

**ESM Results 2: Additional statistical results – Alternative models**

The linear discriminant models showed similar c-statistics to the logistic regression models (T2D v MODY: 0.98 and T1D v MODY: 0.94), but had poorer cross-validated prediction errors (7.6% and 13.8%, respectively). The classification trees had higher total prediction error rates of 9% and 14%, respectively, and were limited as they could only produce a classification rather than a probability for MODY.