Author: Martin Wegman
Date: February 11, 2016

Re: Replication data for “Differences in drug relapse rates between compulsory drug detention centre-based rehabilitation and voluntary drug treatment including methadone in Malaysia”

This document describes the de-identified participant-level data contained in the SAS transport file upload_data.cpt. Note that the data is minimally processed; entry and reporting errors are contained within it. Most of these are addressed in the first SAS program score.sas.The dataset is in wide format, containing one row per participant, where participants are identified by a random id: client_id. The study arm for each participant is represented by site.

Variables beginning with “T” indicate time variables, in units of days. T0 – T15 indicate measurement times for the corresponding outcome indicator variables Y0_xxx - Y15_xxx [where xxx indicates the outcome type: opioid (opi), methadone (mtd), amphetamine (amp), buprenorphine (bup), benzodiazepine (bzd), cannabis (thc), HIV (hiv)]. The numeric digit(s) in the variable name roughly correspond to the month of measurement. For the outcome indicators, 0 indicates absence of the outcome and 1 indicates presence; see the manuscript for more detail on ascertainment methods. T_int variable measures the timing of the baseline interview.

For site = CDDC, T0 – T15 and T_int variables are days from date of release. For site = VTC, T0 – T15 and T_int variables are days from starting MMT. T_mmt indicates days since a fixed arbitrary date prior to the study and the start of MMT (for VTC participants). T_dor indicates days since a fixed arbitrary date prior to the study and the day of release (for CDDC participants). T_ dis indicates days between the start of MMT and the day of discharge from the VTC (for VTC participants).

In addition to time and outcome variables, the series of “F” variables (i.e. F0 – F15) are missing flags, indicating whether the corresponding outcome (i.e. Y0_xxx – Y15_xxx) is lost to follow-up (F=1 or F = 2), by design (F=.) or not missing (F=0). The variable Num_urines_recorded is the sum of measurements where F=0 over all F for each participant.

RA is a de-identified indicator for the primary RA responsible for following-up with the given participant. Version indicates the questionnaire version used, the old version used the SF-12 instead of the WHO-QOL BREF, among a few additional minor changes.

The remaining variable names and values match directly from the companion surveys which have also been uploaded.