

If we wanted to use one column, rather than two, we would need to have 9 feature factor levels. This can be understood with the tables below. (should if be m-1)

Person ID:	DH:	SL:	Diagnosis
1	0	0	Healthy
2	0	1	At risk of SL
3	1	0	At risk of DH
4	1	1	At risk of SL and DH
5	2	0	Has DH
6	0	2	Has SL
7	2	1	Has DH, at risk of SL
8	1	2	Has SL, at risk of DH
9	2	2	Has DH and SL

As you can see in the table, if we include two columns, one for DH status and one for SL status, we only need to have three factor levels; one for being at risk of that disease, one for having that disease, and one for not being at risk of that disease. If we do not separate these diseases into two columns, then we need to have a different factor level for every possible health status a person has in regards to these two diseases.