

Supplementary Table 1: PPMI Clinical Features

Feature	Description	Dataset
Age	ENROLLDT - BIRTHDT	RANDOM (Consent and Enrollment)
Gender	GENDER = 2 is Male GENDER = 0 or 1 is Female	RANDOM (Consent and Enrollment)
Race	RAINDALS, RAASIAN, RABlack, RAHAWOPI, RAWHITE, RANOS Other = RAINDALs, RAHAWOPI, RANOS, or more than one race specified	SCREEN
Family History of PD	BIOMOMPD, BIODADPD, FULSIBPD, HALFSIBPD, MAGPARPD, PAGPARPD, MATAUPD, PATAUPD, KIDSPD Subject has any family history of PD if any one or more of the above variables = '1'. Subject has 1st-degree family history of PD if any one or more of the following variables = '1': BIOMOMPD, BIODADPD, FULSIBPD, HALFSIBPD, KIDSPD	FAMHXPd
MDS-UPDRS Part I	NP1COG, NP1HALL, NP1DPRS, NP1ANXS, NP1APAT, NP1DDS, NP1SLPN, NP1SLPD, NP1PAIN, NP1URIN, NP1CNST, NP1LTHD, NP1FATG Part I Score = sum of these 13 variables 6 in one file, 7 in other. Ignore NPSOURC in both	NUPDRS1, NUPDRS1P
MDS-UPDRS Part II	NP2SPCH, NP2SALV, NP2SWAL, NP2EAT, NP2DRES, NP2HYGN, NP2HWRT, NP2HOBb, NP2TURN, NP2TRMR, NP2RISE, NP2WALK, NP2FREZ Part II Score = sum of these 13 variables	NUPDRS2P
MDS-UPDRS Part III	NP3SPCH, NP3FACXP, NP3RIGN, NP3RIGRU, NP3RIGLU, N3RIGRL, NP3RIGLL, NP3FTAPR, NP3FTAPL, NP3HMOVR, NP3HMOVL, NP3PRSPR, NP3PRSPL, NP3TTAPR, NP3TTAPL,	NUPDRS3

	<p>NP3LGAGR, NP3LGAGL, NP3RISNG, NP3GAIT, NP3FRZGT, NP3PSTBL, NP3POSTR, NP3BRADY, NP3PTRMR, NP3PTRML, NP3KTRMR, NP3KTRML, NP3RTARU, P3RTALU, NP3RTARL, NP3RTALL, NP3RTALJ, NP3RTCON</p> <p>Part III Score = sum of these 33 variables</p>	
MDS-UPDRS Part IV	<p>NP4WDYSK, NP4DYSKI, NP4OFF, NP4FLCTI, NP4FLCTX, NP4DYSTN</p> <p>Part IV Score = sum of these 6 variables</p>	NUPDRS4
MDS-UPDRS Total Score	Sum of MDS-UPDRS Parts I, II, III	NUPDRS1, NUPDRS1P, NUPDRS2P, NUPDRS3
TD / PIGD Classification	<p>First, calculate Tremor and PIGD scores:</p> <p>Tremor score = Mean of the following variables: NP2TRMR, NP3PTRMR, NP3PTRML, NP3KTRMR, NP3KTRML, NP3RTARU, NP3RTALU, NP3RTARL, NP3RTALL, NP3RTALJ, NP3RTCON</p> <p>PIGD score = Mean of the following variables: NP2WALK, NP2FREZ, NP3GAIT, NP3FRZGT, NP3PSTBL</p> <p>Then calculate ratio = Tremor score / PIGD score.</p> <p>If ratio ≥ 1.15, OR if PIGD score = 0 and Tremor score > 0, then subject is TD. If the ratio ≤ 0.9 then the subject is PIGD. If ratio > 0.9 and < 1.15, OR if Tremor score and PIGD score = 0, then subject is Indeterminate.</p>	NUPDRS2P, NUPDRS3
Benton Judgment of Line Orientation Score	<p>Sum of BJLOT1 – BJLOT30</p> <p>Either odd BJLOTs filled or even ones. Just sum overall</p>	LINEORNT
Epworth Sleepiness Scale	<p>Sum of ESS1 - ESS8</p> <p>Subjects with ESS < 10 are “Not Sleepy”</p> <p>Subjects with ESS ≥ 10 are “Sleepy”</p>	EPWORTH
GDS	Add 1 point for each response of “No” (0) to any of	GDSSHORT

	<p>the following variables: GDSSATIS, GDSGSPR, GDSHAPPY, GDSALIVE, GDSENRGY</p> <p>Add 1 point for each response of “Yes” (1) to any of the following variables: GDSDROPD, GDSEMPY, GDSBORED, GDSAFRAD, GDSHLPLS, GDSHOME, GDSMEMRY, GDSWRTLS, GDSHOPLS, GDSBETER</p> <p>Subjects with GDS ≥ 5 are “Depressed” Subject with GDS < 5 are “Not Depressed”</p>	
Hopkins Verbal Learning Test	<p>HVLT Immediate/Total Recall: Sum of HVLTRT1 - HVLTRT3</p> <p>HVLT Discrimination Recognition: HVLTRC - (HVLTFPRL + HVLTFPUN)</p> <p>HVLT Retention: HVLTRDLY / max(HVLTRT2, HVLTRT3)</p>	HVLT
Letter Number Sequencing (LNS)	<p>Sum of LNS1A – LNS7C</p> <p>Can also use LNSTOTRAW</p>	LNSPD
MOCA Total Score	<p>Unadjusted Score = sum of MCAALTTM, MCACUBE, MCACLCKC, MCACLCKN, MCACLCKH, MCALION, MCARHINO, MCACAMEL, MCAFDS, MCABDS, MCAVIGIL, MCASER7, MCASNTNC, MCAVF, MCAABSTR, MCAREC1, MCAREC2, MCAREC3, MCAREC4, MCAREC5, MCADATE, MCAMONTH, MCAYR, MCADAY, MCAPLACE, MCACITY</p> <p>Can use MCATOT as well (Note: Ignore MCAVFNUM column somewhere between these in MOCA dataset)</p> <p>If EDUCYRS ≤ 12 and Unadjusted Score < 30, add 1 more point to the score. If EDUCYRS > 12, do not add any more points to the score.</p>	MOCA, SOCIOECO
QUIP	<p>For Sections A - D, add 1 point if either question has a response of “Yes” (1): Section A: CNTRLGMB, TMGAMBLE Section B: CNTRLSEX, TMSEX</p>	QUIPCS

	<p>Section C: CNTRLBUY, TMBUY Section D: CNTRLEAT, TMEAT For Section E, add 1 point for each response of "Yes" (1): TMTORACT, TMTMTACT, TMTRWD</p>	
REM Sleep Behavior Screening Questionnaire (RBDSQ)	<p>Add 1 point for each response of "Yes" (1) to any of the following variables: DRMVIVID, DRMAGRAC, DRMNOCTB, SLPLMBMV, SLPINJUR, DRMVERBL, DRMFIGHT, DRMUMV, DRMOBJFL, MVAWAKEN, DRMREMREM, SLPDSTRB</p> <p>Add 1 point if any of the following variables has a response of "Yes" (1): STROKE, HETRA, PARKISM, RLS, NARCLPSY, DEPRS, EPILEPSY, RNINFM, CNSOTH</p> <p>If any of the previous variables are missing, then RBD score is missing.</p> <p>Subjects with score ≥ 5 are RBD Positive and Subjects with score < 5 are RBD Negative</p>	REMSLEEP
SCOPA-AUT Total and Subscores	<p>SCAU1 - SCAU25</p> <p>For questions 1-21 (SCAU1 - SCAU21), add 3 points for each response of "9." Otherwise, add the number of points in response.</p> <p>For questions 22-25 (SCAU22 - SCAU25), add 0 points for each response of "9." Otherwise, add the number of points in response.</p> <p>Subscores: Gastrointestinal = questions 1-7 Urinary = questions 8-13 Cardiovascular = questions 14-16 Thermoregulatory = questions 17, 18, 20, 21 Pupillomotor = question 19 Sexual dysfunction = questions 22, 23, 24, 25</p>	SCOPA
Semantic Fluency (SFT)	Sum of VLTANIM, VLTVEG, VLTRUIT	SFT
State Trait Anxiety Index (STAI)	<p>STAIAD1 - STAIAD40</p> <p>Add values for the following questions: 3, 4, 6, 7, 9, 12, 13, 14, 17, 18, 22, 24, 25, 28, 29, 31, 32, 35, 37,</p>	STAI

	<p>38, 40</p> <p>Use reverse scoring for the remaining questions and add to the first score (e.g., if value = 1, add 4 points to score; if value = 2, add 3 points to score, etc.).</p>	
UPSIT	<p>Raw Score = Sum of UPSITBK1 - UPSITBK4</p> <p>Age- and sex-adjusted categories: 0 – 18 = Anosmia (Males and Females) 19 – 33 = Hyposmia (Males) 19 – 34 = Hyposmia (Females) 34 – 40 = Normosmia (Males) 35 – 40 = Normosmia (Females)</p>	UPSIT
Test-based Mild Cognitive Impairment (MCI)	<p>DVT_TOTAL_RECALL, DVT_RECOG_DISC_INDEX, DVS_JLO_MSSAE, DVS_LNS, DVT_SFTANIM, DVT_SDM</p> <p>Subject has MCI if any 2 or more of the following cognitive tests are >1.5 SD below the standardized mean:</p> <ul style="list-style-type: none"> • HVLT Total Recall (DVT_TOTAL_RECALL \leq 35) • HVLT Recognition Discrimination (DVT_RECOG_DISC_INDEX \leq 35) • Benton Judgment of Line Orientation (DVS_JLO_MSSAE \leq 6) • Letter Number Sequencing (DVS_LNS \leq 6) • Semantic Fluency Test (DVT_SFTANIM \leq 35) • Symbol Digit Modalities (DVT_SDM \leq 35) 	<p>HVLT, LINEORNT, LNSPD, SFT, SDM</p>

* Referenced from McGill, M. P., Sullivan, G. M., & Nemeroff, C. B. (2011). Animal models of depression and the mismatch hypothesis of mood disorders. *Progress in Neurobiology*, 93*(3), 213-239.
doi:10.1016/j.pneurobio.2010.10.003