Validity of Melanoma Diagnosis in a Community-based Screening Program

Lin Fritschi¹, Sarah A. Dye², and Paul Katris³

1 Queensland Cancer Fund, Viertel Centre for Research in Cancer Control, Queensland, Australia.
2 School of Population Health, University of Western Australia, Perth, Western Australia, Australia.
3 Western Australian Clinical Oncology Group, The Cancer Council of Western Australia, Perth, Western Australia, Australia.

Received for publication July 26, 2005; accepted for publication February 17, 2006.

Although screening for melanoma is intuitively attractive, evidence of the effectiveness of screening programs for skin cancer is lacking. Since 1990, the Lions Cancer Institute has conducted clinics in Western Australia in which volunteer plastic surgeons and dermatologists undertake full-body skin screens. Advertisements for attendees target people with risk factors for skin cancer. Each person screened between 1994 and 2002 (n = 7,436) completed a questionnaire including basic demographic information, on which the physician added provisional diagnoses. Attendees’ details were linked with the Western Australian Cancer Registry to determine the number of diagnosed melanomas up to 1 and 2 years after screening. The positive predictive value of a screening diagnosis of “any lesion” at a particular body site was 1.5% and that of a screening diagnosis of “melanoma” was 10.0%. The 1-year specificity of the screening test ranged from 95.1% to 99.5%, and 1-year sensitivity ranged from 63.6% to 81.8%. Two-year sensitivity was lower. If body site was not taken into account, the sensitivities were higher and the specificities lower. Findings suggest that the validity of skin screening diagnoses in the general population is reasonable. Body site of the lesion should be taken into account when calculating validity of these diagnoses.

mass screening; melanoma; reproducibility of results; sensitivity and specificity

Abbreviation: PPV, positive predictive value.

Melanoma is one of the most common cancers in Australia, and incidence is increasing in fair-skinned populations around the world (1). Screening for melanoma using visual inspection of skin is intuitively attractive because survival benefit is considerable if melanoma is detected at an early stage. However, there have been no randomized controlled trials of skin screening to our knowledge, so most authorities do not recommend that screening programs be established at this stage (2, 3). 

In Western Australia, the Lions Cancer Institute has been conducting skin screening clinics since 1990 (4). At these clinics, volunteer plastic surgeons and dermatologists conduct full-body examinations of community members. Those with lesions are referred to their own physician for definitive diagnosis and treatment. To assess the validity of the screening diagnoses in a community skin screening clinic, we determined which of the persons screened by the Lions Cancer Institute had had a melanoma diagnosed after their screen by matching lists of names with the Western Australian Cancer Registry.

Several issues need to be addressed in assessing the validity of skin screening diagnoses. First, the appropriate denominator of the validity calculations must be determined. To establish the validity of a screening test, a simple two-by-two table is usually constructed. The screening test can be either negative or positive, as can the “gold standard” test. For cervical cancer, for example, this process is relatively simple: Was the Papanicolaou smear negative or positive?
Was the definitive histologic diagnosis cancer or not? It is usually reasonably clear that the Papanicolaou smear and the histology are from the same area.

However, this process is not so clear with skin cancer because of the size of the area being screened. In the simplest instance, a screen detects a potential melanoma on the arm, the lesion is excised, and the pathologist determines whether the lesion is a melanoma. It is very clear where in the two-by-two table data on this person belong. But what if the screen detects a potential melanoma on the arm but the confirmed melanoma is on the back?

Another approach is to consider that each person undergoes a number of separate screening events and that each can be either negative or positive. Confirmed lesions can then be mapped to those screening events. Because the International Classification of Diseases, Ninth Revision, coding system classifies melanomas as occurring on five body sites (face, scalp and neck, upper limb, trunk, lower limb), we used these body site categories in our analyses. A goal of this project was to contrast the validity of a skin cancer screening using this site-specific approach to methods based on the individual person.

MATERIALS AND METHODS

The Lions Cancer Institute has been conducting skin screening clinics since 1990. A mobile clinic, staffed by volunteers, was set up on weekends in rural and suburban locations around Western Australia. Before each clinic opened, advertisements were run in local newspapers providing a telephone number to call for appointments. Since 1996, the advertisements have listed eight risk factors for skin cancer, and people with three or more have been preferentially targeted (4). These risk factors are 1) family history of melanoma; 2) five or more moles on the forearms; 3) previous removal of benign nevi; 4) previous skin cancer; 5) a lesion that is changing in size, color, or shape; 6) a lesion that does not heal; 7) fair skin that burns rather than tans; and 8) episodes of severe sunburn as child.

At the clinic, each participant completed a questionnaire including demographic details, personal and family history of melanoma, phenotypic characteristics (such as eye and hair color and skin type), and the risk factors. They also completed a consent form allowing future matching of their details with the Western Australian Cancer Registry.

Volunteer dermatologists and plastic surgeons conducted full-body screens on the participants and recorded their provisional diagnoses of any lesions. Included were the type of lesion and the body site (which was marked on a diagram of the body). For this analysis, we divided the types of lesions into three categories: “any lesion,” “malignant lesion,” and “melanoma” (5). The melanoma category includes lesions given the provisional diagnosis of malignant melanoma and also amelanotic melanoma, Hutchinson’s melanotic freckles, and superficial spreading melanoma.

The Western Australian Cancer Registry was provided with an identified list of consenting participants who attended the screening clinics so this information could be matched with registry records. Reporting of all invasive melanomas in Western Australia is mandatory; in addition, data are collected from the register of all hospital admissions in the state, which are cross-matched with the other state cancer registries. Therefore, this list is considered reasonably complete. Only invasive melanomas were included because data on in situ melanomas had not been collected for the entire time period of our study.

Screens conducted between April 16, 1994, and December 31, 2002, were included in the study because participants’ dates of birth were collected after 1993 only, and we wanted to include 2 years of follow-up. Only that information for residents of Western Australia older than age 20 years and for whom complete date of birth data were available was used in these analyses. For one person, two confirmed melanomas of the trunk were excised on the same day. Only one (the thicker one) was considered in these analyses.

For the first analysis, a positive screen was classified as one in which a melanoma had been provisionally diagnosed anywhere on the body at the screening clinic. The gold standard was whether a melanoma anywhere on the body had been reported to the Western Australian Cancer Registry within 2 years of the screening event. Data for each person were entered only once regardless of the number of provisional diagnoses or confirmed melanomas.

Second, each screening attendance was considered five separate screening events: for the face, scalp and neck, trunk, upper limb, and lower limb. Each of the five screening events could be either positive or negative. Thus, if the specialist had marked one possible melanoma occurring on the arm, the screen for the upper limb was considered positive, while those for the face, scalp and neck, trunk, and lower limb were considered negative. The gold standard was whether a melanoma at that body site had been reported to the Western Australian Cancer Registry within 2 years.

All analyses were repeated by using only those melanomas diagnosed within 1 year of the screening event.

We analyzed the data by using the Statistical Package for Social Sciences (6). Measures of validity, including sensitivity, specificity, positive predictive value (PPV), and negative predictive value, were calculated for each of the two scenarios. Exact 95 percent confidence intervals around the observed proportions were determined by the method of Clopper and Pearson (7) as implemented in the Microsoft Excel macros by John C. Pezzullo (Georgetown University, Washington, DC; available at http://statpages.org/confint.html).

Because five measurements were taken from each person in the study, there was the potential for a clustering effect and resulting increased variation. A series of six logistic regression models, comprising combinations of the three screening variables (dependent) by two gold standard variables (independent), were fitted by using SUDAAN (8), a statistical package that adjusts for the cluster study design. The design effect of the odds ratio parameter was examined for each model and was found to be between 1.00 and 1.02, meaning that the increased variation associated with the clustering was about 1–2 percent. Therefore, no further adjustment for the clustering was made in this analysis.
RESULTS

A total of 7,436 persons met the criteria for inclusion in this study (of a total of 9,808 ever screened by the Lions Cancer Institute Skin Cancer Screening Program). There were more females than males, and more than half were older than 50 years of age (table 1). Two thirds of the participants screened lived in areas other than the capital city, in line with the locations of the screening clinics. Most subjects had three or more self-assessed risk factors for skin cancer. The screening detected 201 persons with lesions suspected of being melanoma (2.7 percent of those who were screened); 10 of them had two suspected melanomas.

Thirty-three melanomas were diagnosed within 1 year of a screen, and a further 16 were diagnosed 1–2 years after a screen. Table 2 outlines the validity of the screening diagnoses in the instance when data for each person were entered into the analysis only once. A true-positive screen was classified as one in which a melanoma had been provisionally diagnosed anywhere on the body at the screening clinic and a melanoma anywhere on the body had been reported to the Western Australian Cancer Registry within 1 year (or 2 years) of the screening event. For melanomas diagnosed within 1 year of a screen, the sensitivity of a screening diagnosis of “any lesion” was high (93.9 percent) but at the cost of a large number of false positives and a low PPV (2.2 percent). The PPV of a screening diagnosis of “melanoma” was much higher (11.4 percent), as was the specificity (97.6 percent), but with a concomitant decrease in sensitivity (69.7 percent) and the consequent omission of 10 of the 33 melanomas. The validity of a screening diag-

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>No. of screens</th>
<th>% of total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>3,270</td>
<td>56.0</td>
</tr>
<tr>
<td>Female</td>
<td>4,163</td>
<td>56.0</td>
</tr>
<tr>
<td>Age group (years)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20–39</td>
<td>1,948</td>
<td>26.2</td>
</tr>
<tr>
<td>40–49</td>
<td>1,728</td>
<td>23.2</td>
</tr>
<tr>
<td>50–59</td>
<td>1,709</td>
<td>23.0</td>
</tr>
<tr>
<td>≥60</td>
<td>2,051</td>
<td>27.6</td>
</tr>
<tr>
<td>Region</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metropolitan</td>
<td>2,397</td>
<td>32.2</td>
</tr>
<tr>
<td>Rural</td>
<td>5,039</td>
<td>67.8</td>
</tr>
<tr>
<td>No. of risk factors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–2</td>
<td>2,801</td>
<td>37.7</td>
</tr>
<tr>
<td>3–5</td>
<td>4,147</td>
<td>55.8</td>
</tr>
<tr>
<td>6–8</td>
<td>488</td>
<td>6.6</td>
</tr>
</tbody>
</table>

* Excludes three persons for whom information on gender was missing.

Most of the melanomas diagnosed between 1 and 2 years of the screening event had not been provisionally diagnosed at the screening. Consequently, using the 2-year time period

<table>
<thead>
<tr>
<th>Gold standard and screening lesion diagnosis</th>
<th>Melanoma diagnosed within 1 year</th>
<th>Melanoma diagnosed within 2 years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Any lesion</td>
<td>Melanocytic lesion</td>
</tr>
<tr>
<td>No. of true-positive screens</td>
<td>31</td>
<td>25</td>
</tr>
<tr>
<td>No. of false-positive screens</td>
<td>1,383</td>
<td>296</td>
</tr>
<tr>
<td>No. of false-negative screens</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>No. of true-negative screens</td>
<td>6,020</td>
<td>7,107</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>93.9</td>
<td>75.8</td>
</tr>
<tr>
<td>95% CI*</td>
<td>79.8, 99.3</td>
<td>57.7, 88.9</td>
</tr>
<tr>
<td>Specificity</td>
<td>81.3</td>
<td>96.0</td>
</tr>
<tr>
<td>95% CI</td>
<td>80.4, 82.2</td>
<td>95.5, 96.4</td>
</tr>
<tr>
<td>Positive predictive value</td>
<td>2.2</td>
<td>7.8</td>
</tr>
<tr>
<td>95% CI</td>
<td>1.5, 3.1</td>
<td>5.1, 11.3</td>
</tr>
<tr>
<td>Negative predictive value</td>
<td>0.0</td>
<td>99.9</td>
</tr>
<tr>
<td>95% CI</td>
<td>99.9, 100</td>
<td>99.8, 100</td>
</tr>
</tbody>
</table>

* CI, confidence interval.
as the gold standard resulted in a marked decrease in the sensitivity of the screening test, while the specificity and PPV remained very similar.

When the site of the melanoma was taken into account, the number of true-negative screens increased markedly (table 3). The number of true-positive screens decreased slightly because some of the provisional lesions were not on the same body site as the confirmed lesion, which resulted in a decrease in all measures of screening test validity except for negative predictive value. For the most precise screening diagnosis ("melanoma"), 1-year sensitivity was 63.6 percent, specificity was 99.5 percent, and PPV was 10.0 percent. Twelve of the 33 melanomas were missed when this precise screening diagnosis was used, compared with six that were missed when the least precise screening diagnosis was used.

None of the melanomas diagnosed within 1 to 2 years of the screening event were located on the same part of the body as that found on the screening diagnosis. This difference meant that sensitivity decreased markedly, while specificity and PPV remained the same when a 2-year period was used.

**DISCUSSION**

In this community-based study, accuracy of the screening diagnoses for melanoma was acceptable: 1-year sensitivity was 63.6 percent (95 percent confidence interval: 45.1, 79.6), specificity was 99.5 percent (95 percent confidence interval: 99.4, 99.6), and PPV was 10.0 percent (95 percent confidence interval: 6.3, 14.8). Sensitivity was higher if a screening diagnosis of any lesion was used as the definition of a positive screen, although at a cost of many more false-positive screens and a decreased PPV.

Lesions suspected of being melanoma were found in nearly 3 percent of those persons presenting to the Lions Cancer Institute clinics. The American Academy of Dermatology reported that of over 800,000 people presenting to their community-based skin screening clinic, 0.8 percent had a screening diagnosis of melanoma (9). Clinics that preferentially screen people with more skin cancer risk factors have found prevalences similar to ours, ranging from 1.7 percent (10) to 3.7 percent (11), although one small clinic in San Diego, California, found a prevalence of 7.5 percent positive screens (12).

To our knowledge, this study is the first to follow all negative screens by using a population-based cancer registry to calculate sensitivity of the screening diagnosis. The 1-year sensitivity of a Lions Cancer Institute diagnosis of melanoma was 69.7 percent when body site was ignored and 63.6 percent when body site was considered. The only other study known to follow up negative lesions to calculate sensitivity was a small one that did not report results for melanoma separately (13). One previous study attempted to adjust for false negatives by using a presumed rate of 10 false negatives per 100,000 screens (14). During the first year after the Lions Cancer Institute screening, six melanomas were not diagnosed at all at the screening, which equates to a rate of 82 false negatives per 100,000 screenings. This high rate is quite surprising considering that the screening was performed by specialist medical practitioners, and it underlines the difficulty of skin cancer screening in the general community.
Validity of Melanoma Screening

When body site was considered, the 1-year PPV of a Lions Cancer Institute screening diagnosis of melanoma was 10.0 percent. Other clinics that target people with skin cancer risk factors have found very similar PPVs for screening diagnoses of melanoma, ranging from 5.7 percent to 17.2 percent (10–12). The American Academy of Dermatology clinic, which did not target those at higher risk, found a PPV of 17 percent (15). The PPV for a screening diagnosis of a melanocytic lesion at the American Academy of Dermatology clinic was 6.3 percent (15), which compares well with our finding of 6.6 percent.

The 1-year PPV of a screening diagnosis of any lesion was 1.5 percent, which is lower than previously reported PPVs of 2.7–7.3 percent (10, 13, 16, 17). This discrepancy is probably due to differences in the types of lesions recorded in the different clinics and to the high prevalence of solar keratoses and nonmelanoma skin cancers in the Australian population (18, 19).

Compared with other screening modalities, skin cancer screening diagnoses appear to be reasonably accurate. In a large US study, mammography for invasive breast cancer had an overall sensitivity of 78.7 percent, specificity of 87.7–89.4 percent (20), and PPV of 3.1 percent. Compared with our skin cancer screening, conventional Papanicolaou smear screening has lower sensitivity (about 51 percent) but higher specificity (about 98 percent) (21).

To our knowledge, no one has previously considered the issue of negative screens on different body sites. This is not surprising, since most studies have followed only positive screens over a fairly short time interval from the screening clinic to a diagnostic and treatment clinic. This issue arose for us during the matching process because of several instances in which a screening diagnosis had been made of a lesion that was clearly not on the same body site as the confirmed melanoma. We then debated whether to count that person as having a positive or negative screen since, in reality, he or she had both. Our solution is, we think, appropriate. When we did not take body site into consideration, two screening diagnoses of melanoma were incorrectly thought to be true positives when in fact the actual melanoma was on another part of the body. This issue meant that the true screening sensitivity and PPV were lower than that calculated if we ignored body site completely. Sensitivity and negative predictive value increased slightly when body site was taken into account. Theoretically, there may still have been some misclassification of body site when, for example, the screening lesion was on the right arm and the confirmed melanoma was on the left arm.

In addition, the use of body sites showed us that all melanomas diagnosed more than a year after the screen were not diagnosed at the screening event. This situation was not so clear when body sites were ignored in the analysis; three of the 16 additional melanomas were found on participants who had had positive screens. However, these three confirmed melanomas were found on body sites different from that indicated in the provisional screening diagnosis. This finding suggests that, to avoid missing melanomas, the screening interval should be about 1 year rather than 2.

A large number of screens were performed by the Lions Cancer Institute (n = 7,436), and a reasonable number of melanomas were detected (n = 49). The American Academy of Dermatology clinics have reported on over 250,000 screenings and 364 true melanomas (15). However, all other studies have included between 374 and 4,146 screenings and one to 11 true melanomas only (10–14, 17). The number of screenings meant that confidence intervals around most values were quite narrow, but the confidence intervals around those for sensitivity were still quite wide.

The data used in this study originated from a screening clinic operating in the general community in an area with a high incidence of melanoma. The subjects who attended were self-selected, not randomly selected from the community. As such, they may have been more conscious of their skin health and may have had lesions that they wanted to have checked (5). People were asked to attend only if they had three or more of a given set of risk factors. Their risk level was self-assessed and was not validated.

The screening program has included several different medical specialists; although they were all accredited, they did not have standardized training for this particular clinic. We did not have information on which screenings were performed by each specialist. Anecdotally, it is thought that the threshold for calling a screen positive was quite low at these clinics.

The clinic undertook screening only, not diagnosis or treatment, and there were no mechanisms in place to ensure that people with positive screening diagnoses were followed up by their local general practitioner. Some true melanomas may have been present at screening that were never diagnosed or biopsied during the 2 years after a screen. However, awareness of melanoma is relatively high in Australia (22), and we think it unlikely that many melanomas would have been present for 2 years without having been treated.

The behaviors of people after attending skin cancer clinics have not been studied in detail. It may be that simply having been screened increases awareness of skin lesions and makes people more likely to seek appropriate care when a lesion does occur or change. Additionally, having even a nonmelanocytic lesion provisionally diagnosed may result in closer scrutiny of the skin by both the person and medical professionals. This may mean that, although a particular screen-detected lesion was a false positive, the screening itself is directly responsible for other lesions being diagnosed. In this instance, use of site-specific lesions will slightly underestimate the effectiveness of the screening program.

Because notification of cancers to the Western Australian Cancer Registry is mandatory, it is thought that registration of melanomas occurring in Western Australia is almost complete. Data are routinely linked to the Western Australia Linked Database (which covers all hospital admissions in the state) and to the electoral roll (registration is compulsory for Australian citizens); therefore, name and address changes are usually recorded by the Western Australian Cancer Registry. We may have missed a few people who changed their name without notifying any authority or interacting with the hospital system. In addition, we may have missed people who were screened and moved interstate or overseas within 2 years. Outmigration from Western Australia is about 2.7 percent per year, and migrants tend to be younger than the population in our study (23).
This study has shown that the validity of skin screening diagnoses in a real-life population-based clinic is reasonably good. The level of precision of the screening diagnosis markedly affected accuracy; few melanomas were missed if any lesions were considered a positive screen but at a cost of having to investigate a much larger number of false positives. Accuracy decreased markedly if a longer follow-up time was used. Most of the lesions found in the second year after screening were not on the same body site as noted on the previous positive screen, suggesting that the screening interval should be less than 2 years. Body site should be taken into account when calculating the validity of skin screening.

ACKNOWLEDGMENTS

This study was funded by the Lions Cancer Institute. The authors thank Colin Beauchamp, Chairman of the Lions Cancer Institute, for his consistent support; Dr. Tim Threlfall for matching the screening lists with the Western Australian Cancer Registry; Dr. Peter Baade for running the models to assess clustering; Professors Michael Millward and Peter Leadman, who were medical directors of the Lions Cancer Institute during this study; Lisa van der Vlugt for data entry; Cassandra Clayforth and Heather Williams for assistance with study design and data management; the Board of the Lions Cancer Institute for funding this study; and all the many dedicated volunteers in the Lions Cancer Institute Skin Cancer Screening Program.

Conflict of interest: none declared.

REFERENCES