## Supplementary Table 1: PPMI Clinical Features

| Feature | Description | Dataset |
|---|---|---|
| Age | ENROLLDT - BIRTHDT | RANDOM (Consent and Enrollment) |
| Gender | GENDER = 2 is Male<br>GENDER = 0 or 1 is Female | RANDOM (Consent and Enrollment) |
| Race | RAINDALS, RAASIAN, RABLACK, RAHAWOPI, RAWHITE, RANOS<br><br>Other = RAINDALS, RAHAWOPI, RANOS, or more than one race specified | SCREEN |
| Family History of PD | BIOMOMPD, BIODADPD, FULSIBPD, HALFSIBPD, MAGPARPD, PAGPARPD, MATAUPD, PATAUPD, KIDSPD<br><br>Subject has any family history of PD if any one or more of the above variables = '1'.<br>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<br><br>Part I Score = sum of these 13 variables<br><br>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<br>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 |

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

| | | |
|---|---|---|
| | the following variables:<br>GDSSATIS, GDSGSPIR, GDSHAPPY, GDSALIVE, GDSENRGY<br><br>Add 1 point for each response of “Yes” (1) to any of the following variables:<br>GDSDROPD, GDSEMPTY, GDSBORED, GDSAFRAD, GDSHLPLS, GDSHOME, GDSMEMRY, GDSWRTLS, GDSHOPLS, GDSBETER<br><br>Subjects with GDS ≥ 5 are “Depressed”<br>Subject with GDS < 5 are “Not Depressed” | |
| Hopkins Verbal Learning Test | HVLT Immediate/Total Recall: Sum of HVLTRT1 - HVLTRT3<br><br>HVLT Discrimination Recognition: HVLTREC - (HVLTFPRL + HVLTFPUN)<br><br>HVLT Retention: HVLTRDLY / max(HVLTRT2, HVLTRT3) | HVLT |
| Letter Number Sequencing (LNS) | Sum of LNS1A – LNS7C<br><br>Can also use LNSTOTRAW | LNSPD |
| MOCA Total Score | 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<br><br>Can use MCATOT as well (Note: Ignore MCAVFNUM column somewhere between these in MOCA dataset)<br><br>If EDUCYRS ≤ 12 and Unadjusted Score < 30, add 1 more point to the score.<br>If EDUCYRS > 12, do not add any more points to the score. | MOCA, SOCIOECO |
| QUIP | For Sections A - D, add 1 point if either question has a response of “Yes” (1):<br>Section A: CNTRLGMB,TMGAMBLE<br>Section B: CNTRLSEX, TMSEX | QUIPCS |

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

| | | |
|---|---|---|
| | 38, 40<br><br>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.). | |
| UPSIT | Raw Score = Sum of UPSITBK1 - UPSITBK4<br><br>Age- and sex-adjusted categories:<br>0 – 18 = Anosmia (Males and Females)<br>19 – 33 = Hyposmia (Males)<br>19 – 34 = Hyposmia (Females)<br>34 – 40 = Normosmia (Males)<br>35 – 40 = Normosmia (Females) | UPSIT |
| Test-based Mild Cognitive Impairment (MCI) | DVT_TOTAL_RECALL, DVT_RECOG_DISC_INDEX, DVS_JLO_MSSAE, DVS_LNS, DVT_SFTANIM, DVT_SDM<br><br>Subject has MCI if any 2 or more of the following cognitive tests are >1.5 SD below the standardized mean:<br><br>• HVLT Total Recall (DVT_TOTAL_RECALL ≤ 35)<br>• HVLT Recognition Discrimination (DVT_RECOG_DISC_INDEX ≤ 35)<br>• Benton Judgment of Line Orientation (DVS_JLO_MSSAE ≤ 6)<br>• Letter Number Sequencing (DVS_LNS ≤ 6)<br>• Semantic Fluency Test (DVT_SFTANIM ≤ 35)<br>• Symbol Digit Modalities (DVT_SDM ≤ 35) | HVLT,<br>LINEORNT,<br>LNSPD,<br>SFT,<br>SDM |