**Supplementary Materials belonging to “Multicenter validation of individual preoperative motor outcome prediction for deep brain stimulation in Parkinson’s disease” by Habets et al.**

1. **Literature for deciding minimal clinical important differences of outcome variables (adjusted after supplemental material from Habets et al)[1]**

A PubMed search was done using search terms: (Parkinson\*) AND (UPDRS) AND ((clinical\* relevant) OR (clinical\* significant)) AND (improve\* OR change OR difference).

Papers were selected which aim to define minimal clinical important differences of UPDRS changes. The majority of evidence found describes UPDRS III differences. Table 1 summarizes results of this search.

|  |  |  |
| --- | --- | --- |
|  | **UPDRS scores subject of research** | **Minimal clinical important difference** |
| Kostoglu[2] | III Med-OFF vs. Med-OFF + STIM | 38% |
| Horvath[3] | III | - 3.25  + 4.63 |
| Shulman[4] | III  Total | +/- 2.5 (minimal)  +/- 5.2 (moderate)  +/- 10.8 (large)  +/- 4.3 (minimal)  +/- 9.1 (moderate)  +/- 17.1 (large) |
| Schrag[5] | III  II (H&Y > 2) | +/- 5  3 |
| Makkos[6] | II + III  I + II + III  Total | -4.9  + 4.2  -6.7  +5.2  -7.1  +6.3 |

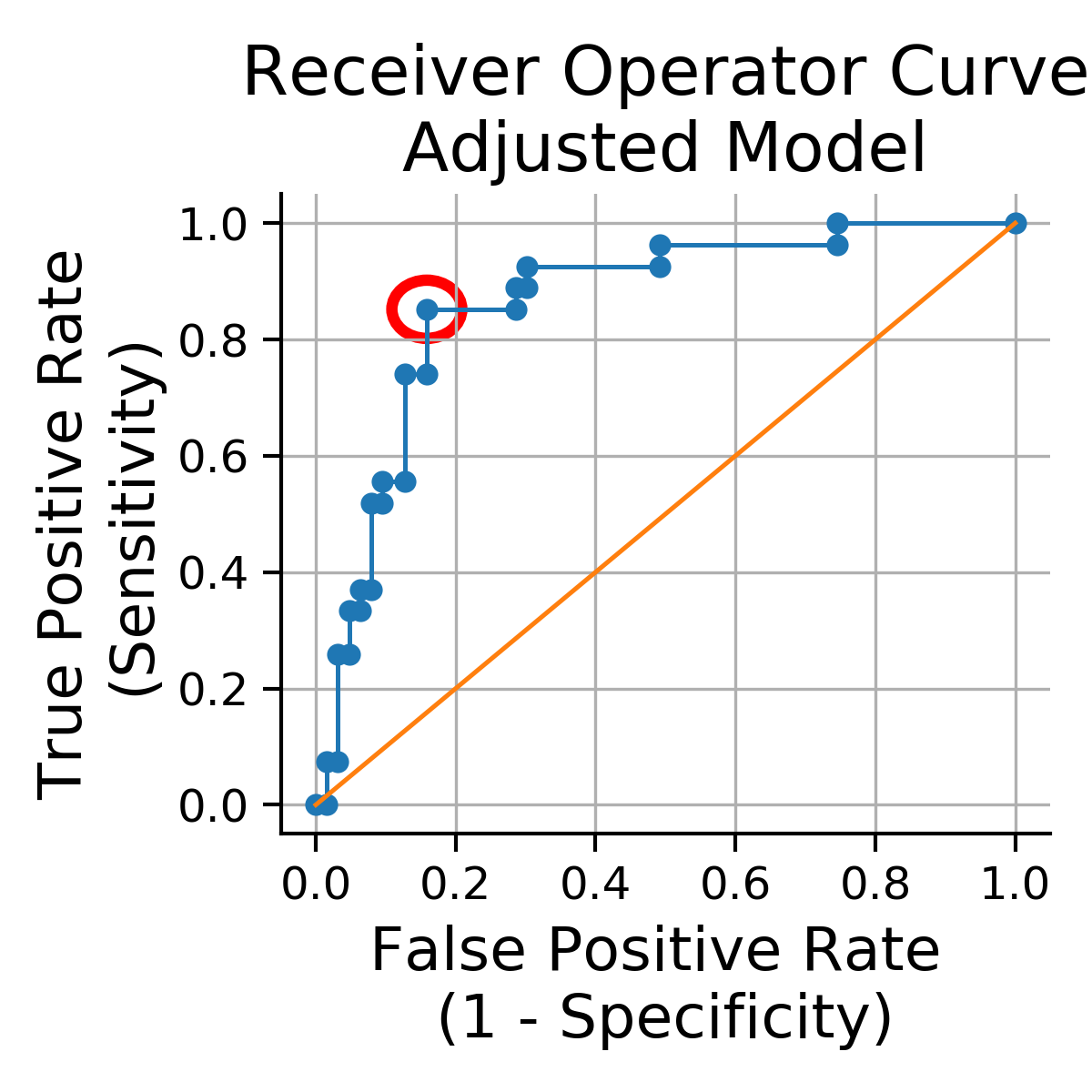
**Table 1: overview clinically significant changes in UPDRS scores reported in literature.**

For the UPDRS III score, the relevant differences range between 2.5 and 5 for the smallest relevant change. In one predictive analysis, a cut off change of 38% was used for UPDRS III off-medication. This translates to an absolute change of >10 in many patients. We averaged these findings and set the cut off for clinically relevant improvement after STN DBS at 5 points.

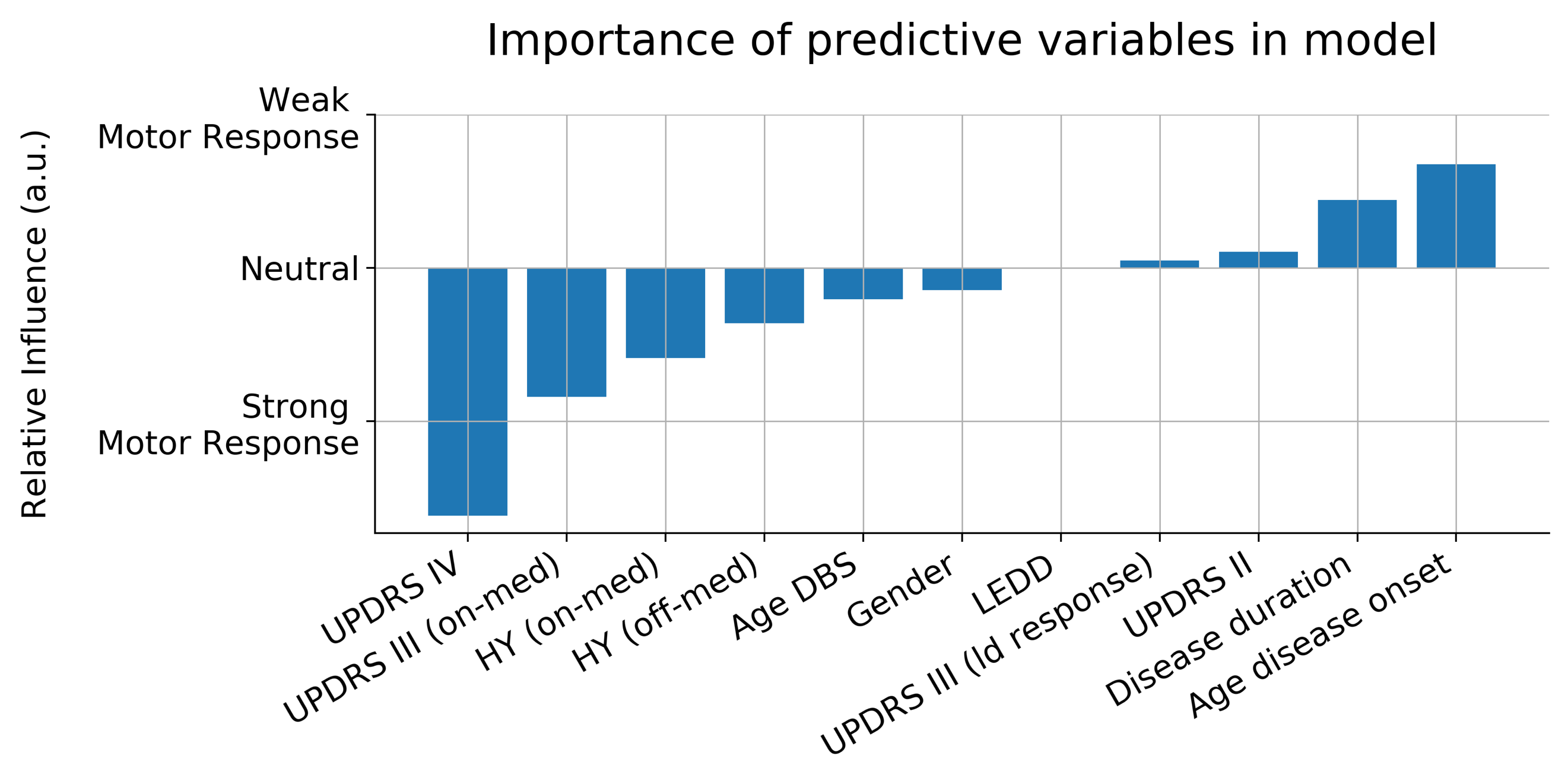
Since the UPDRS II and IV scores consist of fewer points, it is logical that the cut off values are lower. Based on the mentioned cut offs in the literature, we set the cut off for the UPDRS II and IV scores at 3 points.

1. **Adjustment prediction model**

For feasibility reasons of the data collection, we trained our prediction model after excluding neuropsychological variables and without preoperative UPDRS I as preoperative predictors. The performance of the model did not change significantly. The area under the receiver operator curve was 0.85 (standard deviation 0.06), and the optimal classification accuracy was 0.83. The weights of the individual predicting variables were also comparable to the weights of the earlier published model.[1] The adjusted model was preregistered and not optimized anymore based on the newly collected data for the validation cohort.[7]

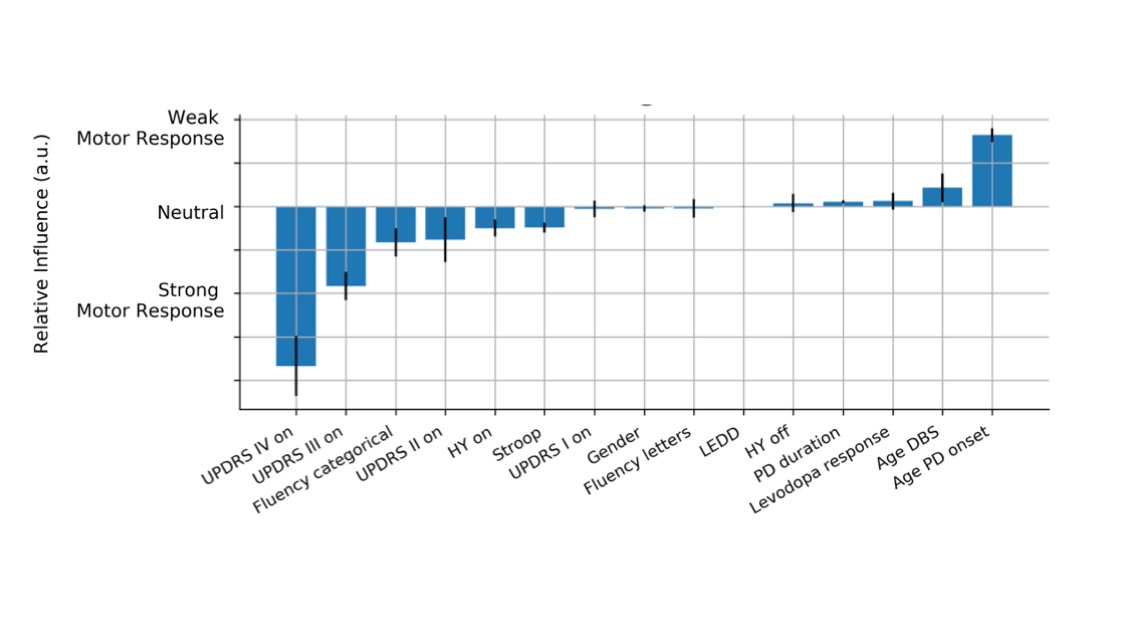


**Supplementary figure 1: Receiver operator curve of adjusted model without neuropsychological variables and UPDRS I.**



**Supplementary figure 2: Variable weights in adjusted model without neuropsychological variables and UPDRS I.**

No relevant changes were seen in the influences of the preoperative variables. Three minor weighting variables (UPDRS II on, HY off, and age DBS) changed directions (contributing to strong or weak response). These changes were not interpret as troublesome to continue the analysis with this model.

****

**Supplementary figure 3: Variable weights in original model published in pre-print (adapted from Habets et al, medrXiv 2019)**

|  |  |  |
| --- | --- | --- |
|  | **Current cohort (mean (standard deviation), unless †)** | **Development cohort (mean (standard deviation), unless †)** |
| Total number of patients collected | 334**†** | 127 **†** |
| Included patients after exclusion due to missing data | 322**†** | 90 **†** |
| Gender (female/ male) | 95/227 **†** | 37/53 **†,††** |
| Age at DBS surgery (years) | 0058 (8) | 0061 (8) |
| Disease duration (years) at DBS surgery | 0011.5 (4.7) | 0010.6 (5.1) |
| LEDD preoperative (mg) | 1525 (648) | 1197 (622) **††** |
| LEDD one year postoperative (mg) | 0731 (444) | 0665 (513) |
| UPDRS II preoperative\* | 0010.7 (6.9) | 0009.8 (6.5) |
| UPDRS II one year postoperative\* | 0008.8 (5.4) | 0009.7 (5.5) |
| UPDRS III preoperative\* | 0017.3 (8.7) | 0021.8 (12.5) **††** |
| UPDRS III preoperative, off-medication | 0044.4 (14.3) | 0039.4 (13.2) **††** |
| UPDRS III preoperative change\*\* | 0 -27.1 (12.2) | 0 -18.9 (13.4) **††** |
| UPDRS III one year postoperative\* | 0015.1 (8.5) | 0016.6 (10.0) |
| UPDRS IV preoperative\* | 0007.8 (3.3) | 0005.5 (4.0) |
| UPDRS IV one year postoperative\* | 0003.5 (2.8) | 0002.7 (2.4) |
| Preoperative HY scales\* (in number of patients) | 1: 01.9 %†  1.5: 02.2 % 2: 69.9 %  2.5: 17.1 %  3: 07.5 %  4: 01.2 %  5: 00.3 % | 1: 02.4 % †  1.5: 02.4 % 2: 16.5 %  2.5: 40.0 %  3: 28.2 %  4: 10.6 %  5: 000 % |
| Preoperative HY scales in off-medication (in number of patients) | 1: 00.0 % †  1.5: 00.3 % 2: 37.0 %  2.5: 16.5 %  3: 33.9 %  4: 09.6 %  5: 02.8 % | NA |

**Supplementary table 1: Descriptive pre- and postoperative variables.**

\*: best medical condition, on-medication for preoperative variables, and on-medication and on-stimulation for postoperative variables; \*\*: preoperative difference between off-medication vs. on-medication; †: numerical description, mean not applicable; ††: statistically significant difference between cohorts, p-value < 0.0038 (Bonferroni corrected for 13 variables). Gender compared with chi-squared test, other variables compared with Mann-Whitney-U test.

DBS: deep brain stimulation, HY: Hoehn and Yahr, LEDD: levodopa equivalent daily dosage, mg: milligram, UPDRS: Unified Parkinson Disease Rating Scale.

**Referenced work**

1. Habets JGV, Janssen MLF, Duits AA, Sijben LCJ, Mulders AEP, De Greef B, et al. Machine learning prediction of motor response after deep brain stimulation in Parkinson’s disease—proof of principle in a retrospective cohort. PeerJ. 2020 2020/11/18;8:e10317.

2. Kostoglou K, Michmizos KP, Stathis P, Sakas D, Nikita KS, Mitsis GD. Classification and Prediction of Clinical Improvement in Deep Brain Stimulation From Intraoperative Microelectrode Recordings. IEEE transactions on bio-medical engineering. 2017 May;64(5):1123-30.

3. Horvath K, Aschermann Z, Acs P, Deli G, Janszky J, Komoly S, et al. Minimal clinically important difference on the Motor Examination part of MDS-UPDRS. Parkinsonism & related disorders. 2015 Dec;21(12):1421-6.

4. Shulman LM, Gruber-Baldini AL, Anderson KE, Fishman PS, Reich SG, Weiner WJ. The clinically important difference on the unified Parkinson's disease rating scale. Archives of neurology. 2010 Jan;67(1):64-70.

5. Schrag A, Sampaio C, Counsell N, Poewe W. Minimal clinically important change on the unified Parkinson's disease rating scale. Movement disorders : official journal of the Movement Disorder Society. 2006 Aug;21(8):1200-7.

6. Makkos A, Kovacs M, Aschermann Z, Harmat M, Janszky J, Karadi K, et al. Are the MDS-UPDRS-Based Composite Scores Clinically Applicable? Movement disorders : official journal of the Movement Disorder Society. 2018 May;33(5):835-39.

7. ClinicalTrials.gov. NCT04093908: Prediction of STN DBS Motor Response in PD (DBS-PREDICT) clinicaltrials.gov2019.