MOST BASELINE SCREENING AND ENROLLMENT

DATASET DESCRIPTION

PUBLIC DATA RELEASE, OCTOBER 2009

This document describes the MOST baseline clinical dataset and data issues relevant to analysts. If you are unfamiliar with the data, it may be useful to begin by reviewing the annotated data collection forms to look for variables of interest.

FORMATS

SAS Format Library

The SAS format library (FORMATS.SAS7BDAT) contains all the formats used for the dataset.

DATASET

Baseline Telephone Screening and Enrollment Visit (V0ENROLL)

The baseline screening and enrollment dataset (V0ENROLL.SAS7BDAT) includes all enrolled participants (N=3026) that completed the baseline clinic visit. Data collected from five points of contact are included: a telephone screening interview, two self-administered questionnaires (one done at home and the other done at the clinic), a clinic interview, and clinic visit.

Variables are sorted in the order of data collection ("creation order") – as if following the participant from the first telephone screening question to the last measurement at the clinic. Page numbers referenced below correspond to the bracketed page numbers located at the bottom of each page of the annotated data collection forms.

Telephone Screening Interview (Variables with 'TS' prefix)

The telephone screening interview was conducted to assess eligibility for enrollment in the study. The dataset includes the following components:

p3	Demographics (age, sex, ethnicity, racial background)[1]
p4	Weight; Knee pain, aching, and stiffness; History of knee surgery
p8	History of rheumatoid arthritis
p9	Joint stiffness, nodules, and swelling

Notes:
[1] Ethnicity and racial background questions in the telephone screening interview were repeated in the SAQ-Home. The released variables ETHNICITY and RACE are derived from the SAQ-Home (see next page).

Self-Administered Questionnaire (SAQ) – Home (Variables with "V0" prefix)

The SAQ-Home was mailed to participants after the telephone interview. Participants were instructed to complete the questionnaire at home prior to departure for the clinic visit. The dataset includes the following components:

p11	Demographics (ethnicity, racial background, marital and living status, education)
p12	Self-reported height[1] and weight history[2]
pp13-15	Joint pain, aching, and stiffness (homunculus diagram)
p16	Back pain and function[3]
p17	History of arthritis; Arthritis medications
p18	Family history of arthritis
pp19-21	Modified Charlson Comorbidity Index – Katz Questionnaire Adaptation[4]
p22	Fracture history
p23	Tobacco use
p24	Employment
pp25-28	Work history
pp29-30	Late Life Function and Disability Instrument (LLFDI) – Modified Disability Component[5,6]
pp31-32	Modified SF-12 U.S. version 1.0[7]
pp33-34	CES-D (Depression scale) – Long version

Notes:

[1] In cases of self-reported height at 25-years-old (V0HTFT, V0HTIN, V0_HT25) greater than 10cm from standing height measured in the clinic (V0HT), the self-reported values and are coded as missing.

[2] In cases of self-reported weight at 25-years-old (V0WT25) greater than self-reported maximum weight (V0WTMAX), the maximum weight value is coded as missing.

[3] Responses to two back pain questions, "How many days did you stay in bed because of your back?" (V0BDDAY) and "How many days did you limit your activities because of your back?" (V0BPLAD) are numbers of days designed to sum to no more than 30. Analysts should be aware that some participants misunderstood and provided responses that sum to greater than 30.

[4] Charlson Comorbidity Index – Katz Questionnaire Adaptation. This measurement was modified to include the option of "Don't know" to accommodate participants unable to answer "Yes" or "No" to any question with certainty. Responses of "Don't know" were scored with a zero value (see the document: Calculated Variable Descriptions and SAS Code).

[5] Late-Life Function and Disability Instrument (LLFDI) – Modified Disability Component. This measurement was shortened to 12 of the authors' 16 disability subscale questions and included only the extent of limitation performing activities ("To what extent do you feel limited in …?"). Frequency performing activities ("How often do you ..?") was not collected and the last of 5 options – [Not at all] [A little] [Somewhat] [A lot] [Completely, cannot do] – was modified with "cannot do" dropped.

[6] Scoring of this measurement was modified to handle missing values in a way that is consistent with how MOST analysts scored the WOMAC™ and SF-12 measures. For more information, refer to the document Calculated Variable Descriptions and SAS Code, and also see Jette AM, et al. Late life function and disability instrument: I. Development and evaluation of the disability component. J Gerontol A Biol Sci Med Sci. 2002 Apr;57(4):M209-16. PMID: 11909885.

[7] SF-12 U.S. version 1.0. This measurement was modified to include a "Don't know" option on questions concerning the extent to which physical health limited work or other regular daily activities in the past 30 days (questions #29 – 32 on annotated forms p26).

Self-Administered Questionnaire (SAQ) – Clinic (Variables with "V0" prefix)

The SAQ-Clinic was administered during the baseline clinic visit. The dataset includes the following components:

–	Modified WOMAC™ knee pain and stiffness[1,2]
–	Modified WOMAC™ degree of difficulty performing daily activities[1,2]
pp39-40	Knee pain visual analog scale (VAS)
p41	Modified KOOS Function in sports and recreational activities subscale[3]
–	Modified WOMAC™ Osteoarthritis Index – Hip pain[1,2]

Notes:

[1] WOMAC Osteoarthritis Index™ Likert version. This measurement was modified to include a "don't do" option for participants who cannot rate severity of pain during a particular activity because they avoid or are unable to do that activity.

[2] The WOMAC™ instrument is not displayed in the annotated forms because it is trademark and copyright protected. Information can be obtained by contacting the author, Nicholas Bellamy, via the WOMAC™ 3.1 Index website (http://www.auscan.org/womac).

[3] KOOS Function in Sports and Recreational Activities Subscale, Likert version. This measurement was modified to include a "don't do" option for participants who cannot rate severity of pain during a particular activity because they avoid or are unable to do that activity.

Clinic Interview (Variables with "V0" prefix)

The clinic interview is an interviewer-administered questionnaire conducted during the baseline clinic visit. The dataset includes the following components:

–	Modified Physical Activity Scale for the Elderly (PASE) [1]
pp43-44	Physical activity – Climbing flights of stairs[2]
pp48-49	Knee pain, aching, and stiffness
p50	Knee buckling
pp51-56	History of knee injury and surgery[3]
pp57-58	Hip pain, aching, and stiffness
p59	History of hip replacement surgery[3]
pp-64	Medication Inventory Form (MIF)[4,5]

Notes:

[1] The PASE© measurement was modified to include a possible response of "Don't know/ Refused" in 7 of the 12 elements that contribute to the total score. All such responses are converted to missing in the calculation of the total score. For a description of the PASE calculation, see the document: Calculated Variable Descriptions and SAS Code. The measurement is not displayed in the annotated forms because it is copyright protected. Information about the measurement can be obtained through the PASE© product information website of New England Research Institutes (NERI) (http://www.neriscience.com/web/MultiPiecePage.asp_Q_PageID_E_253_A_PageName_E_ProductsResearchPhysicalActiv).

[2] The stair climbing question is not part of the PASE measurement and does not contribute to the summary score.

[3] Knee and hip replacement variables (V0R_TKR, V0L_TKR, V0R_THR, V0L_THR) are derived from self-report and radiographic adjudication.

[4] Participants were asked to bring all medications taken in the last 30 days (prescription, non-prescription, vitamins, supplements) to be recorded.

[5] Medication ingredients, coded by the UCSF MIF group using the Iowa Drug Information Service (IDIS) dictionary, are released in Yes/No format, meaning used or not used during the last 30 days. Formulation code, duration, and frequency are not released. For further information about IDIS, see Pahor M, Chrischilles EA, and Guralnik, JM. Drug data coding and analysis in epidemiologic studies. Eur J Epidemiol. 1994 Aug;10(4):405-11.

Clinic Visit (Variables with "V0" prefix)

Selected exams were conducted at the baseline clinic visit. The dataset includes the following components:

p68	Blood pressure
p69	Standing height; Weight
p72	20-meter walk[1]
pp73-74	Chair stands[2]
pp75-79	Isokinetic strength (Cybex)[3]
p80	Leg length; Knee height
pp82	Leg dominance[4]
p84	Hand examination
p98	Knee x-ray[5]
pp95-99	OrthOne 1.0T knee MRI[5]

Notes:

[1] The two clinical sites interpreted the protocol for measuring the stop time in slightly different ways; see page 4, section 5 of the 20-meter Walk Operations Manual.

[2] Trial #2 of the repeated chair stands protocol was discontinued in March 2004; participants with subsequent visit dates are missing values for the variables V0TR2, V0CTIME2, and V0NUM2.

[3] Peak torque measurements obtained from the isokinetic strength exam are not gravity-corrected.

[4] The leg dominance question (V0FOOT) was added in July 2003; participants with prior clinic visit dates are missing a value for that variable.

[5] Some participants returned to the clinic to repeat x-ray and MRI exams when image quality was not adequate for reading. Repeat data is not included in the dataset.