Appendix 1 (as supplied by the authors): Variable importance for the top 20 features, as measured by mean decrease in the Gini index across 500 trees in a Random Forest Model

Variable	Mean Decrease in Gini Index
Everywhere text "type 1"	4.798
Problem list text "type 1"	4.239
Medication A10AB--	2.679
Age at diagnosis less than 40	2.652
Medication A10AB0-	2.476
Encounter text "type 1"	2.152
Medication A10AB04	2.066
Age at diagnosis less than 35	1.977
Medication Name "Lispro"	1.756
Age at diagnosis less than 30	1.575
Age at diagnosis less than 37	1.558
Age at diagnosis less than 25	1.551
Medication Name "insulin"	1.53
Age at diagnosis less than 22	1.512
Medication A10B---	1.446
Medication A10A---	1.379
Referral Name "endocrinology"	1.329
Age at diagnosis less than 32	1.288
Age at diagnosis less than 27	1.279
Medication A10B--- twice in 1 year	1.265