Aims and objectives

In all areas of medicine, the measurement of continuous variables is commonly used as an aid in making diagnostic or prognostic determinations relating to patient treatments and outcomes. Specifically, based on the value of a continuous variable, clinical decision making often involves classifying individuals based on the value of a continuous variable as having or not having a medical condition, or being graded into one of several categories of severity [1].

In medical research, there is often the perception that categorizing a continuous variable simplifies the statistical model and the interpretation of the ensuing results. While by Occam's Razor more parsimonious models are preferred, over-simplification can in fact result in the loss of valuable information, loss of power, and poorer predictive performance [1,2].

The choice of cut-points to use when categorizing continuous variables presents several dilemmas. The rationale for choice of cut-points is typically ad-hoc and unjustified. When considering the relationship between two continuous variables, dichotomization of one of the variables can be equivalent to losing one third of the sample, and dichotomization of both variables can be the equivalent of losing two-thirds of the sample [3]. In addition, as outlined by Altman and Royston [1], dichotomizing continuous variables results in a loss of information and reduced power to detect a relationship between the variable and the clinical outcome; patients with similar measurements found on either side of a cut-point may be described as being very different when they are not; non-linear relationships are concealed; and although less information is lost when variables are split into multiple categories statistical modeling and interpretation becomes more complex. Finally, dichotomizing a variable can lead to false positive findings [4].

Percent mammographic breast density (PMD) is a risk factor for breast cancer. Traditionally, visual assessments of PMD have been used in clinical settings and such measurements have been made using four categories, each spanning a 25% interval of the breast density spectrum (which ranges from 0-100%) [5]. However, it has long been recognized that PMD is a continuous measure and over the past 35 years methods to measure PMD on a continuous scale have evolved from the subjective and time-consuming task of planimetry using screen-film mammograms to fully-automated algorithmic assessments using digital mammograms [6,7].

In an attempt to improve breast cancer risk models, breast density is increasingly used as a predictor of breast cancer; however, the application of breast density in such models is variable, with some models using breast density as a categorical (and sometimes dichotomous) variable and others as a continuous variable [8-11].
While the use of categorical PMD scales is appealing due to the simplicity of interpretation, the use of such a measure of PMD assumes equal risk within a given category [2]. For example, a woman with 76% PMD is assumed to have the same risk as a woman with 95% PMD if the BI-RADS 4th edition density classification scale is used. Using a dichotomous PMD scale, a woman with 49% density is classified as having low density and a woman with a very similar density of 51% is classified as having high density. This categorization introduces a high degree of variability and reduces the power of the model to detect a relationship between breast density and cancer risk, which further leads to inaccurate risk estimates. Additionally, when empirically derived quantiles of PMD are used based on a particular study sample, the external validity is difficult to assess and results can be difficult to compare across studies [2].

The objective of this study was to examine the extent of information loss associated with decreased precision of PMD assessments, and the commensurate loss of predictive power in breast cancer risk models in which screen-detected breast cancer is the outcome of interest.
Fig. 1: BI-RADS 4th edition mammographic density descriptors; a 4-category density scale derived from a continuous density measure.

© Department of Diagnostic Radiology, Nova Scotia Health Authority - Halifax/CA
Methods and materials

Four-view, standard "For presentation" screening mammograms were obtained from 2,374 women attending a population-based breast screening program between 2009 and 2011. These women were part of a case-control study that has been previously described [12].

An automated PMD measurement software (DM-Density 2.0, Densitas Inc.) was used to assess baseline PMD for each mammography study in 1% increments. Less precise PMD measures were subsequently derived by rounding upward to the nearest 2%, 3%, 4%, …, 25% and 50% so that in total, 26 measures of percent density with varying levels of granularity were available for analysis.

R-squared and root mean square error (RMSE) from a linear regression model were used to evaluate the effect of decreasing precision in PMD assessments against the baseline assessment of PMD in 1% increments. Additionally, DeLong's test was used to compare the area under the receiver operator characteristic (AUROC) curve between cancer risk models employing PMD measures of varying precision.
Results

Compared to the baseline PMD measure, when rounded into broader, less precise values, R-squared decreased from 1.00 (1% PMD) to 0.58 (50% PMD), while RMSE increased from 0.00 (1% PMD) to 29.10 (50% PMD).

<table>
<thead>
<tr>
<th>Percent Mammographic Density Scale</th>
<th>R-squared</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1% increments (baseline)</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>5% increments</td>
<td>0.99</td>
<td>2.46</td>
</tr>
<tr>
<td>10% increments</td>
<td>0.97</td>
<td>5.36</td>
</tr>
<tr>
<td>20% increments</td>
<td>0.90</td>
<td>10.60</td>
</tr>
<tr>
<td>25% increments</td>
<td>0.87</td>
<td>13.42</td>
</tr>
<tr>
<td>50% increments</td>
<td>0.58</td>
<td>29.10</td>
</tr>
</tbody>
</table>

Table 1: R-squared and root mean square error (RMSE) based on linear regression model comparing rounded categorical measures of percent mammographic density to a continuous measure of percent mammographic density.

References: Department of Diagnostic Radiology, Nova Scotia Health Authority - Halifax/CA

When used to evaluate cancer risk, the AUROC curve for baseline PMD (0.554) was not significantly different from the AUROC curve for PMD measured in 5% increments (0.551, p = 0.228) but was significantly higher than that for PMD measured in 25% increments (0.535, p = 0.006) and that measured in 50% increments (0.508, p = 0.007).
Fig. 2: Categorizing continuous mammographic density measures leads to increasingly more loss of information as categories become increasingly broad and relates to reduced accuracy of cancer risk models; dichotomizing continuous density measures can lead to loss of a significant amount of information such that risk model performance may be reduced to no better than random guessing (AUC = 0.508).

References: Department of Diagnostic Radiology, Nova Scotia Health Authority - Halifax/CA
Conclusion

Using PMD assessed in 5% vs. 1% increments resulted in no significant loss of information in modelling breast cancer risk. However, less precise density scales, such as PMD measured in 25% or 50% increments, lead to loss of information and lower performing cancer risk models.

Breast cancer risk model performance benefits from the use of continuous or near-continuous measurements of breast density. More course or broadly defined breast density scales, such as the BI-RADS density scale and any dichotomization of that scale into high vs. low density categories, may be less informative for inclusion in breast cancer risk models. While radiologists have demonstrated an ability to assess density using a 1% scale (i.e., using a Visual Analogue Scale) and 5% scale (i.e., visual assessment on a 21-point scale), inter- and intra-rater variability remains a limitation of these methods. Automated breast density algorithms provide standardized, reproducible density assessments on a continuous or near-continuous scale and our results demonstrate that this would provide more information than categorical density scales in breast cancer risk models.
Personal information

Contact Author: Mohamed Abdolell, MSc, PStat, Department of Diagnostic Radiology, Dalhousie University, Halifax, Nova Scotia, Canada, mo@dal.ca
References