

## ATC CODING FOR MEDICATION DATA

SAS DATASET NAME: vr\_meds\_ex01\_3\_0242d

PROTOCOL NAME: vr\_meds\_ex01\_3\_0242d\_protocol.doc

Medications were coded using the World Health Organization's Anatomical Therapeutic Chemical (ATC) classifications. Information about the ATC classification system can be found at <http://www.whooc.no/atcddd/>. Briefly,

The numeric ATC codes assigned to a drug can be broken down into five parts. For example in the code N02BE01:

The first character (N) represents the main anatomical group.

In this example N = Nervous System.

Characters two and three (02) represent the therapeutic subgroup.

In this example N02 = Analgesics.

Character four (B) represents the pharmacological subgroup.

In this example N02B = Other analgesics and antipyretics.

Character five (E) represents the chemical subgroup.

In this example N02BE = Anilides.

Characters six and seven (01) represent the chemical substance.

In this example N02BE01 = Paracetamol (acetaminophen).

If a medication consists of multiple compounds, all unique compounds were coded leaving up to four codes per medication in the coded database (variable names = atc\_cod1,...,atc\_cod4). There were a few exceptions including oral contraceptives, vitamins/supplements and other multi-compound medications not assigned to one code per compound due to the nature of the medication. Looking at the contraceptive example below, if we use the first two codes instead of one of the second two codes we lose the info about fixed or sequential doses and the fact that it is specifically an oral contraceptive:

(1) G03AC03	LEVONORGESTREL
(2) G03CA57	CONJUGATED ESTROGEN
(3) G03AA07	LEVONORGESTREL AND ESTROGEN FIXED
(4) G03AB03	LEVONORGESTREL AND ESTROGEN SEQUENTIAL

The Levodopa example below shows an example of a medication that does not have one code per compound because one of those compounds could not be found in the ATC classification system:

N04BA01	LEVODOPA
N04BA02	LEVODOPA AND DECARBOXYLASE INHIBITOR
N04BA03	LEVODOPA AND DECARBOXYLASE INHIBITOR COMT INHIBITOR

Medication data were collected during the participant physical exam using free text fields for medication name and strength. Note that the same drug may be recorded in more than

one record for an individual participant. This is because the participant took the medication at more than one strength and/or dose.

Medication coding was carried out in two steps. (1) Medication names that matched those in the coding dictionary were coded automatically. (2) If the medication name did not match (e.g. because of misspelling) or if there were more than one use for the drug, a code was assigned manually with physician input based on the use of the drug. An EXCEL spreadsheet ‘meds1\_3s\_coding\_manual\_supplement.xls’ exists that contains a cross reference for drugs that have more than one code in the ATC dictionary. For example acetylsalicylic acid can be coded in 3 ways, depending on its use:

MEDNAME	OTHER_CODE	ATC_CODE	SYSTEM	THER_GP	PHRM_GP	CHEM_GP
ACETYLSALICYLIC ACID	B01AC06	A01AD05	A	A01	A01A	A01AD
ACETYLSALICYLIC ACID	N02BA01	A01AD05	A	A01	A01A	A01AD
ACETYLSALICYLIC ACID	A01AD05	B01AC06	B	B01	B01A	B01AC
ACETYLSALICYLIC ACID	N02BA01	B01AC06	B	B01	B01A	B01AC
ACETYLSALICYLIC ACID	A01AD05	N02BA01	N	N02	N02B	N02BA
ACETYLSALICYLIC ACID	B01AC06	N02BA01	N	N02	N02B	N02BA

If one is searching the medication database for participants taking an analgesic (THER\_GP = N02), one may consider searching participants taking A01AD05 or B01AC06 as well. This EXCEL file is sorted by atc\_code. The ATC codes that correspond to that chemical name are contained in the column ‘other\_code’.

The ATC classification database contains only generic drug names. Brand names have been added and continue to be added to the FHS coding dictionary as they arise. During the process, the free text for medication name that is in the exam data sets was not changed from that which was originally entered.

Note: This data set was coded between July 2008 and January 2009. The ATC codes available during that time were used.

## Variables:

As discussed above, the ATC codes assigned to a drug can be broken down into five parts. For example in the code N02BE01:

The first character (N) represents the main anatomical group.

In this example N = Nervous System.

Variables System1 – System4 are descriptions/labels for the main anatomical group. Characters two and three (02) represent the therapeutic subgroup.

In this example N02 = Analgesics.

Variables Ther\_gp1 – Ther\_gp4 are descriptions/labels for the therapeutic group. Character four (B) represents the pharmacological subgroup.

In this example N02B = Other analgesics and antipyretics.

Variables Phrm\_gp1 – Phrm\_gp4 are descriptions/labels for the pharmacological subgroup.

Character five (E) represents the chemical subgroup.

In this example N02BE = Anilides.

Variables Chem\_gp1 – Chem\_gp4 are descriptions/labels for the chemical subgroup. Characters six and seven (01) represent the chemical substance.

In this example N02BE01 = Paracetamol (acetaminophen).

Variables Chem\_nm1 – Chem\_nm4 are the chemical substance name.

If specific chemical name/origin was not known, that record contains a partial code (i.e., less than 7 characters). For example, “ANTI-HISTAMINE” was coded R06A for systemic use or D04AA for topical use. “ANTACID” was coded A02A since source was unknown.