NHLBI grant: Ancillary to HCHS/SOL: Cardiometabolic Outcomes in Multi-ethnic Physical Activity & Sedentary Behavior Study (R01 HL136266; Y Mossavar-Rahmani, RC Kaplan, & V Ramachandran)

HCHS/SOL COMPASS Ancillary Derived Variable Dictionary

October 2020
(Documentation V1.1)
INV1 data

Prepared by
HCHS/SOL Coordinating Center
Collaborative Studies Coordinating Center
UNC Department of Biostatistics

Franklyn Gonzalez II
Daniela Sotres-Alvarez

COMPASS Derived Variable Dictionary INV1

Table of Contents

Updates to COMPASS datasets

DATA Version	Date	What changed	Documentation
INV1	4/30/2020	1st Data Distribution to Investigators	Ver. 1.0
INV1	10/12/2020	2nd Data Distribution to Investigators. There are NO changes to the definitions of any of the variables. The only change is the update in section 6 to increase the number of adherent participants by four.	Ver. 1.1

I. Participant Derived Variable File

1. DESIGN

For the following sample weight variable(s): Please see HCHS/SOL Manuscript “Sample design and cohort selection in the Hispanic Community Health Study/Study of Latinos” (Annals of Epidemiology, Volume 20, Issue 8, August 2010, Pages 629-641) for information on how the sample weights were created. Also see the HCHS/SOL Analyses Methods at Baseline (Sept. 2016) and Analyses Methods for Visit 2 (May 2020) documentation that describes the proper use of sample weights in statistical analyses.

Ins COMPASS, the sampling weights are calculated based on visit 2 non-response adjusted sampling weights and the nonresponse rates by gender and 6-level age groups for COMPASS. The sampling weights for COMPASS are then trimmed (to handle extreme values), calibrated (to the US 2010 Census population in the target population census tracks), and normalized (so that the sum of the sampling weights adds to the total sample size).

1.1. WEIGHT_EXPANDED_COMPASS (Expanded Census 2010 Calibrated, Trimmed, Nonresponse Adjusted Weights)

The COMPASS_WEIGHT_EXPANDED sample weight is calibrated, trimmed, non-response adjusted (by gender and age groups) reciprocal of a participant’s probability of selection into the HCHS/COMPASS Ancillary Study. It was calibrated using first the Hispanic background distribution and then the 12-category age group/gender distribution using the 2010 Decennial Census.

NOTE that for San Diego participants, the WEIGHT_EXPANDED_COMPASS is identical to WEIGHT_ EXPANDED _CASAS.

1.2. WEIGHT_NORM_CENTER_COMPASS (Normalized to field center sample size, Census 2010 Calibrated, Trimmed, NonResponse Adjusted Weights)

This sampling weight is the expanded sampling weight normalized to COMPASS sample size within each field center. The sum of the weights is 4,656.

NOTE that for San Diego participants, the WEIGHT_NORM_CENTER_COMPASS is identical to WEIGHT_NORM_CENTER_CASAS.

1.3. WEIGHT_NORM_OVERALL_COMPASS (Normalized to overall sample size, Census 2010 Calibrated, Trimmed, NonResponse Adjusted Weights)

This sampling weight is the expanded sampling weight normalized to the overall COMPASS sample size. The sum of the weights is 4,656.

2. ADMINISTRATIVE

2.1. CLINDATE_COMPASS (Date of COMPASS clinic visit)

The date of the participant's COMPASS clinic visit is derived from the COMPASS Informed Consent Tracking form (ICT) and from the CASAS Eligibility/Participation Checklist form (CEL).

CLINDATE_COMPASS= ICT0a (COMPASS) or CEL2a (CASAS)

Source variable(s):
ICT0A. Date of COMPASS Appointment
CEL2A. Date of CASAS Appointment

3. SOCIO-DEMOGRAPHIC

3.1. AGE_COMPASS (Age in years at COMPASS or CASAS clinic visit)

The age of the participant in years determined from the participant's date of birth (from HCHS/SOL Visit 2) and the COMPASS or CASAS clinic visit date.

Source variable(s):
DOB_V2. Date of Birth from HCHS/SOL Visit 2
CLINDATE_COMPASS

3.2. AGEGROUP_C5_COMPASS (5-level grouped age of participant)

This is a 6-level ordinal variable determined from the derived variable AGE_COMPASS. Note that COMPASS participants came during/after HCSH/SOL visit 2. The youngest COMPASS participant was 24 years old at the COMPASS clinic visit.

AGE (yrs)	AGEGROUP_C6_COMPASS
18 – 34	1
35 – 44	2
45 – 54	3
55 – 64	4
65 +	5

Response Format:
1=Ages 18-34
2=Ages 35-44
3=Ages 45-54
4=Ages 55-64
5=Ages 65+

Source variable(s):
AGE_COMPASS. Age of participant at the time of COMPASS clinic visit

4. GLOBAL PHYSICAL ACTIVITY QUESTIONNAIRE (GPAQ)

The GPAQ variables were defined identically to variables in HCHS/SOL baseline. For completeness we include here their definitions (identical to HCHS/SOL documentation).

The GPAQ was developed by the World Health Organization (WHO). It collects information on physical activity participation in three settings (or domains: work, travel, leisure) and sedentary behavior. Scoring information, available at www.who.int/chp/steps , was used to define the variables below.

A MET (Metabolic Equivalent) is the ratio of the metabolic rate for a given activity relative to resting metabolic rate. One MET is defined as 1 kcal/kg/hour and is equivalent to the energy cost of sitting quietly. A MET is also defined as oxygen uptake in ml/kg/min with one MET equal to the oxygen cost of sitting quietly, around 3.5 ml/kg/min.

For calculations described below, MET values were assigned value corresponding to an average MET for typical types and intensities of activity.

Domain	METS value
Work	• Moderate MET value = 4.0 • Vigorous MET value = 8.0
Transport	• Cycling and walking MET value = 4.0
Recreation	• Moderate MET value = 4.0 • Vigorous MET value = 8.0

Analyses information can be found in the following WHO documentation:
http://www.who.int/chp/steps/resources/GPAQ_Analysis_Guide.pdf

4.1. GPAQ_REC_MOD (Recreational physical (moderate) activity (min/day))

This is a numeric variable.

If days of activity and time spend at different intensity levels are not missing, then
GPAQ_rec_mod = (PAE16 × PAE17A_MIN)/7

Source variable(s):
In a typical week, …
PAE16. Days of moderate-intensity sports, fitness or recreational activities?
PAE17A_MIN. Time spent doing moderate-intensity recreational activities (min)

4.2. GPAQ_REC_VIG (Recreational physical (vigorous) activity (min/day))

This is a numeric variable.

If days of activity and time spend at different intensity levels are not missing, then
GPAQ_rec_vig = (PAE12 × PAE13A_MIN)/7

Source variable(s):
In a typical week, …
PAE12. Days of vigorous-intensity sports, fitness or recreational (leisure) activities?
PAE13A_MIN. Time spent doing vigorous-intensity recreational activities (min)

4.3. GPAQ_REC (Recreational physical activity (min/day))

This is a numeric variable.

If days of activity and time spend at different intensity levels are not missing, then
GPAQ_REC = sum(GPAQ_REC_MOD, GPAQ_REC_VIG).

Source variable(s):
GPAQ_rec_mod. Recreational physical (moderate) activity (min/day)
GPAQ_rec_vig. Recreational physical (vigorous) activity (min/day)

4.4. GPAQ_WORK_MOD (Work-related (moderate) physical activity (min/day))

This is a numeric variable.

If days of activity and time spent at different intensity levels are not missing, then
GPAQ_WORK = (PAE5 × PAE6A_MIN)/7

Source variable(s):
In a typical week, …
PAE5. Days of moderate-intensity activities as part of your work?
PAE6A_MIN. Time spent doing moderate-intensity activities at work (min)

4.5. GPAQ_WORK_VIG (Work-related (vigorous) physical activity (min/day))

This is a numeric variable.

If days of activity and time spent at different intensity levels are not missing, then
GPAQ_WORK = (PAE2 × PAE3A_MIN)/7

Source variable(s):
In a typical week, …
PAE2. Days of vigorous-intensity activities as part of your work?
PAE3A_MIN. Time spent doing vigorous-intensity activities at work (min)

4.6. GPAQ_WORK (Work-related physical activity (min/day))

This is a numeric variable.

If days of activity and time spent at different intensity levels are not missing, then
GPAQ_WORK = sum(GPAQ_WORK_MOD, GPAQ_WORK_RIG).

Source variable(s):
GPAQ_work_mod. Work-related (moderate) physical activity (min/day)
GPAQ_work_vig. Work-related (vigorous) physical activity (min/day)

4.7. GPAQ_TOTAL_MOD (Total physical (moderate) activity (min/day))

This is a numeric variable.

If days of activity and time spent at different intensity levels are not missing, then
GPAQ_TOTAL_MOD = sum((PAE5 × PAE6A_MIN), (PAE8 × PAE9A_MIN), (PAE16 × PAE17A_MIN))/7

Source variable(s):
In a typical week, …
PAE5. Days of moderate-intensity activities as part of your work?
PAE6A_MIN. Time spent doing moderate-intensity activities at work (min)
PAE8.Days walk or bicycle to get to and from places?
PAE9_MIN. Time spent walking or bicycling for travel (min)
PAE16. Days of moderate-intensity sports, fitness or recreational activities?
PAE17A_MIN. Time spent doing moderate-intensity recreational activities (min)

4.8. GPAQ_TOTAL_VIG (Total physical (vigorous) activity (min/day))

This is a numeric variable.

If days of activity and time spent at different intensity levels are not missing, then
GPAQ_TOTAL_VIG = sum((PAE2 × PAE3A_MIN),(PAE12 × PAE13A_MIN))/7

Source variable(s):
In a typical week, …
PAE2. Days of vigorous-intensity activities as part of your work?
PAE3A_MIN. Time spent doing vigorous-intensity activities at work (min)
PAE12. Days of vigorous-intensity sports, fitness or recreational (leisure) activities?
PAE13A_MIN. Time spent doing vigorous-intensity recreational activities (min)

4.9. GPAQ_TRSPORT (Transportation-related physical activity (min/day))

This is a numeric variable.

If days of activity and time spend at different intensity levels are not missing, then
GPAQ_TRSPORT = (PAE8 × PAE9A_MIN)/7

Source variable(s):
In a typical week, …
PAE8.Days walk or bicycle to get to and from places?
PAE9_MIN. Time spent walking or bicycling for travel (min)

4.10. GPAQ_TOTAL
(Total physical activity (min/day))

This is a numeric variable.

If activity time for each domain is non-missing, then
GPAQ_TOTAL = sum(GPAQ_WORK, GPAQ_TRSPORT, GPAQ_REC)

Source variable(s):
GPAQ_work. Work-related physical activity (min/day)
GPAQ_trsport. Transportation-related physical activity (min/day)
GPAQ_rec. Recreational physical activity (min/day)

4.11. GPAQ_TOTAL_MET
(Total physical activity (MET-min/day))

This is a numeric variable.

If activity time for each domain is non-missing, then
GPAQ_TOTAL_MET = ((PAE2 × PAE3A_MIN × 8) + (PAE5 × PAE6A_MIN × 4)
+ (PAE8 × PAE9A_MIN × 4)
+ (PAE12 × PAE13A_MIN × 8) + (PAE16 × PAE17A_MIN × 4))/7

Source variable(s):
In a typical week, …
PAE2. Days of vigorous-intensity activities as part of your work?
PAE3A_MIN. Time spent doing vigorous-intensity activities at work (min)
PAE5. Days of moderate-intensity activities as part of your work?
PAE6A_MIN. Time spent doing moderate-intensity activities at work (min)
PAE8.Days walk or bicycle to get to and from places?
PAE9A_MIN. Time spent walking or bicycling for travel (min)
PAE12. Days of vigorous-intensity sports, fitness or recreational (leisure) activities?
PAE13A_MIN. Time spent doing vigorous-intensity recreational activities (min)
PAE16. Days of moderate-intensity sports, fitness or recreational activities?
PAE17A_MIN. Time spent doing moderate-intensity recreational activities (min)

4.12. GPAQ_MOD_WEEK
(GPAQ - Total physical (moderate) activity (min/week)

This is a numeric variable that finds the average amount of time spent per week doing moderate physical activity.

GPAQ_MOD_WEEK = GPAQ_TOTAL_MOD*7

Source variable(s):
GPAQ_TOTAL_MOD. Total physical (moderate) activity (min/day)

4.13. GPAQ_VIG_WEEK
(GPAQ - Total physical (vigorous) activity (min/week))

This is a numeric variable that finds the average amount of time spent per week doing vigorous physical activity.

GPAQ_VIG_WEEK = GPAQ_TOTAL_VIG*7

Source variable(s):
GPAQ_TOTAL_VIG. Total physical (vigorous) activity (min/day)

4.14. GPAQ_level
(Physical activity level)

This is a categorical variable which assigns a value of 1 assigned to those with high total physical activity levels, a value of 2 to those with moderate total physical activity levels and a value of 3 to those with low total physical activity levels. Update 12/2013: corrected a typo in code such that participants who were actually "moderate" were being set to "low". Set participants to missing if they reported more than 1440 total minutes per day of active + sedentary.

Level of total physical activity	Physical activity cutoff value
High	• If vigorous PA (work+rec) ≥ 3 days & GPAQ_total_MET ≥ 1500 **OR** • If moderate+vigorous PA ≥ 7 days & GPAQ_total_MET ≥ 3000
Moderate	• If vigorous PA (work+rec) ≥ 3 days & GPAQ_total_vig × 7 ≥ 60 min **OR** • If moderate PA (work+transport+rec) ≥ 5 days & GPAQ_total_mod × 7 ≥ 150 min **OR** • If moderate+vigorous PA ≥ 5 days & GPAQ_total × 7 ≥ 600 min
Low	If the value does not reach the criteria for either high or moderate levels of physical activity

Source variable(s):
PAE2. Days per week of vigorous-intensity activities as part of your work
PAE5. Days per week of moderate-intensity activities as part of your work
PAE8. Days per week of walk or bicycle to get to and from places
PAE12. Days per week of vigorous-intensity sports, fitness or recreational (leisure) activities
PAE16. Days per week of moderate-intensity sports, fitness or recreational (leisure) activities
GPAQ_TOTAL_VIG. Total physical (vigorous) activity (min/day)
GPAQ_TOTAL_MET. Total physical activity (MET-min/day)
GPAQ_TOTAL. Total physical (moderate) activity (min/day)

4.15. GPAQ_PAG2008
(Activity level per 2008 PA guidelines (categorical - 4 levels))

This is an ordinal variable that contains the level of physical activity per the 2008 PA guidelines (http://www.health.gov/paguidelines/guidelines/default.aspx). There are 4 mutually exclusive levels: high, medium, low or inactive based on all counts. Set participants to missing if they were missing GPAQ_MOD_WEEK and GPAQ_VIG_WEEK. Set participants to missing if they reported more than 1440 total minutes per day of active + sedentary.

GPAQ_PAG2008	Definition from 2008 Physical Activity Guidelines (for adults)
4 = High activity	"more than the equivalent of 300 minutes of moderate-intensity physical activity a week" or more than 150 minutes of vigorous activity, or an equivalent combination of both. Activity should be performed in episodes of at least 10 minutes.
3 = Medium activity	"150 minutes to 300 (5 hours) minutes of moderate-intensity activity a week (or 75 to 150 minutes of vigorous-intensity physical activity a week)" or the equivalent combination of moderate and vigorous activity. Activity should be performed in episodes of at least 10 minutes.
2= Low activity	"activity beyond baseline but fewer than 150 minutes (2 hours and 30 minutes) of moderate-intensity physical activity a week or the equivalent amount (75 minutes, or 1 hour and 15 minutes) of vigorous-intensity activity" or the equivalent combination of moderate and vigorous activity.
1 = Inactive	"no activity beyond baseline activities of daily living"

Note that all borderline cases are given the higher level of activity. If none of these criteria are met, GPAQ_PAG2008 is set to missing.

```
if GPAQ_MOD_WEEK>=300 | GPAQ_VIG_WEEK>=150 | GPAQ_MV_WEEK>=300
then GPAQ_PAG2008=4;

else if 150<=gpaq_mod_week<300 | 75<=gpaq_vig_week<150 |
150<=gpaq_mv_week<300 then GPAQ_PAG2008=3;

else if 1<=gpaq_mod_week<150 | 1<=gpaq_vig_week<75 | 1<=gpaq_mv_week<150
then GPAQ_PAG2008=2;

else if (0<=gpaq_mod_week<1 & 0<=gpaq_vig_week<1) then GPAQ_PAG2008=1;

else GPAQ_PAG2008=.;
```

Response format: 1 = Inactive
2 = Low Activity
3 = Medium Activity
4 = High Activity

Variable GPAQ_MV_WEEK is an intermediate variable created for the purposes of this definition only. It is defined as below:

GPAQ_MV_WEEK=(GPAQ_TOTAL_MOD*7) + (2*(GPAQ_TOTAL_VIG*7))

Source variable(s):
GPAQ_MOD_WEEK. GPAQ - Total physical (moderate) activity (min/week).
GPAQ_VIG_WEEK. GPAQ - Total physical (vigorous) activity (min/week).
GPAQ_MV_WEEK. GPAQ - Total physical activity (min/week).

4.16. GPAQ_PAG2008YN
(Meets 2008 activity level guidelines (1=Yes, 0=No))

This is a 0/1 categorical variable that determines whether or not high or medium activity level has been met based on GPAQ_PAG2008 (all counts). Update 12/2013: changes reflected from GPAQ_PAG2008

GPAQ_PAG2008YN = 1 if GPAQ_PAG2008 = 3 **or** GPAQ_PAG2008 = 4
Else GPAQ_PAG2008YN = 0

Response format: 0 = No
1 = Yes

Source variable(s):
GPAQ_PAG2008. Activity level per 2008 PA guidelines (categorical - 4 levels)

4.17. GPAQ_SED
(Sedentary behavior (min/day)

This is a numeric variable

Convert PAEA18A to minutes (i.e. PAEA18A*60) and add it to PAEA18B. If GPAQ_SED < 10 then it was set to missing

Source variable(s):
PAE18A
PAE18B

II. Physical Activity/Actical Data

5. COMPASS_PA_EPOCH (15 sec EPOCH)

See Chapter 6 below for a brief description of the accelerometer protocol. Details are described in section 7 of the COMPASS Manual of Procedures. Same protocol was followed in CASAS.

5.1. COMPASS_PA_EPOCH

This dataset has the counts for each 15sec EPOCH (each one a record in the dataset). Hence, it has multiple records per participant. It also has data from multiple devices per participant if they wore them more than once. Almost all participants worn the device once. However, in COMPASS 23 participants worn it twice, and in CASAS 82 participants worn it two times and three participants worn it three times.

6. COMPASS_PA_MIN (counts at the minute level)

The protocol for the **accelerometer component** is described in section 7 of the COMPASS Manual of Procedures. Briefly, participants were asked to wear an accelerometer (Actical) for 7 days to assess the frequency, duration, and intensity of their physical activity during that time period. Field center staff gave participants the accelerometers after all physical examinations during the clinic visit were completed. Participants were fitted with a belt with the appropriate size waist and instructed to wear the accelerometer above the iliac crest, the location most sensitive to vertical movements consistent with ambulation. Participants were told to undertake their usual activities for 7 days while wearing the monitor, and to remove it only for swimming, showering, and sleeping. They were also provided written instructions and a phone number to call if any questions arose and were called mid-week to answer any questions about the device, to make sure the instructions were clear, and to remind them to wear the monitor. Participants left the clinic visit wearing the accelerometer. Most of the participants came back to return the device, where the data were downloaded and the monitor was re-initialized for reuse.

The COMPASS protocol stated that the accelerometer (Actical) should be **initialized (start recording the counts)** at midnight (00:00) of the next day of the clinic visit and participants should wear it for 7 days during walking hours. For participants the start time was different to midnight and, hence, minutes from 0:00 to start time were classified as non-wear time.

Although data was collected at the 15-sec epoch it was processed at the minute level. First data was aggregated to the minute level and then non-wear time was determined using the Choi algorithm (Choi et al., 2011), defined as at least 90 consecutive minutes of zero counts, with allowance of 1 or 2 minutes of nonzero counts if no counts were detected in a 30-minute window upstream and downstream of the 90-minute period. After determining non-wear, a day was defined as **ADHERENT** if it has at least 10 hours of wear time. To summarize data at the participant level, at least 3 adherent days of accelerometer data were required. The **COMPASS_PA_MIN_DERV dataset** has an indicator variable **(ADHRENTYN**) that identifies participants with at least three adherent days (i.e. each day with at least 10hrs of wear). It also includes an indicator variable (**WEEKEND_INCLUDED**) that identifies whether an adherent weekend day was included or not in case investigators wish to conduct sensitivity analyses to further exclude participants that did not provide an adherent weekend day. For those

participants with multiple devices (108 participants worn it two times and three participants worn it 3 times) data was summarized across all devices (e.g. two adherent days from wear #1 and three adherent days from wear #2).

As done in HCHS/SOL baseline, intensity of activity was defined using the following intensity level cutpoints (Colley et al., 2011; Wong et al., 2011):
Sedentary: <100 counts/minute
Light activity: 100-1534 counts/minute
Moderate activity: 1535-3961 counts/minute
Vigorous activity: >=3962 counts/minute

In summary, among all participants, 96.4% (4,491 participants out of 4,656) returned the Actical, 90.4% (n=4,208 out of 4,656) had at least three adherent days, and hence summarized physical activity data (i.e. averaged across adherent days and multiple devices when available).

Description of Exclusion	**Excluded**	**Included**	
	N	**N**	% from 4,656
COMPASS (including CASAS in SD)		**4,656**	100.0
No Actical	165	**4,491**	96.4
3 or more adherent days	283	**4,208**	90.4
3 or more adherent days including a weekend day	322	**3,886**	83.4

The dataset COMPASS_PA_EPOCH has the following variables: STARTDATE, STARTTIME, DAY, EPOCH, COUNTS, MIDNIGHT, DIFF_START_ISSUED, DIFF_START_REISSUED, ON_DATE, DIFF, DIFFTIME, and indicator variables (0/1): LIGHT, MODERATE, SEDENTARY and VIGOROUS. The variables that uniquely identify a record in the dataset are: ID + STARTDATE + EPOCH.

The dataset COMPASS_PA_MIN has the following variables. The variables that uniquely identify a record in the dataset are: ID + STARTDATE + MIN.

6.1. STARTDATE
Actical start date (KEY)

6.2. STARTTIME
Actical start time

6.3. SERIALNUM
Actical serial number

6.4. DAY
Consecutive day number (1=next day of clinic visit). The minimum number of wear days is X and the maximum number of days is within one wear. Those with multiple devices have more wear days.

6.5. MIN
60-sec interval. Record number (KEY). This is a derived variable by combining the 15-second EPOCH data.

6.6. COUNTS
Counts per minute

6.7. ON_DATE
Date participant starts wearing the Actical (Day 1)

6.8. SEDENTARY
Indicator (0/1) of Sedentary Activity

6.9. LIGHT
Indicator (0/1) of Light Activity

6.10. MODERATE
Indicator (0/1) of Moderate Activity

6.11. VIGOROUS
Indicator of (0/1) Vigorous Activity

7. COMPASS_MIN_SEDBOUTS (Sedentary bout variables at the Bout level)

After aggregating 15-sec epoch data to the minute level, a **sedentary bout** is identified by consecutive sedentary minutes (i.e., < 100 counts/min). The duration of a sedentary bout (**BOUT_DURATION**) is the length of this period of sedentary activity.

Based on the COMPASS_PA_MIN dataset, sedentary bouts were determined separately for each participant and wear attempt, across worn days. **For a sedentary bout around midnight (i.e. has minutes before midnight and after midnight),** all the minutes in the bout are assigned to the day the bout starts and minutes after midnight are not counted in the following day. For example, for a sedentary bout that starts on day 3 and extends to day 4, this sedentary bout belongs to day 3. The **COMPASS_MIN_SEDBOUTS** dataset has a record for each **bout** and with the bout duration in minutes (**SED_BOUT_LENGTH**). In this dataset, **NDAY** is the newly defined day-level variable to tract bouts across multiple worn days (i.e. crossing at midnight), indicating on which day the sedentary bout belongs to. In the COMPASS Overview document, the Table 4.1 summarizes the PA dataset characteristics (e.g., key identifier variables, number of records, and other). In summary, 919,445 sedentary bouts were determined for 1,725 participants, with 82 participants had two wear attempts and 3 participants with three attempts. There are 1,445 sedentary bouts that crossed multiple days, with the longest bout being 2040 minutes (34 hours) and crossing three days. This dataset was then used to create sedentary bout variables at the day level (**COMPASS_PA_MIN_DAYS)** and at the participant level (**COMPASS_PA_MIN_DERV)**.

7.1. STARTDATE
Actical start date (KEY).

7.2. DAY
Consecutive day.

7.3. ADHERENT
Actical serial number.

7.4. WEARORDER
Order number for multiple devices worn by a participant (1 to 3).

7.5. ADHERENTYN_COMPASS
Indicator for having at least 3 days with 10hrs each.

7.6. SED_BOUT_LENGTH
Bout Duration in minutes.

7.7. NDAY
New Day variable to tract bouts across two days (midnight). The total minutes of the bout are included only in the day before midnight.

7.8. BOUT (KEY)
Consecutive number for the order of the bout within each day. For example, bout=1 is the first bout of any given day. If a participant on day 1 has only 10 sedentary bouts then bout variable takes values 1 to 10.

8. COMPASS_PA_MIN_DAYS (Actical variables at the day level)

The COMPASS_PA_MIN_DAYS dataset has multiple records per participant. Specifically, one record per worn DAY per ID and STARTDATE (some participants have multiple wears). DAY 1 is the next day after the clinic visit day (protocol). Each ID has as many days as the device was worn (range from 4 to 29 days).

There are three types of variables:

a) **General variables.** These include day, date, device serial number, number or wear attempts (1, 2, or 3), wear order, number of wear days, number of adherent days, weekend indicator, wear hours.

b) **Activity variables.** These include derived variables with the number of minutes in each intensity (sedentary, light, moderate, and vigorous).

c) **Sedentary bout variables.** These include sedentary bouts of different lengths: at least 1, 5, 10, 20, 30, 40, 50, 60, or 90 minutes. It also includes sedentary bouts between a certain interval: 1 to 30 min, 30-60 min, and 60 to 90 min.

8.1. STARTDATE
Actical start date (KEY)
See general description at the beginning of section 5.

8.2. STARTTIME
Actical start time
See general description at the beginning of section 5.

8.3. SERIALNUM
Actical serial number
See general description at the beginning of section 5.

8.4. DAY
Consecutive day (KEY)
See general description at the beginning of section 5.

8.5. ON_DATE
Date participant starts wearing the Actical (Day 1)
See general description at the beginning of section 5.

8.6. TOT_CNTS
Total counts
See general description at the beginning of section 5.

8.7. TOTCNTS_SED
Total counts in sedentary behavior within a day
See general description at the beginning of section 5.

8.8. TOTCNTS_LIGHT
Total counts in light activity within a day
See general description at the beginning of section 5.

8.9. TOTCNTS_MOD
Total counts in moderate activity within a day
See general description at the beginning of section 5.

8.10. TOTCNTS_VIG
Total counts in vigorous activity within a day
See general description at the beginning of section 5.

8.11. TOTSED
Total minutes/day of sedentary activity (<100 counts/min)
See general description at the beginning of section 5.

8.12. TOTLIGHT
Total minutes/day of light activity (26-384 counts/min)
See general description at the beginning of section 5.

8.13. TOTMOD
Total minutes/day of moderate activity (385-923 counts/min)
See general description at the beginning of section 5.

8.14. TOTVIG
Total minutes/day of vigorous activity (>=924 counts/min)
See general description at the beginning of section 5.

8.15. TOTMV
Total minutes/day of moderate or vigorous activity (>=385 counts/min)
See general description at the beginning of section 5.

8.16. TOTHRS
Total hours monitor was worn per day
See general description at the beginning of section 5.

8.17. CNTS_HR
Average counts per hour
See general description at the beginning of section 5.

8.18. CNTS_MIN

Average counts per minute

See general description at the beginning of section 5.

8.19. DAY_OF_WK

Day of the week (1=Sunday,2=Monday...7=Saturday)

See general description at the beginning of section 5.

8.20. WEEKDAY

Week day indicator (1=Monday to Friday,0=Saturday or Sunday)

See general description at the beginning of section 5.

8.21. ADHERENT

Actical was worn>=10 hours(1=Yes,0=No)

See general description at the beginning of section 5.

8.22. N_WEAR_DAYS_ORIG

Count for number of wear days (Original)

See general description at the beginning of section 5.

8.23. N_ADH_DAYS_ORIG

Count for number of adherent days (>=10 hrs) (Original)

See general description at the beginning of section 5.

8.24. WEARORDER

Order number for multiple devices worn by a participant

See general description at the beginning of section 5.

8.25. NUMWEARS

Number of times a participant wore a device

See general description at the beginning of section 5.

Sedentary Bout Variables

Dataset COMPASS_PA_MIN_SEDBOUTS has all sedentary bouts for each participant ordered by time of the day. For example, NDAY=1 has the first sedentary bout in day 1, NDAY=2 has the second, and sequentially. Based on the bout-level dataset **COMPASS_PA_MIN_BOUTS,** five sedentary bout variables were defined for nine different bout lengths and three intervals.

A **bout length** (denoted with &CUTPOINT) contains sedentary bouts with bout duration lasting over certain threshold, or within certain boundary. For definition of sedentary bout, and bout duration (SED_BOUT_LENGTH), please refer to section 6.

For example, a bout length ">=5" contains all sedentary bouts >= 5 min (SED_BOUT_LENGTH >= 5). A bout category "[1,30)" contains all sedentary bouts between 1 and 30 minutes (1 <= SED_BOUT_LENGTH < 30).

	Bout length &CUTPOINT	**Variable naming for &CUTPOINT**	**Description**
1	>=1	1	Contain all sedentary bouts >= 1 min
2	>=5	5	Contain all sedentary bouts >= 5 min
3	>=10	10	Contain all sedentary bouts >= 10 min
4	>=20	20	Contain all sedentary bouts >= 20 min
5	>=30	30	Contain all sedentary bouts >= 30 min
6	>=40	40	Contain all sedentary bouts >= 40 min
7	>=50	50	Contain all sedentary bouts >= 50 min
8	>=60	60	Contain all sedentary bouts >= 60 min
9	>=90	90	Contain all sedentary bouts >= 90 min
10	[1,30)	1TO30	Contain all sedentary bouts between 1 and 30 min
11	[30,60)	30TO60	Contain all sedentary bouts between 30 and 60 min
12	[60,90)	60TO90	Contain all sedentary bouts between 60 and 90 min

Note: First 9 bout categories are nested. For e.g., bout category ">=30" is a subset of bout category ">=20"; while bout categories [1,30), [30,60), and [60,90) are mutually exclusive.

For each of the 12 bout CUTPOINTS, five sedentary bout variables were defined at the day level: SUM_&CUTPOINT, N_D_&CUTPOINT, MEAN_&CUTPOINT, PCT_N_&CUTPOINT, PCT_SUM_&CUTPOINT. For example, SUM_&CUTPOINT refers to 12 sedentary bout variables created: SUM_1, SUM_5, SUM_10, SUM_20, SUM_30, SUM_40, SUM_50, SUM_60, SUM_90, SUM_1TO30, SUM_30TO60, SUM_60TO90 at the day level. Note that they were created based on **NDAY** (see section 6). Thus, there are a total of 12*5 = 60 sedentary bout variables in COMPASS_PA_MIN_DAYS.

8.26. SUM_&CUTPOINT

Total bout duration for sedentary bouts of &CUTPOINT (min/day)

Computed on the day level, the sum of bout duration of all sedentary bouts belonging to the bout category &CUTPOINT. If there is no sedentary bout in the bout category &CUTPOINT on one day, this variable is set to 0.

8.27. N_D_&CUTPOINT

Number of bouts for sedentary bouts of &CUTPOINT (per day)

Computed on the day level, the number of sedentary bouts belonging to the bout category &CUTPOINT. if there is no sedentary bout in the bout category &CUTPOINT on one day, this variable is set to 0.

8.28. MEAN_&CUTPOINT

Mean bout duration for sedentary bouts of &CUTPOINT (min/day)

Computed on the day level, this variable equals to
SUM_&CUTPOINT / N_D_&CUTPOINT
If division by 0 is to occur, meaning there is no sedentary bout in the bout category &CUTPOINT on one day, then set this variable to missing.

8.29. PCT_N_&CUTPOINT

Percent of all bouts for sedentary bouts of &CUTPOINT (%/day)

Computed on the day level, this variable equals to
N_D_&CUTPOINT / N_D_1
It represents the proportion of all sedentary bouts that is of the bout category &CUTPOINT. If division by 0 is to occur, meaning there is no sedentary bout of any duration on one day, then set this variable to missing.

8.30. PCT_SUM_&CUTPOINT

Percent of total bout duration for sedentary bouts of &CUTPOINT (%/day)

Computed on the day level, this variable equals to
SUM_&CUTPOINT / SUM_1
It represents the proportion of the total bout duration that is from sedentary bouts of the bout category &CUTPOINT. If division by 0 is to occur, meaning there is no sedentary bout of any duration on one day, then set this variable to missing.

9. COMPASS_PA_MIN_DERV (Actical variables at the participant level)

This dataset summarizes the data at the participant level by averaging data only from adherent days (i.e. with 10 or more wear hours) and across devices when multiple were available. Below is a list of variables included in the dataset.

9.1. STARTDATE_COMPASS

9.2. START TIME_COMPASS

9.3. SERIALNUM_COMPASS

9.4. ON_DATE_COMPASS
(Date participant starts wearing the Actical (Day 1))

9.5. NUMWEARS
(Number of times a participant wore a device)

Response Format:
1 = One wear
2 = Two times worn
3 = Three times worn

9.6. ADHERENTYN_COMPASS
(Participant has at least 4 adherent (>=10 hrs) days (1=Yes, 0=No))

Response Format:
0 = No
1 = Yes

Source variable(s):
ADHERENT from COMPASS_PA_MIN_DAYS

9.7. WKENDDAY_INCLUDED_COMPASS
(Summarized activity includes at least one adherent weekend day)
Indicator of whether or not one of the adherent days is a weekend day.

Response Format:
0 = No
1 = Yes

Source variable(s):
ADHERENT and WEEKDAY from COMPASS_PA_MIN_DAYS

9.8. N_WEAR_DAYS_COMPASS
(Count for number of wear days)
Total number of days the Actical was worn by the participant.
Source variable(s):
N_WEAR_DAYS from COMPASS_PA_MIN_DAYS

9.9. N_ADH_DAYS_COMPASS
(Count for number of adherent (>=10hrs) days)
Number of days the participant is adherent (≥10 hours of wear time).
Source variable(s):
ADHERENT from COMPASS_PA_MIN_DAYS

9.10. N_ADH_WKENDDAYS_COMPASS
(Count for number of adherent weekend days)
Number of weekend days that are adherent (≥10 hours of wear time) days.
Source variable(s):
ADHERENT and WEEKDAY from COMPASS_PA_MIN_DAYS

9.11. N_ADH_WKDAYS_COMPASS
(Count for number of adherent week days)
Number of week days that are adherent (≥10 hours of wear time) days.
Source variable(s):
ADHERENT and WEEKDAY from COMPASS_PA_MIN_DAYS

9.12. CNTS_DAY_COMPASS
Average cnts/day monitor was worn
Average of TOT_CNTS (hours that the monitor was worn within a day), among adherent days.
Source variable(s):
TOT_CNTS and ADHERENT from COMPASS_PA_MIN_DAYS

9.13. HRS_DAY_COMPASS
(Average hrs/day monitor was worn)
Average of TOTHRS (hours that the monitor was worn within a day), among adherent days.
Source variable(s):
TOTHRS and ADHERENT from COMPASS_PA_MIN_DAYS

9.14. CNTS_MIN_DAY_COMPASS
(counts/min per day)
Average of CNTS_MIN (averaged counts per minute within a day) among adherent days.

Source variable(s):
CNTS_MIN and ADHERENT from COMPASS_PA_MIN_DAYS

9.15. SED_DAY_COMPASS
(min/day of sedentary activity (<100 counts/min))
Average of TOTSED (sum of sedentary minutes within a day) across adherent days.
Source variable(s):
TOTSED and ADHERENT from COMPASS_PA_MIN_DAYS

9.16. LIGHT_DAY_COMPASS
(min/day of light activity (100-1534 counts/min))
Average of TOTLIGHT (sum of light minutes within a day) across adherent days.
Source variable(s):
TOTLIGHT and ADHERENT from COMPASS_PA_MIN_DAYS

9.17. MOD_DAY_COMPASS
(min/day of moderate activity (1535-3961 counts/min))
Average of TOTMOD (sum of moderate minutes within a day) across adherent days.
Source variable(s):
TOTMOD AND ADHERENT from COMPASS_PA_MIN_DAYS

9.18. VIG_DAY_COMPASS
(min/day of vigorous activity (>=3962 counts/minute))
Average of TOTVIG (sum of vigorous minutes within a day) across adherent days.
Source variable(s):
TOTVIG and ADHERENT from COMPASS_PA_MIN_DAYS

9.19. MV_DAY_COMPASS
(min/day of moderate or vigorous activity (>=1535 counts/min))
Average of TOTMV (sum of moderate or vigorous minutes within a day) across adherent days.
Source variable(s):
TOTMV and ADHERENT from COMPASS_PA_MIN_DAYS

9.20. MOD_WEEK_COMPASS
(min/week of moderate activity (1535 – 3961 counts/min))
Defined as MOD_DAY * 7 among adherent participants (ADHERENTYN=1).
Source variable(s):
MOD_DAY and ADHERENTYN from COMPASS_PA_MIN_DERV

9.21. VIG_WEEK_COMPASS
(min/week of vigorous activity (>= 3962 counts/min))
Defined as VIG_DAY*7 among adherent participants (ADHERENTYN=1).
Source variable(s):
VIG_DAY and ADHERENTYN from COMPASS_PA_MIN_DERV

9.22. MV_WEEK_COMPASS
(Total min/week of moderate or vigorous activity (>= 1535 counts/min))
Defined as 7*(MOD_DAY + 2*VIG_DAY) among adherent participants (ADHERENTYN=1).
The min/day of vigorous activity is multiplied by two according to the 2008 PA Guidelines for Americans. So, for example 60 min/day of moderate activity counts the same as 30 min/day of vigorous activity.
Source variable(s):
MOD_DAY, VIG_DAY and ADHERENTYN from COMPASS_PA_MIN_DERV

9.23. PAG2008_COMPASS (Activity level per 2008 PA guidelines (categorical - 4 levels))

This is an ordinal variable that contains the level of physical activity per the 2008 PA guidelines (http://www.health.gov/paguidelines/guidelines/default.aspx). There are 4 mutually exclusive levels: high, medium, low or inactive based on all counts. This variable is defined only for participants with at least 3 adherent days (i.e. ADHERENTYN=1).

PAG2008	Definition from 2008 Physical Activity Guidelines (for adults)	Translation for HCHS/SOL study using Actical data
4 = High activity	"more than the equivalent of 300 minutes of moderate-intensity physical activity a week" or more than 150 minutes of vigorous activity, or an equivalent combination of both. Activity should be performed in episodes of at least 10 minutes.	If MOD_WEEK >=300 or VIG_WEEK>=150 or MV_WEEK>=300 then PAG2008=4; else go to next row.
3 = Medium activity	"150 minutes up to 300 (5 hours) minutes of moderate-intensity activity a week (or 75 to 150 minutes of vigorous-intensity physical activity a week)" or the equivalent combination of moderate and vigorous activity. Activity should be performed in episodes of at least 10 minutes.	If MOD_WEEK >=150 but <300 or VIG_WEEK >=75 but <150 or MV_WEEK >=150 but <300 then PAG2008=3; else go to next row.
2= Low activity	"activity beyond baseline but fewer than 150 minutes (2 hours and 30 minutes) of moderate-intensity physical activity a week or the equivalent amount (75 minutes, or 1 hour and 15 minutes) of vigorous-intensity activity" or the equivalent combination of moderate and vigorous activity.	If MOD_WEEK>=1 but <150 or VIG_WEEK>=1 but <75 or MV_WEEK >=1 but < 150 then PAG2008=2; else go to next row.
1 = Inactive	"no activity beyond baseline activities of daily living"	If 0 <= MOD_WEEK<1 and 0 <= VIG_WEEK<1 then PAG2008=1; else set to missing.

The guidelines that we used to help define these categories are:
http://www.health.gov/paguidelines/guidelines/summary.aspx

- For substantial health benefits, adults should do at least 150 minutes (2 hours and 30 minutes) a week of moderate-intensity, or 75 minutes (1 hour and 15 minutes) a week of vigorous-intensity aerobic physical activity, or an equivalent combination of moderate- and vigorous intensity aerobic activity. Aerobic activity should be performed in episodes of at least 10 minutes, and preferably, it should be spread throughout the week.

- For additional and more extensive health benefits, adults should increase their aerobic physical activity to 300 minutes (5 hours) a week of moderate intensity, or 150 minutes a week of vigorous intensity aerobic physical activity, or an equivalent combination of moderate- and vigorous-intensity activity. Additional health benefits are gained by engaging in physical activity beyond this amount.

Source variable(s):
MOD_WEEK, VIG_WEEK, MV_WEEK and ADHERENTYN from COMPASS_PA_MIN_DERV

9.24. PAG2008YN_COMPASS
(Meets 2008 activity level guidelines (1=Yes, 0=No))
High or medium activity level based on PAG2008_COMPASS (all counts).

Source variable(s):
PAG2008_COMPASS from COMPASS_PA_MIN_DERV

9.25. SUM_&CUTPOINT
Total bout duration for sedentary bouts of &CUTPOINT (min/day)
Average of SUM_&CUTPOINT across adherent days.

Source variable(s):
SUM_&CUTPOINT and ADHERENT from COMPASS_PA_MIN_DAYS

9.26. N_D_&CUTPOINT
Number of bouts for sedentary bouts of &CUTPOINT (per day)
Average of N_D_&CUTPOINT across adherent days.

Source variable(s):
N_D_&CUTPOINT and ADHERENT from COMPASS_PA_MIN_DAYS

9.27. MEAN_&CUTPOINT
Mean bout duration for sedentary bouts of &CUTPOINT (min/day)
Average of MEAN_&CUTPOINT across adherent days.

Source variable(s):
MEAN_&CUTPOINT and ADHERENT from COMPASS_PA_MIN_DAYS

9.28. PCT_N_&CUTPOINT
Percent of all bouts for sedentary bouts of &CUTPOINT (%/day)
Average of PCT_N_&CUTPOINT across adherent days.

Source variable(s):
PCT_N_&CUTPOINT and ADHERENT from COMPASS_PA_MIN_DAYS

9.29. PCT_SUM_&CUTPOINT
Percent of total bout duration for sedentary bouts of &CUTPOINT (%/day)
Average of PCT_SUM_&CUTPOINT across adherent days.

Source variable(s):
PCT_SUM_&CUTPOINT and ADHERENT from COMPASS_PA_MIN_DAYS

IV. References

Choi, L., Liu, Z., Matthews, C.E., Buchowski, M.S., 2011. Validation of accelerometer wear and nonwear time classification algorithm. *Med Sci Sports Exerc* 43, 357-364.

Colley, R., Garriguet, D., Janssen, I., Craig, C., Clarke, J., Tremblay, M., 2011. Physical activity of Canadian adults: Accelerometer results from the 2007 to 2009 Canadian Health Measures Survey. Statistics Canada no. 82-003-XPE, Health Reports 22, 7-14.

Troiano R, Berrigan D, Dodd K, Masse L, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. *Med Sci Sports Exerc* 2008;40(1):181-188.

2008 US Physical Activity Guidelines: http://www.health.gov/paguidelines/guidelines/default.aspx

Wong, S.L., Colley, R., Connor Gorber, S., Tremblay, M., 2011. Actical accelerometer sedentary activity thresholds for adults. J Physical Activity Health 8, 587-591.

World Health Organization (WHO) Global Physical Activity Questionnaire (GPAQ) http://www.who.int/chp/steps/GPAQ/en/index.html