter has a commercial research grant from Merck and GSK. No potential conflicts of interest were disclosed by the other authors.

**Authors' Contributions**

**Conception and design:** E.F. Dunne, A.L. Naleway, R. Baxter, J. Gee, L. Markowitz

**Development of methodology:** D. Surie, E.F. Dunne, A.L. Naleway, S. Weinmann

**Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.):** E.F. Dunne, A.L. Naleway, S. Weinmann, N.P. Klein, R. Baxter

**Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis):** D. Surie, E.F. Dunne, A.L. Naleway, K. Hutchins

**Writing, review, and/or revision of the manuscript:** D. Surie, E.F. Dunne, A.L. Naleway, S. Weinmann, N.P. Klein, R. Baxter, J. Gee, L. Markowitz

**Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases):** E.F. Dunne, S. Weinmann, R. Baxter, J. Gee

**Study supervision:** E.F. Dunne, N.P. Klein, L. Markowitz

**Disclaimer**

The findings and conclusions in this article are those of the authors and do not necessarily represent the views of the CDC.

**Grant Support**

This study was funded through a National Vaccine Program Office award received by the Epidemiology and Surveillance Branch, Division of STD Prevention, NCHHSTP, CDC, American Health Insurance Plans (AHP) contract number: 200–2002–00732.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received December 15, 2012; revised February 22, 2013; accepted March 18, 2013; published OnlineFirst April 5, 2013.

**References**

Cervical Intraepithelial Neoplasia Grade 3 and Adenocarcinoma *In Situ*: Comparison of ICD-9 Codes and Pathology Results—Kaiser Permanente, United States, 2000–2005


*Cancer Epidemiol Biomarkers Prev* Published OnlineFirst April 5, 2013.

Updated version Access the most recent version of this article at: doi:10.1158/1055-9965.EPI-12-1413

E-mail alerts Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.