**Algorithm description and justification:**

We know there are 7862 I10 codes that each code mapped to many I9 codes. This algorithm takes advantage of 2015 California State ICD9 diagnosis codes to figure out which I10 1:many code can be converted to 1:1 code. First, we figured out which I10 code is 1:many code and store them into a dictionary. We use the 1:many I10 code as key and mapped I9 codes as value. Then, for each 1:many I10 code, we summed up its mapped I9 codes’ principal diagnosis cases(dx0). If any of those mapped I9 code’s dx0 divided by total sums of those I9 code’s dx0s larger than 0.5, which means this I9 code is the one that most frequently used as the principal diagnosis among all mapped I9 codes, even more frequent than the sum of the rest. I think in this situation, it is reasonable to adopt this most frequently used I9 code as the pair of 1:1 mapping of the I10 code. First, it can somewhat represent the group of mapped I9 codes, since its frequency is larger than the sum of the rest. Next, since our goal is to convert as many 1:many mappings as possible, so I set the threshold as larger than 0.5.

**Quantification of converted 1:many I10 codes:**

5666 one-to-many mappings were reduced to one-to-one mapping.