the sensitivity and specificity observed among the diseased and nondiseased, respectively.

REFERENCES

Frédéric Lagarde
Lars Alfredsson
Institute of Environmental Medicine
Karolinska Institute
S-171 77 Stockholm, Sweden

APPENDIX

Let \( C_T, C_Z, \) and \( C_W \) be the vectors of cell counts in the correctly classified table, in the table based on the alloyed gold standard, and in the table based on the usual measure, respectively.

Let \( M_Z \) and \( M_W \) be invertible misclassification matrices corresponding to the alloyed gold standard and to the usual measure, respectively, i.e.,

\[
C_Z = M_Z C_T, \quad (A1)
\]

\[
C_W = M_W C_T, \quad (A2)
\]

and

\[
C_T = M_Z^{-1} C_Z, \quad (A3)
\]

\[
C_T = M_W^{-1} C_W. \quad (A4)
\]

Let \( M_{WZ} \) be an invertible misclassification matrix corresponding to the misclassification of the alloyed gold standard in the usual measure, i.e.,

\[
C_W = M_{WZ} C_Z.
\]

Now equations A2 and A3 imply that \( C_W = M_W M_Z^{-1} C_Z \), giving the misclassification matrix resulting from validation of the usual measure against the alloyed gold standard \((M_W M_Z^{-1} = M_{WZ})\).

Using, as specified by the correction method (3), the inverse of that matrix to obtain corrected cell counts gives

\[
M_{WZ}^{-1} C_W = M_Z M_W^{-1} C_W = M_Z C_T = C_Z
\]

(using equations A4 and A1), which are the cell counts that would be expected using the alloyed gold standard.

THE AUTHORS REPLY

We thank Drs. Lagarde and Alfredsson (1) for contributing to our understanding of the effects of misclassification when using an alloyed gold standard (2). They point out that the misclassification probabilities for the usual measure \( W \) with respect to the alloyed gold standard \( Z \) can be differential between cases and controls, even when \( W \) and \( Z \) are nondifferential with respect to the true gold standard \( X \). The apparent sensitivity of \( W \) using \( Z \) as a gold standard depends on the relative frequencies of \((X = 1, Z = 1)\) and \((X = 0, Z = 1)\), because \( Z = 1 \) arises from \( X = 0 \) and \( X = 1 \) with fixed proportions. If the distribution of \( X \) differs between cases and controls, the relative frequencies also differ, and the result will be differential misclassification of \( W \) with respect to \( Z \). Lagarde and Alfredsson (1) further contend that applying the misclassification matrix from the validation study in the controls to the cases when the error is differential is the source of the potential exaggeration of effect; however, using estimates of misclassification matrix for \( W \) with respect to \( Z \) from cases and controls separately will lead to the same estimates as if \( Z \) had been used.

While the argument is essentially correct, one of the authors’ claims and the penultimate paragraph of their Appendix (1, p. 1176) may be misleading. When nondifferential misclassification is assumed, as is the usual practice, the bias in the corrected estimate does depend on the correlation between the errors in \( Z \) and \( W \), as demonstrated clearly in table 5 of our original paper (2, p. 1254). Secondly, in their Appendix, the matrix \( M_W M_Z^{-1} \), which they equate with \( M_{WZ} \), may not even be a misclassification matrix. More importantly, \( M_{WZ} \) cannot be inferred from \( M_{WZ} \) and \( M_{ZX} \) in fact, more than one \( M_{WZ} \) is consistent with the same \( M_{WX} \) and \( M_{ZX} \). For example, if \( X = 1 \) and \( X = 0 \) each have the probability 0.5 and \( W \) and \( Z \) each have a sensitivity and specificity of 0.6 for \( X \) in cases and in controls, there are infinitely many possibilities for \( M_{WZ} \); these include, when \( Z \) equals \( W \), \( M_{WZ} \) being the identity matrix (100 percent sensitivity and specificity), and, when the errors in \( Z \) and \( W \) are independent, \( M_{WZ} \) being the matrix with entries 0.52 = 0.62 + 0.42 on the diagonals and 0.48 = 2 \( \times (0.6 \times 0.4) \) for the nonzero off-diagonals. Notwithstanding this point, Lagarde and Alfredsson are correct in saying that, in expectation, the odds ratio estimate obtained after correction of \( W \) with respect to \( Z \) differentially equals what would be observed if \( Z \) had been used throughout, regardless of the specific \( M_{WZ} \).

We continue to be wary of correction methods based on validation studies using alloyed gold standards that are far from the truth. The cost of the validation study may become prohibitive when data from cases and controls must be used separately (3). However, when feasible, allowing for differential error may be more robust than assuming nondifferential, and therefore is worth considering if one plans to use a validation study to correct for misclassification.

REFERENCES

Sholom Wacholder
Patricia Hartge
National Cancer Institute
Executive Plaza North
Bethesda, MD 20892

Ben Armstrong
London School of Hygiene and Tropical Medicine
Keppel Street
London WC1E 7HT, England

Am J Epidemiol Vol. 143, No. 11, 1996