

Appendix***Analysis of Male and Female Cohorts from the BHR Group***

An analysis of the demographic characteristics of the study patients demonstrated a significant difference between the ratio of male patients to female patients in the group that had adverse reactions to metal debris and the group that did not have adverse reactions to metal debris for the BHR subgroup of patients ($p = 0.0001$), but not for the Corail-Pinnacle subgroup ($p = 0.5967$). ROC analysis was therefore carried out on the separate male and female groups for the BHR group to see whether different thresholds might be applicable to each sex. The additional analysis was restricted to cobalt metal ions because cobalt had been found to produce the best diagnostic test results in the main body of work.

There were twenty-two female patients and four male patients in the BHR group who had adverse reactions to metal debris.

Table E-1 shows that the optimal thresholds established from the ROC analysis were 2.27 $\mu\text{g/L}$ for the male group and 2.12 $\mu\text{g/L}$ for the female group in comparison with 2.15 $\mu\text{g/L}$ for the BHR group as a whole. Although the optimal cobalt thresholds appear similar for both male and female patients, the very low number of failed implants in male patients mean that the reliability of the results is not as good as when the BHR group was analyzed as a whole. In the future, if a greater cohort of patients with failed BHR implants becomes available, this analysis could be repeated to see whether any major differences between male and female patients are observed.

TABLE E-1 Summary of ROC Analysis for Cobalt in Patients with the BHR Implant Stratified by Sex

	Group Analyzed for Cobalt		
	All (N = 309)	Male (N = 164)	Female (N = 145)
AUC* (%)	90.5 (82.8 to 98.1)	90.3 (82.5 to 98.1)	76.3 (65.8 to 86.8)
Optimal threshold for cobalt ($\mu\text{g/L}$)	2.15	2.27	2.12
Sensitivity* (%)	88.5 (76.9 to 100)	88.5 (76.9 to 100)	57.7 (38.5 to 76.9)
Specificity* (%)	84.5 (80.2 to 88.3)	79.2 (74.2 to 83.8)	87.3 (83.4 to 90.8)
Positive predictive value* (%)	34.3 (24.1 to 46.3)	28.0 (19.5 to 38.6)	29.4 (18.7 to 43.0)
Negative predictive value* (%)	98.8 (96.4 to 99.6)	98.7 (96.2 to 99.5)	95.7 (92.5 to 97.6)
Misclassification of patients (%)	15.2	20.1	15.2
Positive likelihood ratio*	5.69 (4.19 to 7.72)	4.24 (3.25 to 5.54)	4.54 (2.90 to 7.11)
Negative likelihood ratio*	0.14 (0.05 to 0.40)	0.15 (0.05 to 0.42)	0.49 (0.31 to 0.76)

*The values are given as the mean percentage, with the 95% CI in parentheses.