

Prospective evaluation of a risk scoring system for cervical neoplasia in primary care

CLARE E WILKINSON

TIM J PETERS

NIGEL C H STOTT

IAN M HARVEY

SUMMARY

Background. Against a background of concern over the costs of the cervical screening programme in the United Kingdom, increased precision in targeting groups at high risk of having an abnormal cervical smear offers a means of increasing efficiency. Previous papers have described the development of a risk scoring system and its feasibility and reliability in primary care.

Aim. A study was carried out to assess the validity of the scoring system by testing its predictive ability on a prospective data set.

Method. Consecutive attenders for cervical smear tests at seven practices and three clinics were recruited for the study. The women completed a questionnaire from which their risk scores could be calculated. The scores were compared with cytology and histology results. Various performance statistics were obtained.

Results. In terms of cervical intraepithelial neoplasia (CIN) 2 or 3, there was an 11 fold increased risk among the low risk group (scores of four or five) compared with the very low risk group (scores of three or less). The system enabled the identification of 75% (95% confidence interval 62% to 84%) of cases of CIN 2 or 3 among the 21% of the 3629 women with known histology who had a score of four or five.

Conclusion. Given the ease with which risk status can be ascertained (a risk score could not be calculated for only 23 of 3661 women) and the magnitude of difference in risk, the risk scoring system appears to have potential for assisting the targeting of screening resources. Studies of risk perception and behaviour, and ultimately a randomized controlled trial, are required to assess the effectiveness and cost effectiveness of risk targeting.

Keywords: cervical cancer; at risk groups; risk assessment; cervical screening.

Introduction

IN the current climate where suggestions are being sought to reduce the costs of the cervical screening programme in the United Kingdom, increased precision in targeting groups at high

risk of developing an abnormal cervical smear offers a means of increasing efficiency.¹ Case-control studies have demonstrated substantial variation in risk status among individual women.²⁻⁶ Moreover, although the programme is equally targeted towards all sexually active women aged 20 to 65 years with a recommended five year interval between tests,⁷ there is evidence that cervical screening in the UK remains unevenly distributed, which may make health gain targets difficult to achieve.^{8,9}

Previous evaluation of risk-related cervical screening was concerned with selective screening, that is, the exclusive screening of high risk women; not surprisingly, the concept was dismissed because of the penalty in terms of missed cases.^{10,11}

Two previous papers have in turn described the current dilemmas of the national cervical screening programme and the development of a risk scoring system for cervical neoplasia.^{1,12} This scoring system was designed for simplicity so as to have the potential for use by women themselves, and the four items included are independent risk factors, that is, risk information is not being duplicated within the scale. In the risk scoring system, women educated to A-level or higher scored zero; women educated to a different level scored one. Non-smokers scored zero and smokers scored one. Women who had used oral contraceptives for fewer than five years scored zero; women who had used them for five years or more scored one. Women who had had only one sexual partner scored zero; women who had had two partners scored one; and women who had had three or more partners scored two. Scores could therefore be between zero and five.

A pilot interview survey in primary care was conducted to assess the feasibility and test-retest reliability of the tool.¹³ Women found the questionnaire (from which the scoring system could be derived) easy to complete and were not uncomfortable answering sensitive questions in this setting. There was excellent test-retest reliability.

Dividing the pilot sample into two risk categories (scores of four and five being at low risk and scores of zero to three at very low risk) led to about 20% of women being in the first of these groups. These labels were chosen since, regardless of the explicit outcome intended in any application of the risk scale, women may link their risk status to invasive cervical cancer, which in the UK is rare.¹⁴ In addition, there is evidence that women in the UK are particularly anxious about the disease and therefore frightening labels are likely to be falsely interpreted.¹⁵⁻¹⁷

Having constructed the scoring system and demonstrated the feasibility of its use in a primary care setting, the next stage was to demonstrate the validity of the tool by testing its predictive ability on a new data set. A prospective study was therefore undertaken of the current risk status and subsequent smear result of a sample of women attending for cervical cytology screening.

Method

The sample consisted of consecutive attenders for cervical smear tests in seven practices, two district health authority clinics and the genitourinary medicine clinic in Cardiff. The relative proportion of the types of sites recruited into the study was designed to reflect the overall pattern of cervical cytology carried out in the area, and to represent a wide spread across the social classes. When considering social class, the occupation of the partner was

C E Wilkinson, MRCP and N C H Stott, FRCP, Department of General Practice, University of Wales College of Medicine, Cardiff. T J Peters, PhD and I M Harvey, MFFHM, Health Care Evaluation Unit, Department of Epidemiology and Public Health Medicine, University of Bristol. Submitted: 7 April 1993; accepted: 30 August 1993.

© British Journal of General Practice, 1994, 44, 341-344.

used where this was stated; if not the woman's own occupation was used.¹⁸

The sample size was chosen to achieve reasonably precise confidence intervals for a projection of acceptable sensitivity in the detection of a dyskaryotic smear result. Specifically, if the sensitivity obtained were to be 70%, the aim was to estimate this with at most a 10% margin of error, that is, to obtain a 95% confidence interval for true sensitivity of 60% to 80%. Using exact confidence intervals based on the binomial distribution¹⁹ this required approximately 80 dyskaryotic smear results in total. Dyskaryosis of any degree appears to have a higher rate in South Glamorgan (about 3%) than the UK average.²⁰ For this calculation, however, 2% was used as a conservative estimate. Hence the total cohort size required for this study was around 4000 women. As a corollary, if as in the pilot study, the proportion of women in the low risk band was about 20%, then given this sample size the positive predictive value would be estimated as 7% with a margin of error of 2%. This was also felt to be acceptable. Given a target of 4000 patients the projected length of study for the number of sites envisaged was approximately 12 months.

Data collection and management

In each site, the practice nurse or clinic nursing sister who carried out the majority of the cervical smear tests took a lead role in the data collection. Women were invited to participate in the study during their consultation for a cervical smear test. A self-report form, described previously,¹³ contained questions designed to gather, among other information, the risk factors chosen for the scoring system and this was completed by each woman. Confidentiality was preserved by using numbers instead of names on the self-report forms, which were sent to the laboratory alongside, but separated from, the cytology form. The risk score was not calculated explicitly until the data were analysed for the purposes of this study.

Data collection commenced in December 1989 and was completed in January 1991. The denominator for the sample was determined from the manual records which were kept of all tests processed through the South Glamorgan laboratory.

Histology results were collected for those women who had an abnormal smear result followed by colposcopy examination. All abnormal smear tests in the study were re-read by a consultant cytopathologist to identify any false positives in the sample; none was identified.

Outcome measures

Three outcomes were used: any dyskaryosis, moderate or severe dyskaryosis and cervical intraepithelial neoplasia (CIN) grades 2 or 3. These are the most influential in terms of broadly accepted patient management: all women with dyskaryosis will at the least receive increased surveillance; all women with moderate or severe dyskaryosis proceed to an outpatient appointment for colposcopy; all patients with CIN 2 or 3 are likely to receive operative treatment.²¹

Statistics

Since the aim was to assess the predictive value of the scoring system, sensitivity, specificity and predictive values were calculated. Although the use of a single (pre-specified) cut-off point for the risk score would accord with its probable practical application, a more complete picture of performance could be depicted by using a receiver operating characteristic chart, that is, a plot of sensitivity against 1-specificity across the possible cut-off points on the risk score.²² The charts were used to compare, for the three different outcomes, the performance statistics of the scor-

ing system across the full range of the risk scale.

The positive and negative predictive values for the specified cut-off point (below four; four and above) were important for assessing the practical value of the scoring system. These are presented in the form of pre- and post-test probabilities, that is, the proportion with the outcome, first for the entire sample (the prevalence) and then separately for the two risk groups (positive predictive value and 1-negative predictive value). The comparison of the two post-test probabilities is presented in terms of their ratio (risk ratio) and their difference (attributable risk).

A further question relating to the performance of the risk scale was how this varied across subgroups of women according to various sociodemographic characteristics. In part this is an issue of its generalizability, but it is also apposite to the wider question of whether the many items collected but not part of the previously proposed scoring system might have confounded or interacted with the risk score. Therefore age, social class and practice type were considered by both stratified analysis and by multiple logistic regression.²²

Data analysis was performed using the *SPSSPC+* version 3.0 for basic frequencies and cross-tabulations, the *EGRET* package for multiple logistic regression and the *CIA* package for confidence intervals.²³

Results

According to the manual records held for the South Glamorgan cervical screening programme, 4912 cervical smear tests were performed at the study sites during the data collection period. Of these, 314 had been repeat smears owing to inadequate initial smears. Of the 4598 women tested, 319 were already under surveillance owing to previous dyskaryosis. Therefore, 4279 women were eligible for inclusion in the study. Fifty four women chose not to take part, 196 study forms were not completed, and in 368 cases the study protocol was not followed. Thus 3661 women (85.6%) had data available for analysis. The participation rate varied across the sites from 72% in one of the practices to 97% in the two clinics. The percentage contributions of the 10 sites to the total number of women in the study varied from 5% to 17%.

The age distribution of the sample was positively skewed (mean age 36 years, median 33 years, range 14 to 72 years). Compared with the 1981 census data for England and Wales, women in social class 1 were over-represented (10%), and women in social classes 3M (15%), 4 (6%) and 5 (4%) were under-represented.²⁴

Risk scores could not be calculated for 23 women, 10 because they reported being virgins at the time of attending for the smear test. Among the 3638 with a risk score, 280 women scored zero (7.7%), 868 scored one (23.9%), 904 scored two (24.8%), 801 scored three (22.0%), 593 scored four (16.3%), and 192 scored five (5.3%). Therefore, 785 women (21.6%) were in the low risk category and 2853 (78.4%) were in the very low risk category.

The cytology and histology results are shown in Table 1. Women who initially had inadequate smears were given a repeat test and then classified according to the result of the repeat smear. The cytology figures from South Glamorgan Health Authority for 1990 show that 94.7% of smears were normal, 2.2% were borderline and 3.1% were dyskaryotic, compared with 93.2%, 3.0% and 3.9%, respectively, in this study.

The relationship between risk score (scores of four or five being low risk and scores of zero to three being very low risk) and cytology and histology (any dyskaryosis, moderate or severe dyskaryosis and CIN 2 or 3) is shown in Table 2. The performance statistics for the risk scale using the cut-off points for the

Table 1. Cytology results for the study women, and histology results for the study women with a dyskaryotic smear result.

| Result | No. (%) of women with result |
|--|------------------------------|
| <i>Cytology (n = 3661)</i> | |
| Normal ^a | 3411 (93.2) |
| Borderline | 109 (3.0) |
| Dyskaryosis: | |
| Mild | 86 (2.3) |
| Moderate | 30 (0.8) |
| Severe | 25 (0.7) |
| <i>Histology (n = 114)^b</i> | |
| Normal | 32 (28.1) |
| CIN: | |
| 1 | 15 (13.2) |
| 2 | 18 (15.8) |
| 3 | 47 (41.2) |
| Invasive cervical cancer | 2 (1.8) |

n = number of women in group. ^aIncludes one case of glandular neoplasia. ^bExcludes 18 cases where biopsy was not performed and nine cases for whom biopsy result was unavailable.

three outcome measures, and the risk ratios and attributable risks, are shown in Table 3. In particular, these indicate that 75% of the cases of CIN 2 or 3 were identified among just 21.4% of the total sample with known histology who had a risk score of four or five; equivalently, that in this 21.4% the proportion with the outcome CIN 2 or 3 was 64 per 1000, compared with six per 1000 among the remaining 78.6% of the sample.

The receiver operating characteristic charts comparing the performance of the risk scale across all possible cut-off points for each of the three outcome measures are shown in Figure 1. A progressively increasing shift towards the top left corner is apparent for moderate or severe dyskaryosis and CIN 2 or 3 in turn compared with any dyskaryosis. This shows the improved performance of the scoring system as the outcome becomes more focused on the higher grade cervical lesions.

Investigation of the generalizability of the scoring system across different ages, socioeconomic groups and primary care settings revealed that none of these factors altered the performance statistics of the risk scale for each of the three outcome measures. That is, they did not confound the risk score. As regards the factors themselves, there was a statistically significant linear decline with age in the risk of 'any dyskaryosis', and increased risks of the two more serious outcomes among 30–49 year olds. All of these patterns were reduced in magnitude and, for the more serious outcomes, were of borderline or no statistical significance after controlling for the risk score (not surprisingly since age is implicit within some components of the risk score). However, there was an element of systematic variation in risks across socioeconomic groups which was not wholly accounted for by the risk score.

Although there was a suggestion that the performance of the risk scale was enhanced among older women, this was compens-

Table 2. Cytology and histology results according to risk status.

| Result | Number of women | | |
|---------------------------------------|-----------------|----------|-------|
| | Risk status | | Total |
| | Low | Very Low | |
| <i>Any dyskaryosis</i> | | | |
| Yes | 83 | 58 | 141 |
| No | 702 | 2795 | 3497 |
| Total | 785 | 2853 | 3638 |
| <i>Moderate or severe dyskaryosis</i> | | | |
| Yes | 40 | 15 | 55 |
| No | 745 | 2838 | 3583 |
| Total | 785 | 2853 | 3638 |
| <i>CIN 2 or 3</i> | | | |
| Yes | 50 | 17 | 67 |
| No | 728 | 2834 | 3562 |
| Total | 778 | 2851 | 3629 |

CIN = cervical intraepithelial neoplasia.

ated for by the declining absolute risks. There was no evidence of differential performance across either social class groups or type of site.

Discussion

The results from this prospective survey are encouraging, with an 11 fold change in risk of CIN 2 or 3 between the women in the low and very low risk groups. Given the ease with which risk status could be ascertained, and the magnitude of difference in risk for the outcome of CIN 2 or 3, the scale appears to have potential for the targeting of screening resources.

The performance statistics of the scoring system were better for the more serious outcomes of moderate or severe dyskaryosis and CIN 2 or 3 than for the wider outcome of any dyskaryosis. This may be because minor degrees of dyskaryosis and dysplasia are less easy to define than the more severe end of the disease spectrum.

Although younger women and those in higher social classes were over-represented in this sample, this is characteristic of a sample of women recruited while attending for cervical screening.⁸ It is reassuring that the scoring system appeared to be generalizable across ages, social classes and participating sites, allowing application across a wide variety of primary care settings.

Nevertheless further research is needed to answer fully the question of whether or not the performance statistics are sufficient to achieve the ultimate aim of improving the cost-effectiveness of the screening programme. The acceptability of the levels of sensitivity and positive predictive value will be determined by the precise change in management accruing from identification of risk status. The particular issues which would need to be

Table 3. Performance statistics for the three outcome measures.

| Outcome | Sensitivity (%) (95% CI) | Specificity (%) (95% CI) | Prevalence per 1000 | Positive predictive value per 1000 (95% CI) | 1- negative predictive value per 1000 (95% CI) | Risk ratio | Attributable risk (per 1000) |
|--------------------------------|--------------------------|--------------------------|---------------------|---|--|------------|------------------------------|
| Any dyskaryosis | 59 (50 to 67) | 80 (78 to 81) | 39 | 106 (80 to 120) | 20 (15 to 26) | 5.2 | 85.4 |
| Moderate or severe dyskaryosis | 73 (59 to 83) | 79 (78 to 80) | 15 | 51 (36 to 68) | 5 (3 to 9) | 9.7 | 45.7 |
| CIN 2 or 3 | 75 (62 to 84) | 80 (78 to 80) | 18 | 64 (48 to 83) | 6 (3 to 10) | 10.8 | 58.3 |

CIN = cervical intraepithelial neoplasia. CI = confidence interval.

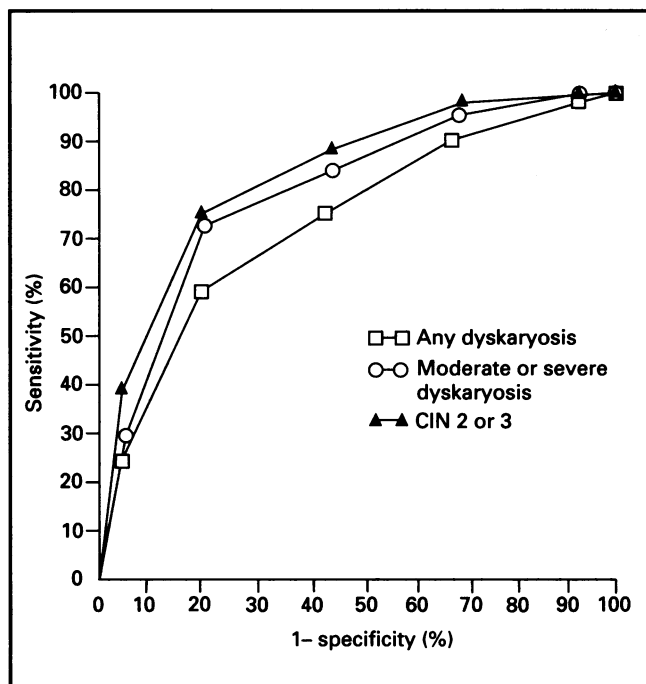


Figure 1. Receiver operating characteristic charts for the three outcomes across all possible cut-off points on the risk scale.

addressed include the cost implications of alternative management strategies and the potentially differing perspectives of the women, primary care professionals and those involved in the management of the programme.

In South Glamorgan there is evidence that up to 40% of women are rescreened at less than a three-year interval,²⁵ and this is unlikely to be unique in the UK. In these instances, the scoring system might be used as an educational intervention to demonstrate the very low risk status of the majority of women and hence may encourage more appropriate screening behaviour. In turn, this would enable resources to be targeted more towards those women in the higher of the two risk categories.

Screening intervals appropriate to risk status could improve the cost efficiency of the programme, presuming that such recommendations are reflected in actual screening behaviour and that the logistic problems of implementation are overcome. This could be attempted in a number of ways: influencing women's behaviour directly, influencing those responsible for taking the smears, altering the call-recall process either at the family health services authority or laboratory level, or a combination of these (C Wilkinson, unpublished MD thesis). Any of these approaches would be complementary to the suggestion that screening be ceased beyond the age of 50 years for women with a negative screening history.²⁶

It would be important to ensure that women who fell into the higher risk category were not unduly frightened by the scoring system, since their real risk of invasive cervical cancer would still be very small since the incidence of the disease in the UK is low.²⁴

Before the scoring system is put into practice, further research is essential to investigate women's individual perception of their risk status for cervical abnormalities and their reaction to the scoring system. The design of a prototype scoring tool is currently underway to allow further field testing of this simple system, following which a practice-based randomized controlled trial should be conducted to assess the impact and efficiency of risk-sensitive screening.

References

1. Wilkinson CE, Peters TJ, Harvey IM, Stott NCH. Risk targeting in cervical screening: a new look at an old problem. *Br J Gen Pract* 1992; **42**: 435-438.
2. Harris RWC, Brinton LA, Cowdell RH, *et al*. Characteristics of women with dysplasia or carcinoma in situ of the cervix uteri. *Br J Cancer* 1980; **42**: 359-369.
3. Clarke EA, Morgan RW, Newman AM. Smoking as a risk factor in cancer of the cervix: additional evidence from a case control study. *Am J Epidemiol* 1982; **115**: 59-65.
4. Clarke EA, Hatcher J, McKeown-Eyssen GE, Lickrish GM. Cervical dysplasia: association with sexual behaviour, smoking and oral contraceptive use? *Am J Obstet Gynecol* 1985; **151**: 612-616.
5. La Vecchia C, Franceschi S, Decarli A, *et al*. Sexual factors, venereal diseases and the risk of intraepithelial and invasive cervical neoplasia. *Cancer* 1986; **58**: 935-941.
6. Brinton LA, Hamman RF, Huggins GR, *et al*. Sexual and reproductive risk factors for invasive squamous cell cervical cancer. *J Natl Cancer Inst* 1987; **79**: 23-30.
7. Department of Health and Social Security Health Services Department. *Cervical cancer screening health circular. HC(88) 1*. London: HMSO, 1988.
8. Austoker J, McPherson A. *Cervical screening*. Oxford University Press, 1992.
9. Secretary of State for Health. *The health of the nation: a strategy for health in England (Cm 1986)*. London: HMSO, 1992.
10. Hakama M, Pukkala E. Selective screening for cervical cancer. Experience of the Finnish mass screening system. *Br J Prev Soc Med* 1977; **31**: 238-244.
11. Hakama M, Pukkala E, Saastamoinen P. Selective screening: theory and practice based on high-risk groups of cervical cancer. *J Epidemiol Community Health* 1979; **33**: 257-261.
12. Wilkinson C. Abnormal cervical smear test results: old dilemmas and new directions. *Br J Gen Pract* 1992; **42**: 336-339.
13. Wilkinson CE, Peters TJ, Harvey IM, Stott NCH. Feasibility, reliability and women's views of a risk scoring system for cervical neoplasia in primary care. *Br J Gen Pract* 1994; **44**: 306-308.
14. Cancer Research Campaign. *Annual report 1989/handbook 1990*. London: CRC, 1990.
15. Posner T, Vessey M. *Prevention of cervical cancer: the patients' view*. London: King Edward's Hospital Fund, 1988.
16. Wilkinson C, Jones JM, McBride J. Anxiety caused by abnormal result of cervical smear test: a controlled trial. *BMJ* 1990; **300**: 440.
17. Lee JM. Screening and informed consent (sounding board). *N Engl J Med* 1993; **328**: 438-440.
18. Office of Population Censuses and Surveys. *Standard occupational classification. Volume 3*. London: HMSO, 1991.
19. Lentner C (ed) *Geigy scientific tables Volume 2*. 8th edition. Basle, Switzerland: Ciba-Geigy, 1982.
20. British Cooperative Clinical Group. Cervical cytology screening in sexually transmitted disease clinics in the United Kingdom. *Genitourin Med* 1987; **63**: 40-43.
21. National Coordinating Network of the National Health Service Cervical Screening Programme. Duncan ID (ed). *Guidelines for clinical practice and programme management*. Oxford: National Coordinating Network, 1992.
22. Altman DG. *Practical statistics for medical research*. London: Chapman and Hall, 1991.
23. Gardner MJ, Altman DG. *Statistics with confidence: confidence intervals and statistical guidelines*. London: British Medical Journal, 1989.
24. Government Statistical Service. *1981 census: key statistics for local authorities*. London: HMSO, 1984.
25. Jones J. *South Glamorgan cervical cytology programme newsletter*. Cardiff: South Glamorgan Health Authority, 1992.
26. Van Wijngaarden WJ, Duncan ID. Rationale for stopping cervical screening in women over 50. *BMJ* 1993; **306**: 967-971.

Acknowledgements

We are grateful to the Welsh Scheme for Health and Social Research and to the Scientific Foundation Board of the RCGP for funding this project. We thank Jane Jones, Nick Dallimore, Alma Barter, Penny Moore, the women of South Glamorgan and the primary care teams, genitourinary medicine clinic and health authority clinic staff involved in the project.

Address for correspondence

Dr C E Wilkinson, Department of General Practice, University of Wales College of Medicine, Llanedeyrn Health Centre, Maelfa, Cardiff CF3 7PN.