

Appendix

Detailed Statistical Methods

Internal Validation of the Final Multivariable Model

Robust internal validation of the final multivariable model was performed using calibration, discrimination, and bootstrapping techniques³⁷⁻³⁹. Calibration measures how closely predicted risk agrees with observed risk. Calibration was assessed for each tenth of predicted risk with use of 10 equally sized groups, with a Hosmer-Lemeshow goodness-of-fit test performed to quantify calibration; a nonsignificant p value indicated a well-calibrated model. Discrimination is the model's ability to differentiate between metal-on-metal hip resurfacings (MoMHRs) with and without evidence of a pseudotumor, which is assessed by calculating the area under the receiver operating characteristic curve (AUC). The AUC for a useful prognostic model is between 0.60 and 0.85³⁸. The final multivariable regression model was further validated by using bootstrapping with backward stepwise variable deletion³⁷. Two hundred bootstrap repetitions were performed to obtain a bias-corrected estimate of the AUC with use of a modified data set.

Developing a Clinical Risk Scoring Tool

With use of the patient and radiographic predictors that were included in the final multivariable logistic regression model, a clinical risk scoring tool was developed using previously described methods⁴⁰. The calculated overall score represents a patient's risk of having evidence of a pseudotumor, with higher scores associated with increased pseudotumor risk. The respective regression coefficient for each predictor from the final multivariable model was converted to an integer risk score. The reference group for all variables was assigned a risk score of 0. For all variables other than inclination, the risk scores for the nonreference groups were calculated by multiplying the respective regression coefficient by 10 and then rounding to the nearest whole number. Inclination data were grouped with the midpoint for each group taken; the midpoint for the 30° to 39° group was 34.5°. The midpoint for nonreference inclination groups was then multiplied by the respective regression coefficient, multiplied by 10, and rounded to the nearest whole number. The discriminatory ability of the final overall risk score was assessed by calculating the AUC. The optimal risk score threshold for identifying MoMHRs with evidence of a pseudotumor was also calculated.