# Supporting Information for "Corrected ROC Analysis for Misclassified Binary Outcomes"

Matthew Zawistowski, Jeremy B. Sussman, Timothy P. Hofer, Douglas Bentley, Rodney A. Hayward and Wyndy L. Wiitala

| Baseline Risk $e^{\beta_0}/(1+e^{\beta_0})$ | Effect Size $\beta_1$ | False Positive Rate $\gamma_0$ | False Negative Rate $\gamma_1$ | True Outcome ROC $AUC(T, \hat{P})$ | Misclass. Outcome ROC $AUC(Y, \hat{P}^I)$ | Mean Bias (% Change) |
|---|---|---|---|---|---|---|
| 0.6 | 0.1 | 0.05 | 0.05 | 0.527 | 0.524 | 0.57 |
| | | | 0.2 | 0.527 | 0.519 | 1.53 |
| | | 0.2 | 0.05 | 0.527 | 0.521 | 1.13 |
| | | | 0.2 | 0.527 | 0.515 | 2.29 |
| | 1.0 | 0.05 | 0.05 | 0.739 | 0.714 | 3.41 |
| | | | 0.2 | 0.739 | 0.675 | 8.75 |
| | | 0.2 | 0.05 | 0.739 | 0.689 | 6.87 |
| | | | 0.2 | 0.739 | 0.641 | 13.30 |
| 0.8 | 0.1 | 0.05 | 0.05 | 0.528 | 0.523 | 0.99 |
| | | | 0.2 | 0.528 | 0.514 | 2.70 |
| | | 0.2 | 0.05 | 0.528 | 0.520 | 1.47 |
| | | | 0.2 | 0.528 | 0.510 | 3.42 |
| | 1.0 | 0.05 | 0.05 | 0.742 | 0.704 | 5.15 |
| | | | 0.2 | 0.742 | 0.640 | 13.76 |
| | | 0.2 | 0.05 | 0.742 | 0.687 | 7.39 |
| | | | 0.2 | 0.742 | 0.618 | 16.79 |

Table S1: Bias in Area Under the ROC curve (AUC) for misclassified outcomes in a population cohort in which cases outnumber controls. In this scenario, misclassifying cases as controls (false negative) has a greater impact on AUC bias than does misclassifying true controls.

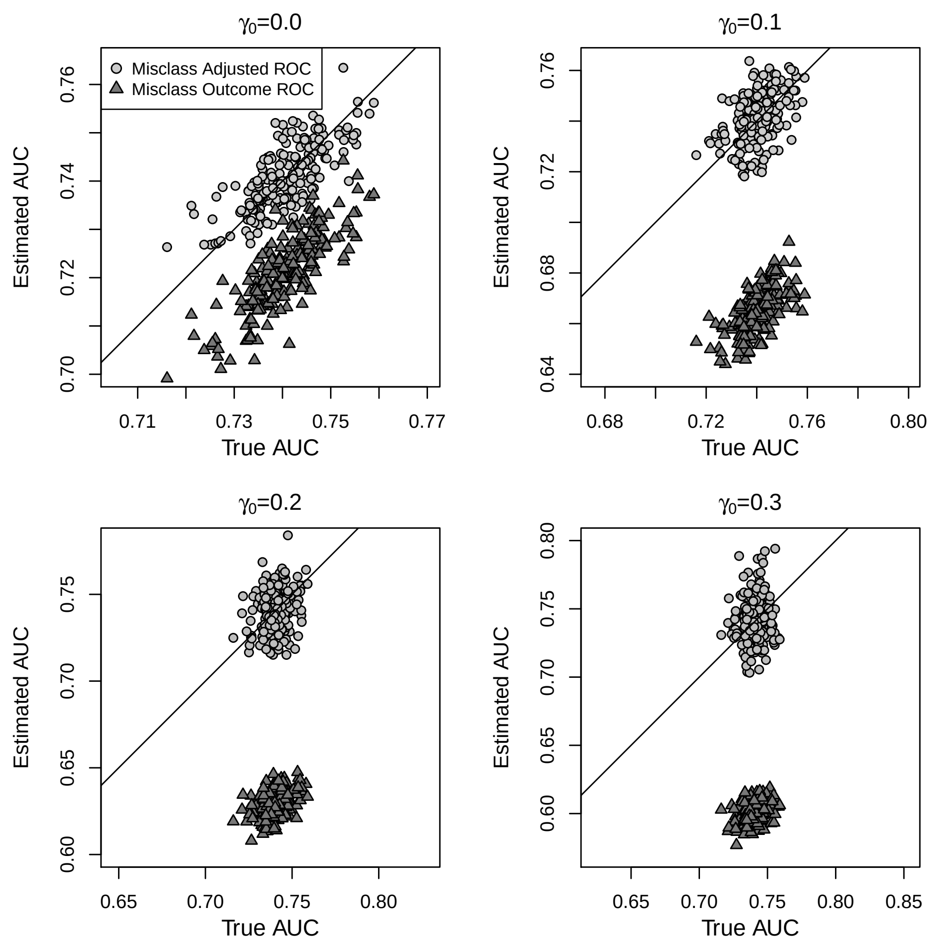


Figure S1: Scatterplots of standard AUC estimates and misclassification-adjusted AUC estimates versus the true AUC values for various levels of misclassification. In each figure, simulations were for model parameters $\beta_0 = -1$ and $\beta_1 = 1$ and fixed false-negative misclassification rate of $\gamma_1 = 0.2$. As misclassification increases, bias increases in standard estimates of AUC (triangles). The misclassification-adjusted AUC estimators (circles) have bias of nearly zero but increasing variance as misclassification increases.

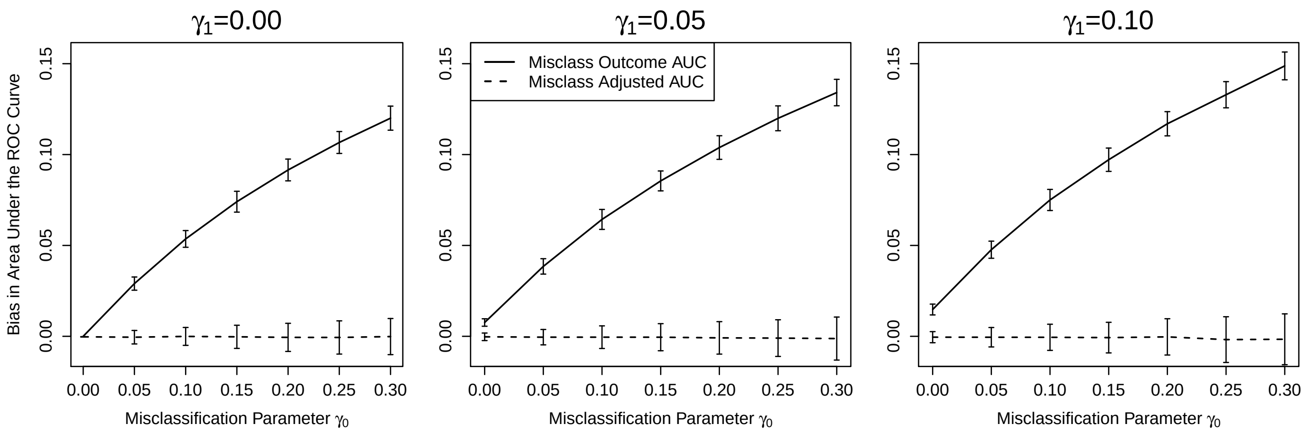


Figure S2: Bias in AUC for standard and misclassification-adjusted ROC procedures over a range of constant misclassification values for a model with multiple predictor variables ($\beta_0 = -1, \beta_1 = 1, \beta_2 = 0.5, \beta_3 = -1$). For each simulated dataset, we computed bias as the difference between the AUC value based on the true outcome data $AUC(\mathbf{T}, \hat{P})$ and the estimates using the observed outcome data $AUC(\mathbf{Y}, \hat{P}^M)$ and $AUC_M$. Mean bias in AUC for standard ROC computation (solid lines) increases with increasing amount of misclassification. The proposed misclassification-adjusted ROC procedure (dashed lines) has bias of nearly zero over all combinations of false positive and false negative values. The vertical bars give standard errors for the AUC estimates.

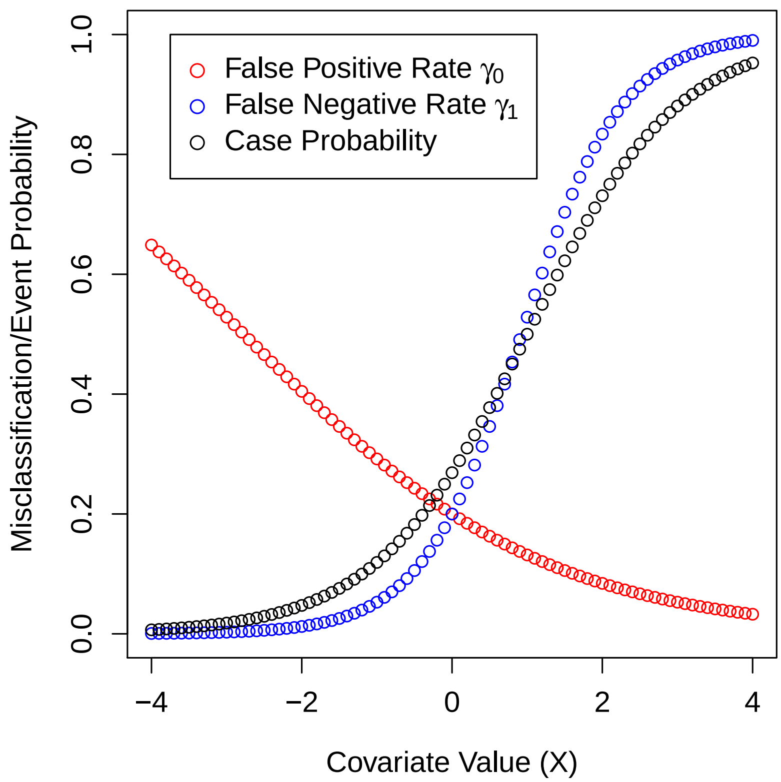


Figure S3: Graphical display of false positive, false negative and event probabilities for differential misclassification scenario 1 in which the false positive rate increases with the probability of being a control and false negative rate increases with the probability of being a case. The misclassification in this scenario leads to an underestimate of the true AUC value.

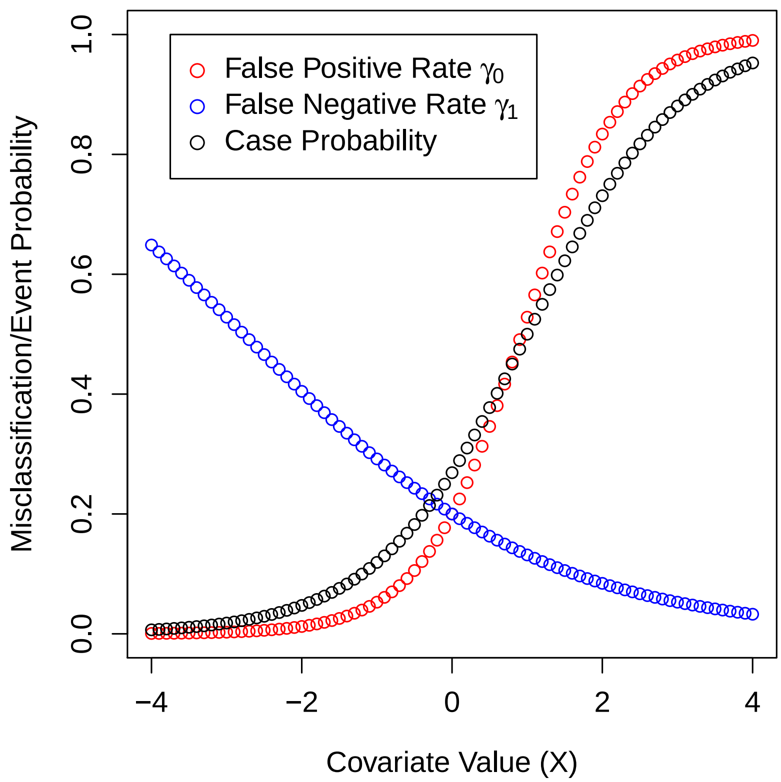


Figure S4: Graphical display of false positive, false negative and event probabilities for differential misclassification scenario 2 in which the false positive rate increases with the probability of being a case and the false negative rate increases with the probability of being a control. The misclassification in this scenario leads to an overestimate of the true AUC value.

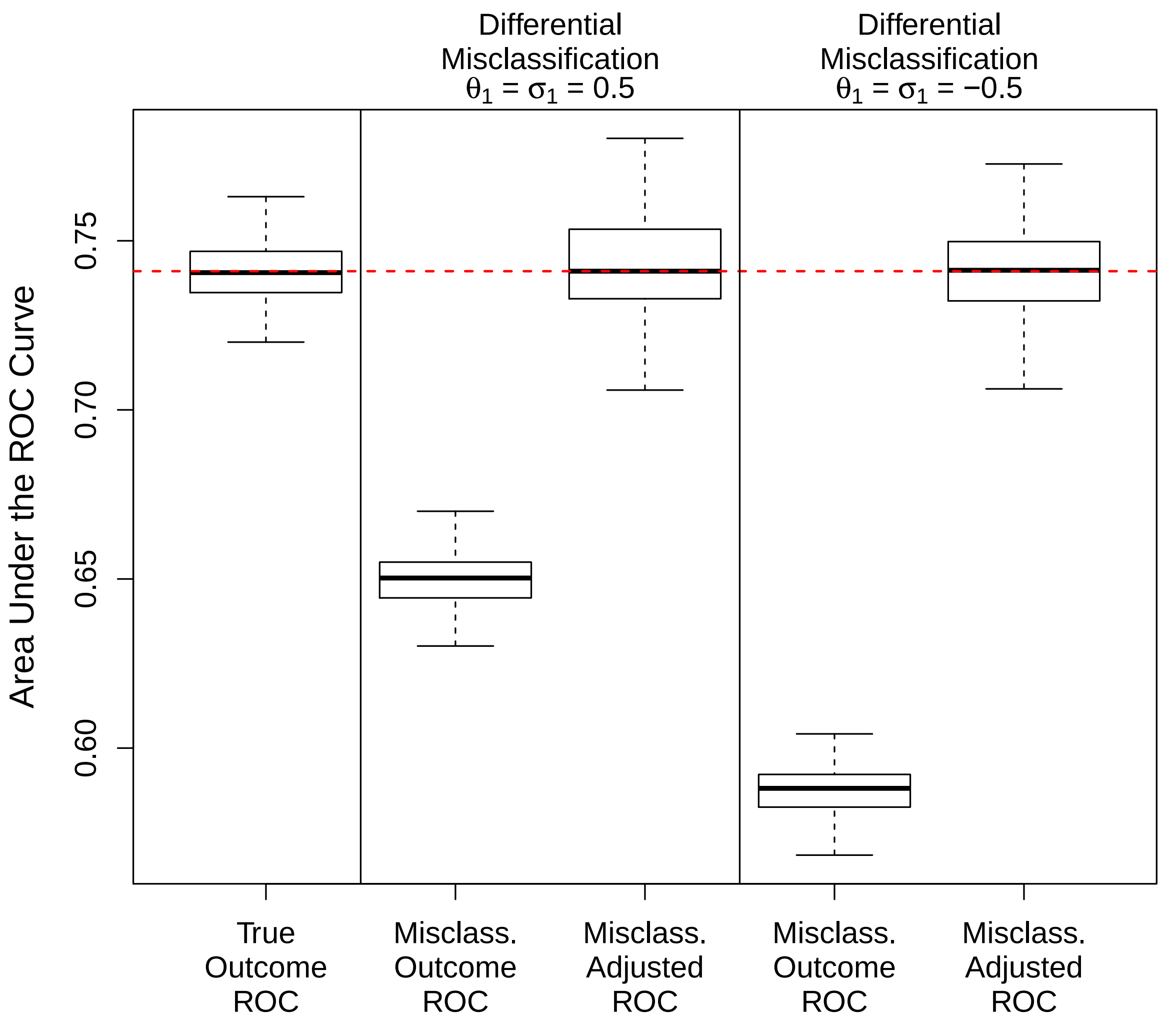


Figure S5: Examples of differential misclassification in which false positive and false negative rates have same direction. When $\theta_1 = \sigma_1 = 0.5$, both the false positive and false negative rates increase with the probability of being a case and the misclassification leads to an underestimate of the true AUC. When $\theta_1 = \sigma_1 = -0.5$, both the false positive and false negative rates decrease with the probability of being a case and, again, the misclassification leads to an underestimate of the true AUC. In both cases, the misclassification-adjusted ROC method substantially reduces AUC bias.

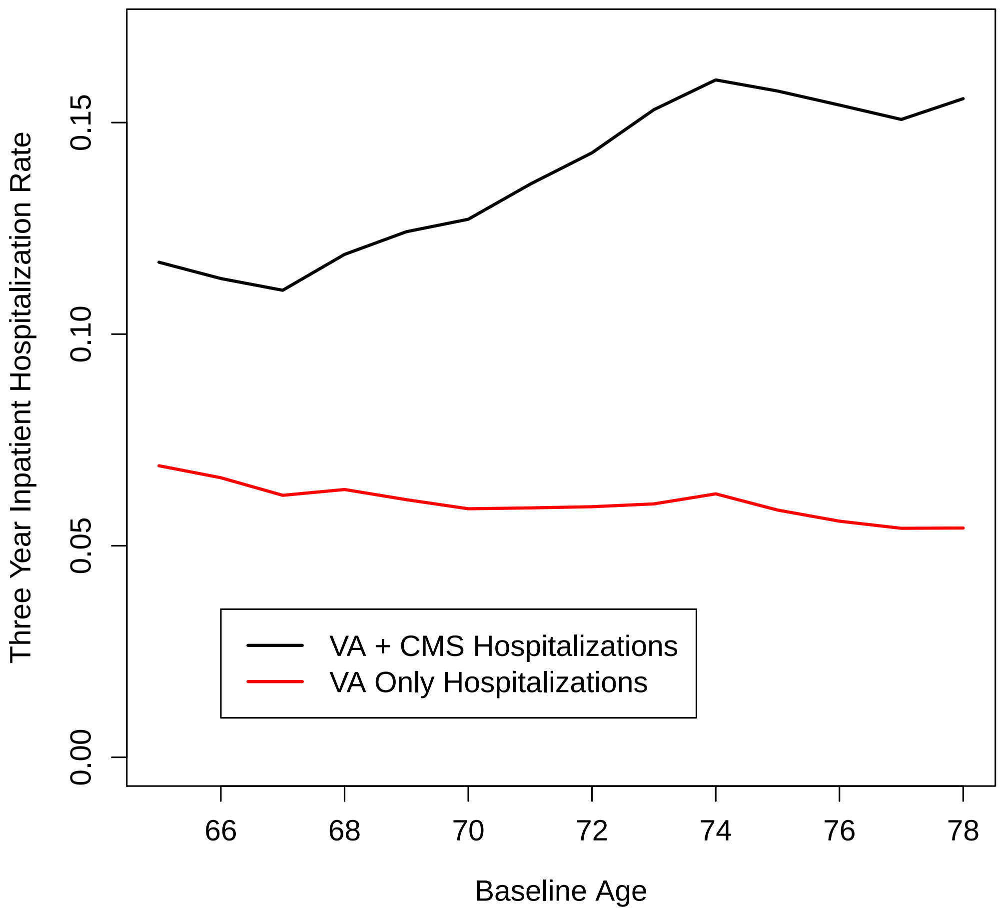


Figure S6: Three year inpatient hospitalization rates by baseline age using events recorded in either VHA or CMS records (black line) and only VHA EHR events (red line). Using only the VHA EHR leads to a misclassification of hospitalization outcome events as non-events. The change in difference between event rates by age indicates covariate dependent misclassification.

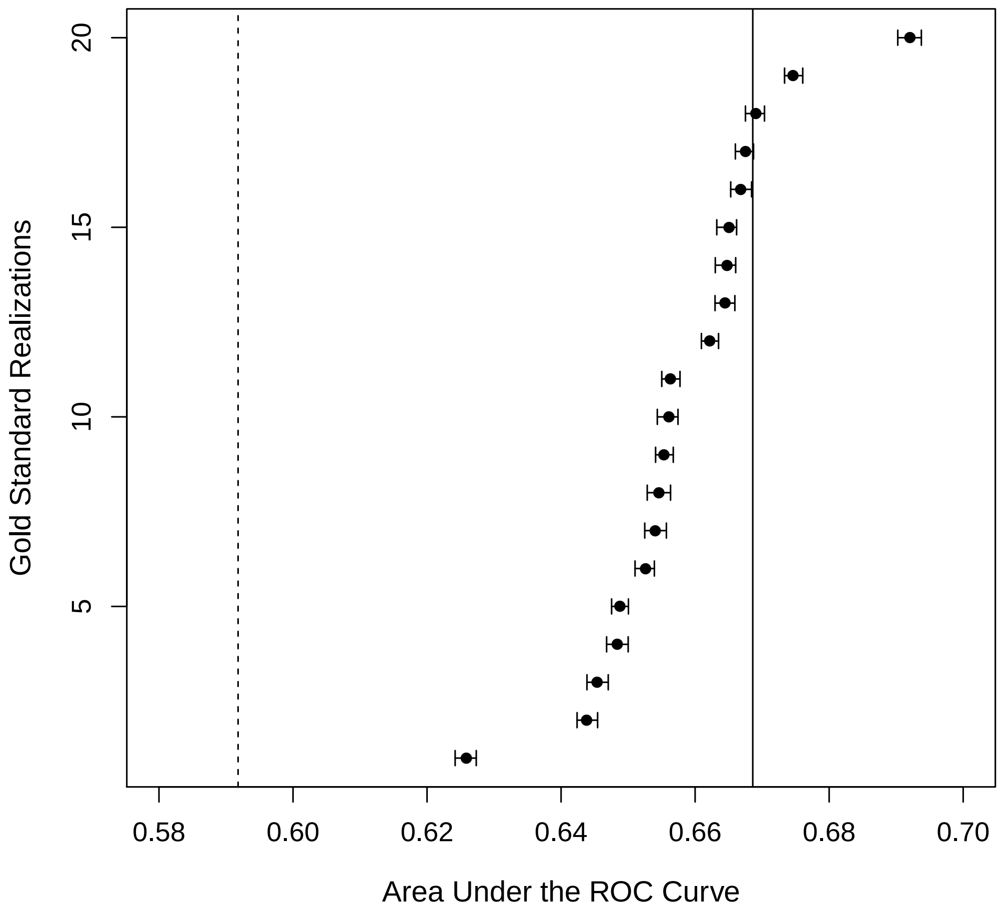


Figure S7: Bootstrap-based 90% confidence intervals of $AUC_M$ for 20 realizations of the internal validation cohort in the data example. The solid black line at AUC=0.669 indicates the true AUC value and the dashed line at AUC=0.592 the nave analysis of misclassified outcomes. Coverage of the bootstrap-based confidence intervals suffered due to the misclassification rates being estimated rather than known exactly.