Published in final edited form as:

J Low Genit Tract Dis. 2013 July ; 17(3): 242–247. doi:10.1097/LGT.0b013e3182730fec.

A cervical abnormalities risk prediction model: can we use clinical information to predict which patients with ASCUS/LSIL Pap tests will develop CIN2/3 or AIS?

Brittany M. Charlton, MSc^a, Jenny L. Carwile, MPH^a, Karin B. Michels, ScD, PhD^{a,b}, and Sarah Feldman, MD, MPH^c

^aDepartment of Epidemiology, Harvard School of Public Health, 677 Huntington Ave, 9th Fl, Boston, MA, 02115

^bObstetrics and Gynecology Epidemiology Center, Brigham and Women's Hospital, Department of Obstetrics and Gynecology, 221 Longwood Ave, Boston, MA, 02115

^cDivision of Gynecologic Oncology, Department of Obstetrics and Gynecology, Brigham and Women's Hospital, 75 Francis St, Boston, MA, 02115

Abstract

Objective—HPV infections and abnormal Pap tests are common, and most do not progress to cervical cancer. Since it is difficult to predict which mild Pap abnormalities will develop into precancerous lesions, many women undergo painful and costly evaluations, and even unnecessary treatment. The objective of this study was to develop a risk prediction model based on clinical and demographic information to identify women most likely to develop significant precancerous lesions (CIN2/3 or AIS) among women with mild Pap abnormalities (ASCUS/LSIL).

Materials and Methods—The Abnormal Pap Smear Registry includes women who received treatment at the Brigham and Women's Hospital/Dana Farber Cancer Institute Pap Smear Evaluation Center beginning in 2006. It includes 1,072 women with mild cervical dysplasia (ASCUS or LSIL Pap tests) on their referral Pap test. We derived a clinical prediction model to predict the probability of developing CIN2/3 or AIS using multivariate logistic regression with a split-sample approach.

Results—By the end of follow-up, 93 of the 1,072 women developed CIN2/3 or AIS (8.7%). There were several differences between women who developed CIN2/3 or AIS and women who did not. However, once we put these into the regression model, the only variable that was significantly associated with CIN2/3 or AIS was having a prior history of an abnormal Pap or biopsy [OR=2.44, 95% CI (1.03 to 5.76)]. The resulting prediction model had poor discriminative ability and was poorly calibrated.

Conclusions—Despite accounting for known risk factors, we were unable to predict individual patients' probability for progression on the basis of available data.

Keywords

Vaginal Smears; Uterine Cervical Dysplasia; Colposcopy; Decision Support Techniques

Corresponding author: Brittany M. Charlton, Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, 9th Floor, Boston, MA 02115. Phone: (617) 432-1050, bcharlto@hsph.harvard.edu.

6. Disclosures of interest

None of the authors have any conflicts of interest.

Introduction

Since the introduction of the Pap test in the 1950s, cervical cancer rates have declined by more than 60% in the United States [1]. However, because many preinvasive lesions will never progress to invasive cancer [2, 3], it would be helpful to be able to predict which lesions will progress in order to better target, evaluate, and treat patients at highest risk. Diagnostic tools and treatments such as colposcopies, cold knife conization, loop electrosurgical excision procedure (LEEP), and cervical surgery have been associated with pain, sexual dysfunction, poor obstetric outcomes, reduced quality of life, and are costly to the health care system [4–6].

If clinicians were able to predict the risk of a particular woman with a mildly abnormal Pap test progressing to high-grade dysplasia or cervical cancer, we might thus be able to improve patient experience, expose fewer patients to these procedures, and also reduce health care cost. Although new screening lab tests have been developed, such as human papillomavirus (HPV) co-testing (different from reflex testing), that add additional information about risk [7], these tests are not highly specific and add cost, often increasing the number of patients who undergo further evaluation, rather than reducing that number.

In contrast, risk prediction models for other diseases, such as cardiovascular disease [8–10], have helped clinicians triage care and improve decision making, using only information available in the clinical history. A number of risk factors have been associated with an increased risk for cervical cancer, and may be incorporated into a risk prediction model. These include risk factors that lead to both the development and progression of precancerous lesions. While infection with HPV is the primary, however, possibly transient risk factor, other known predictors of cervical cancer include: age, cigarette smoking, number of sexual intercourse partners, gravidity, number and regularity of Pap tests, weakened immune system (e.g., HIV infection), and long-term use of oral contraceptives [11]. Previously constructed cervical cancer prediction models have used molecular markers and dynamic Bayesian networks of dozens of laboratory variables to predict risk [12]. However, such models are relatively inaccessible to clinicians, and a more parsimonious model based on readily assessable patient characteristics could be more useful clinically.

The purpose of this study was to develop a risk prediction model based on prospectively obtained clinical and demographic information to identify women most likely to develop biopsy confirmed cervical intraepithelial neoplasia grades 2/3 (CIN2/3) or adenocarcinoma in situ (AIS) among women presenting with atypical squamous cells of undetermined significance (ASCUS) or low-grade squamous intraepithelial lesions (LSIL) at their referral Pap test. Such a model could suggest various options, based on a women's risk, including different time intervals between routine Pap tests, retesting, or more invasive treatment options including cold knife conization or LEEP.

Methods

Study Design, Setting, and Participants

The Abnormal Pap Smear Registry includes women who received treatment at the Brigham and Women's Hospital/Dana Farber Cancer Institute Pap Smear Evaluation Center beginning in 2006. Of the 1,890 women who presented to the clinic between 2006 and 2010, 1,299 had mild cervical dysplasia (ASCUS-HPV positive or LSIL Pap tests) on their referral Pap test. The referral Pap test was defined as the most recent Pap test within one year of the first colposcopy appointment at the Pap Smear Evaluation Center. Patients with ASCUS HPV-negative tests were not included in this dataset. Additionally, women presenting at the referral visit with the following cytologic abnormalities were not included in the analysis, as

they already had evidence of more serious abnormalities necessitating evaluation: 154 had high grade dysplasia (HSIL), 64 had atypical squamous cells with possible high grade abnormalities (ASC-H), 14 had atypical glandular cells (AGC), and four had AIS. The outcome measure for the prediction model was: biopsy proven high grade lesions (including both CIN2/3) and adenocarcinoma in situ (AIS). The BWH pathology lab does not report CIN2 and CIN3 separately so these were grouped together as CIN2/CIN3.

The database was created for the purpose of developing evidence-based management and treatment options, and studying clinical risk and demographic factors affecting the care of women with abnormal Pap tests or precancerous lesions of the lower genital tract. Each woman's demographic and reproductive history was obtained using a standardized questionnaire. The woman completed the questionnaire during her first appointment at the Pap Smear Evaluation Center and the physician clarified any questions before the nurse entered the information into the registry. On subsequent visits, the woman had the opportunity to update questionnaire data as necessary. After each physical exam, the physician entered the procedure details including any tests (e.g., HPV, Pap test, or pathology) performed and the nurse recorded the test result after reviewing the formal pathology report. Prior medical records and pathology reports were obtained on all patients and entered by the physician/nursing staff at the time of the first visit. This analysis was restricted to women with complete data for all predictor variables. Women also needed to have a referral Pap test result of ASCUS or LSIL followed by at least one Pap test or biopsy completed at the Pap Smear Evaluation Center between the database creation in 2006 and when analyses began in June of 2010 (N=1,072). This study was approved by the Brigham and Women's Hospital institutional review board and written informed consent was obtained from all research participants at the time of enrollment.

Variables

Demographic information included age (continuous), race (white or nonwhite), and smoking history (ever/never). Reproductive history included the referral Pap test result (ASCUS or LSIL). Gravidity was categorized as 0 or ≥ 1 . The number of lifetime sexual partners was modeled using the following categories: 1–3, 3–5, 7–10, ≥ 10 partners. Because the error in category creation in number of sexual partners (overlap and missing number of six) was not corrected in the first year, data collection for the lifetime sexual partners variable continued in this manner for consistency and was modeled as 1–3 or ≥ 4 . Current oral contraceptive use and history of an abnormal Pap test were also collected through the questionnaire. Additional variables used in the risk prediction model included the following variables extracted from medical records: history of HPV infection (negative/positive/untested or unknown from reflex testing) and history of biopsy proven CIN2/3 or AIS or history of abnormal Paps (negative/positive/untested or unknown). The HPV infection status was taken from the most recent HPV test within one year of the first colposcopy appointment. In all cases, the lab report was obtained for confirmation of HPV test result. Finally, the outcome variable of CIN2/3 or AIS was recorded from pathology results. All slides of patients with CIN2/3 or AIS were either read primarily or reviewed and confirmed by pathologists from Brigham and Women's Hospital.

Statistical Methods

We derived a clinical prediction model to predict the probability of developing CIN2/3 or AIS using multivariate logistic regression. Due to the size of the registry, the model was validated using a split-sample approach, in which one-half of the data were used for model building and the other half of the data were used for model testing [13]. Variables were chosen for inclusion in the model based on a priori hypothesis or univariate association with the outcome of CIN2/3 or AIS, and included: age, history of smoking, current oral

contraceptive use, prior HPV infection, prior abnormal Pap or biopsy, and referral Pap result. Analyses were restricted to women with complete data for all covariates of interest (N=1,072). Receiver-operating curves (ROC) were constructed [14], and the ability of the risk prediction model to discriminate between women who would and would not develop CIN2/3 or AIS was calculated as the Area Under the Curves (AUC) (an AUC of 1 indicates perfect prediction whereas an AUC of 0.5 is the same as chance). The AUC captures the discriminative ability of the model to distinguish between a CIN2/3 or AIS case and a control. We also examined the model's calibration, or how close the predicted probability was to the actual risk [15]. All analyses were conducted with SAS statistical software 9.2 [16].

Results

Women were followed for a median of two visits (range: 1–11). The mean age of the women was 30.5 years (range: 16–66 years), and the cohort was racially heterogeneous with 43.6% (N=467) identifying as white and 56.4% (N=605) identifying as non-white. Thirty-eight percent (N=407) of women reported a history of smoking. In regards to reproductive history, a referral Pap test result of ASCUS HPV positive (N=551) was slightly more common than LSIL (N=521). More than half of the women (N=627) were parous, 65% (N=698) had at least four lifetime sexual partners, and nearly one-third (N=320) currently used oral contraceptives. Additionally, 94.6% (N=1,014) of the women had a previous abnormal Pap test or biopsy result before their referral Pap test at the Pap Smear Evaluation Center. By the end of follow-up, 93 of the 1,072 women who presented with ASCUS or LSIL developed CIN2/3 or AIS (8.7%). Table 1 shows the demographic and reproductive characteristics stratified by the CIN2/3 or AIS outcome and Table 2 shows the results from the risk prediction model including the beta estimates. There were several differences between women who developed CIN2/3 or AIS and women who did not. Women who developed CIN2/3 or AIS were more likely to have smoked cigarettes (46.2% compared to 37.2%), reported having 4 lifetime sexual intercourse partners (76.3% compared to 64.0%), and were more likely to have had a previous abnormal Pap test or biopsy (97.8% compared to 94.3%). However, once we put these into the regression model, the only variable that was significantly associated with CIN2/3 or AIS was having a prior history of abnormal Pap or biopsy [OR=2.44, 95% CI (1.03 to 5.76)].

Next, we quantified the ability of the model to separate cases from controls. The AUC was 0.63 in the training set and 0.53 in the testing set, indicating poor discriminative ability of the model to distinguish between a CIN2/3 or AIS case and a control. Calibration using the Hosmer-Lemeshow goodness-of-fit test indicated low agreement between a woman's actual outcome and her predicted outcome in both the training and testing set. In the training set, the Hosmer-Lemeshow $X^2=5.27$ with a p -value=0.73 and in the testing set $X^2=14.32$ with a p -value=0.05. The null hypothesis is that the risk prediction model is a good model, so it is desirable to have a low X^2 value and a high p -value. Therefore, the model was not able to closely quantify the predicted probability compared to the actual risk.

Discussion

Our risk prediction model developed from data from over 1,000 patients seen at the Brigham and Women's Hospital/Dana Farber Cancer Institute Pap Smear Evaluation Center did not perform well given the available covariates. Building a risk prediction model based on easily obtainable clinical data to distinguish between a woman who presents with a mildly abnormal Pap test and subsequently develops CIN2/3 or AIS and a woman who does not go on to develop CIN2/3 or AIS proved difficult. Nonetheless, as expected, some covariates were associated with developing CIN2/3 or AIS. Women who developed CIN2/3 or AIS

reported a higher number of lifetime sexual intercourse partners, suggesting that a woman is more likely to be exposed to the HPV virus with an increasing number of sexual contacts. Women who developed CIN2/3 or AIS were also more likely to smoke, which supports the role of cigarette smoke in inhibiting the clearance of an HPV infection or possibly even in progressing a precancerous lesion. The odds ratio for smoking among women who progressed to CIN2/3 or AIS is similar to other studies [17]. A history of smoking was associated with a large increase [RR=1.47, 95% CI (0.81 to 2.68)], though this was not statistically significant. Women with a prior history of an abnormal Pap or biopsy were also more likely to subsequently develop CIN2/3 or AIS, even if the current Pap showed ASCUS or LSIL, again confirming the long-term elevation of risk for cervical cancer/dysplasia associated with a history of CIN 2/3 or AIS [18].

HPV infection cofactors can be classified into three groups (1) environmental or exogenous factors, including oral contraceptive use, tobacco use, diet, and sexually transmitted infections; (2) host factors including endogenous hormones, genetic factors, and immune response; and (3) viral factors, including HPV type and variant, viral load, and viral integration [19]. Of these, our risk prediction model included data on only the first two. Risk prediction models using molecular markers such as mRNA and DNA have had better discriminative ability [20], supporting the idea that to arrive at a model with good predictive abilities, researchers must incorporate all three cofactor groups. However, more research is needed to understand the feasibility and cost-effectiveness of such models.

Our study had a number of strengths including the large sample size and data on many predictors of progression to cervical cancer. However, we were limited by the few cases of CIN2/3 or AIS since these are relatively rare outcomes. Nonetheless, the size of our database allowed us to capture 93 cases, giving us adequate power. Additionally, there was the potential for misclassification of risk factors. Women may have misreported the number of lifetime sexual partners. However, due to the prospective design, this misclassification is likely to be nondifferential and therefore not differ between those with the CIN2/3 or AIS and those without. Additionally, we only collected data on current oral contraceptive use, which may have led to some misclassification of long-term use. Women presenting with LSIL or ASCUS Pap tests could have underlying CIN2/3 or AIS so some further misclassification is possible. The origin of CIN2/3 and AIS could differ but excluding the AIS cases in sensitivity analyses did not change the results. We also could not examine CIN2 and CIN3 separately due to the pathology reporting. Finally, our study population was likely similar to that of other teaching hospitals in developed countries, thus our results should be generalizable to women in developed nations. Additionally, there is no reason to believe that biological mechanisms should greatly differ between populations, giving further justification for generalizability.

Current guidelines recommend additional evaluation for women with LSIL and ASCUS Pap tests [21]. Since most of these women will not go on to develop more serious cervical dysplasia, it is desirable to use clinical and demographic factors to further stratify women by risk of progression. In this study, however, consideration of additional risk factors did not improve prediction of CIN2/3 or AIS among women presenting with ASCUS/LSIL. Further research is needed to explore additional parameters to predict cervical dysplasia progression. Such a model could improve a major public health problem using a quick, simple, and inexpensive approach that is both safe and acceptable to the target population.

Acknowledgments

Funding

Ms. Charlton was supported by the Training Program in Cancer Epidemiology under grant T32 CA 09001. Ms. Carwile was supported by the Training Program in Environmental Epidemiology under grant T32 ES 007069 and the Training Program in Reproductive, Perinatal, and Pediatric Epidemiology under grant T32 HD 060454.

We thank all the women who agreed to participate in this study. We also acknowledge the Obstetrics and Gynecology residents as well as nurses at Brigham and Women's Hospital for diligently collecting the data used in this study.

References

1. Borrito, F.; De Ridder, M. HPV and Cervical Cancer Achievements in Prevention and Future Prospects. New York, NY: Springer New York; 2012.
2. Insinga RP, Dasbach EJ, Elbasha EH. Epidemiologic natural history and clinical management of Human Papillomavirus (HPV) Disease: a critical and systematic review of the literature in the development of an HPV dynamic transmission model. BMC Infect Dis. 2009; 9:119. [PubMed: 19640281]
3. Melnikow J, Nuovo J, Willan AR, Chan BK, Howell LP. Natural history of cervical squamous intraepithelial lesions: a meta-analysis. Obstet Gynecol. 1998; 92(4 Pt 2):727–35. [PubMed: 9764690]
4. Crane JM. Pregnancy outcome after loop electrosurgical excision procedure: a systematic review. Obstet Gynecol. 2003; 102(5 Pt 1):1058–62. [PubMed: 14672487]
5. Flanagan SM, Wilson S, Luesley D, Damery SL, Greenfield SM. Adverse outcomes after colposcopy. BMC Womens Health. 2011; 11:2. [PubMed: 21251278]
6. Kyrgiou M, Koliopoulos G, Martin-Hirsch P, Arbyn M, Prendiville W, Paraskeva E. Obstetric outcomes after conservative treatment for intraepithelial or early invasive cervical lesions: systematic review and meta-analysis. Lancet. 2006; 367(9509):489–98. [PubMed: 16473126]
7. Mayrand MH, Duarte-Franco E, Rodrigues I, Walter SD, Hanley J, Ferenczy A, et al. Human papillomavirus DNA versus Papanicolaou screening tests for cervical cancer. N Engl J Med. 2007; 357(16):1579–88. [PubMed: 17942871]
8. Anderson KM, Wilson PW, Odell PM, Kannel WB. An updated coronary risk profile. A statement for health professionals. Circulation. 1991; 83(1):356–62. [PubMed: 1984895]
9. D'Agostino RB Sr, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. Circulation. 2008; 117(6):743–53. [PubMed: 18212285]
10. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. Circulation. 1998; 97(18):1837–47. [PubMed: 9603539]
11. Wang SS, Zuna RE, Wentzensen N, Dunn ST, Sherman ME, Gold MA, et al. Human papillomavirus cofactors by disease progression and human papillomavirus types in the study to understand cervical cancer early endpoints and determinants. Cancer Epidemiol Biomarkers Prev. 2009; 18(1):113–20. [PubMed: 19124488]
12. Austin RM, Onisko A, Druzdzel MJ. The Pittsburgh Cervical Cancer Screening Model: a risk assessment tool. Arch Pathol Lab Med. 2010; 134(5):744–50. [PubMed: 20441506]
13. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med. 2008; 27(2):157–72. discussion 207–12. [PubMed: 17569110]
14. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 1982; 143(1):29–36. [PubMed: 7063747]
15. Hosmer DW, Hjort NL. Goodness-of-fit processes for logistic regression: simulation results. Stat Med. 2002; 21(18):2723–38. [PubMed: 12228887]
16. SAS Statistical Software, Release 9.2. Cary, NC: SAS Institute Inc;
17. Harris TG, Kulasingam SL, Kiviat NB, Mao C, Agoff SN, Feng Q, et al. Cigarette smoking, oncogenic human papillomavirus, Ki-67 antigen, and cervical intraepithelial neoplasia. Am J Epidemiol. 2004; 159(9):834–42. [PubMed: 15105176]

18. Ostor AG. Natural history of cervical intraepithelial neoplasia: a critical review. *Int J Gynecol Pathol.* 1993; 12(2):186–92. [PubMed: 8463044]
19. Castellsague X, Munoz N. Chapter 3: Cofactors in human papillomavirus carcinogenesis--role of parity, oral contraceptives, and tobacco smoking. *J Natl Cancer Inst Monogr.* 2003; (31):20–8. [PubMed: 12807941]
20. Scheurer ME, Guillaud M, Tortolero-Luna G, Follen M, Adler-Storthz K. Epidemiologic modeling of cervical dysplasia with molecular and cytopathological markers. *Gynecol Oncol.* 2007; 107(1 Suppl 1):S163–9. [PubMed: 17727928]
21. Saslow D, Solomon D, Lawson HW, Killackey M, Kulasingam SL, Cain J, et al. American Cancer Society, American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology Screening Guidelines for the Prevention and Early Detection of Cervical Cancer. *J Low Genit Tract Dis.* 2012 [Epub ahead of print].

Table 1

Baseline characteristics by CIN2/3 or AIS outcome* (N=1,072)

	No CIN 2/3 or AIS (n=979)	CIN 2/3 or AIS (n=93)
<i>Demographic</i>	% (n)	
Age, years [(mean), range: 16–66]	30.6	29.7
Race		
White	43.2 (423)	47.3 (44)
Nonwhite	56.8 (556)	52.7 (49)
Smoking (ever)	37.2 (364)	46.2 (43)
<i>Reproductive History</i>		
Referral Pap Result		
ASCUS	51.1 (500)	54.8 (51)
LSIL	48.9 (479)	45.2 (42)
Gravidity		
0	41.8 (409)	38.7 (36)
1	58.2 (570)	61.3 (57)
Lifetime number of sexual partners		
1–3	36.0 (352)	23.7 (22)
4	64.0 (627)	76.3 (71)
Current oral contraceptive use	29.4 (288)	34.4 (32)
<u>Previous abnormal Pap test or biopsy</u>	94.3 (924)	97.8 (90)

* Denotes histology obtained on biopsy

Table 2

Model for predicting progression to CIN2/3 or AIS *

	β	OR (95% CI)
Age (years)	-0.002	1.00 (0.97 to 1.03)
History of smoking	0.39	1.47 (0.81 to 2.68)
Current oral contraceptive use	0.41	1.51 (0.79 to 2.89)
Prior HPV infection [†]		
Negative	ref	ref
Positive	1.12	3.06 (0.38 to 24.78)
Untested/unknown	0.69	2.00 (0.23 to 17.57)
Prior abnormal Pap test or biopsy		
Negative	ref	ref
Positive	0.89	2.44 (1.03 to 5.76)
Untested/unknown	-0.08	0.92 (0.46 to 1.87)
Referral Pap test [‡]		
LSIL	ref	ref
ASCUS	-0.28	0.76 (0.25 to 2.27)

Abbreviations: CI, confidence interval; OR, Odds Ratio

* Denotes histology obtained on biopsy

[†] Most recent HPV result within one year of colposcopy appointment[‡] Most recent Pap test within one year of colposcopy appointment