

**Table 2**

Misclassified Estimates of the Effect of Smoking on Exposure and Outcome Using ICD-9 Codes and Natural Language Processing

Estimates	OR	95% CI
ICD-9 and readmission (outcome), Yes vs. No	0.94	0.90–0.98
ICD-9 and gender (exposure), Male vs. Female	0.68	0.63–0.74
NLP and readmission (outcome)	1.04	0.96–1.14
NLP and gender (exposure)	1.08	0.92–1.26

*Abbreviations:* CI, confidence interval; ICD-9, *International Classification of Diseases, Ninth Revision*; NLP, natural language processing; OR, odds ratio.