

Table 2

Proportion of Mismatches between Data Discrepancy Subtypes

Data Element	Percentage of Total Duplications (n = 398,939)	Percentage of Duplicate Pairs with Discrepancy by Data Element
First Name	22.6%	FN–MN swap: 4.0% FN–LN swap: 7.8% FN blank in one record: 31.3%
Middle Name	58.1%	FN–MN swap: 1.6% LN–MN swap: 0.8% MN blank in one record: 62.0% Middle initial in one record and full MN in the other record: 20.3%
Last Name	23.7%	FN–LN swap: 7.4% LN–MN swap: 2.0% LN blank in one record: 5.1%
Gender	6.0%	Gender blank in one record: 16.5%
SSN	53.4%	Default SSN: 50.2% SSN blank in one record: 58.2%
DOB	6.2%	DOB blank in one record: 0.4% Month does not match: 22.1% Day does not match: 48.3% Year does not match: 40.0% Month and day swapped: 2.6% Month and year swapped: 0% Day and year swapped: 0%

Abbreviations: DOB = date of birth, FN = first name, MN = middle name, LN = last name, SSN = Social Security number.