ReadMe file for accelerometry data set from prospective, longitudinal cohort of persons undergoing outpatient rehabilitation after stroke or Parkinson Disease
June 2022, CE Lang

These accelerometry data were collected from 2017-2021 from persons receiving outpatient rehabilitation services to improve upper limb function (stroke only) or to improve walking (stroke and Parkinson Disease [PD]). Data collection was first stopped in March 2020 with the onset of the COVID 19 pandemic. Followed by a slow re-opening in 2021. Criteria for inclusion for enrollment were:

- Neurologist diagnosis of stroke or idiopathic PD (Hoehn-Yahr score 2-3), but not both diagnoses in the same individual;
- Referral for outpatient physical or occupational therapy;
- Anticipated to receive rehabilitation services for at least one month;
- Documented therapy goal(s) to improve upper limb function or walking mobility;
- Able to follow 2-step commands and participate in testing; and
- For persons with PD, stable dose of PD medication >2 weeks prior to enrollment and no medication changes anticipated during the time of therapy services.

Potential participants were excluded if they met one or more of the following criteria:

- Other neurological or psychiatric conditions, including deep brain stimulator implants;
- Other orthopedic conditions that limit UL capacity or mobility (e.g. amputation, severe arthritis, significant pain);
- Other co-morbid conditions such that the physician or therapy documentation indicates minimal chance for improvement in function (e.g. end-stage cancer diagnosis);
- Upper limb or walking capacity that is already near normal, as indicated by ARAT scores ≥ 52 or self-selected gait speeds ≥ 1.2 m/s).

Participants were enrolled in one of three subgroups: stroke UL, stroke walking, or PD walking, such that each subgroup could be considered statistically independent from one another. Participants were assessed at admission to outpatient services and then monthly thereafter until discharge or until 6 months, whichever came first. Assessments included measures of capacity for activity (Action Research Arm test for stroke upper limb subgroup, 10m walking speed for stroke and PD walking subgroups) and performance of activity in daily life (wearable sensor measurements on bilateral upper limbs with Actigraph GT9X Links or on the less affected ankle with Modus Health StepWatch SW1002 activity monitors). Details of the sample and data collection procedures can be found in:

- Lang CE, Holleran CL, Strube MJ, Ellis TD, Newman CA, Fahey M, DeAngelis TR, Nordahl TJ, Reisman DS, Earhart GM, Lohse KR, Bland MD (2023) Improvement in the capacity for activity vs. improvement in the performance of activity in daily life during outpatient rehabilitation. Journal of Neurologic Physical Therapy, 47: 16-25.

This paper should be cited in any cases of presentation or publication with the data.

The file, *LongitudinalOutPatientCohort_UL&Walking_20220731.csv*, is structured in long format, where the columns are the variables and each participant has multiple rows representing the time points. The first row for each participant includes the demographic information (left most column) and a variable “Include record in analysis” (right most column) indicating whether or not that individual was included in the primary results paper listed above. Rows are for the time points and include the activity capacity measures along with the activity

performance variables derived from the accelerometers or step activity monitors (sensor variables in columns X – AO).

Participants in the stroke upper limb subgroup wore the Actigraph accelerometers on both wrists for 3 days at each time point. Participants in the stroke or PD walking subgroups wore the Step Activity Monitors on the less affected ankle for 7 days at each time point.

Files are stored in folders by participant number. Each participant folder contains folders from the assessment time points.

For the stroke upper limb subgroup:
Within each time point are folders for the two upper limbs, "LUE" and "RUE", along with some files. The other files in each time point folder that may or may not be useful. The other files contain output variables and graphs, generated via MATLAB. Within each limb folder, there will be at least four files. Not all participants have all time points.

*.gt3x is the original Actigraph file and needs to be opened with Actigraph software.
*.agd is also an Actigraph file. It is used for visualizing data in the Actigraph software. There may be two *.agd files that were downsampled to different bin widths (e.g. 10 sec).
*RAW.csv is a comma-separated values file of the 30 Hz raw data, with accelerometry values in gravitational units (m/s^2).
*1sec.csv is a comma-separated values file with filtered and resampled data. Using the proprietary ActiLife software, data were bandpass filtered between 0.25 and 2.5 Hz and binned into 1 second epochs, where each second is the sum of the values within that second. Values are in activity counts, defined by the software as 1 activity count = 0.001664 gravitational units (m/s^2).

For the stroke and PD walking subgroups:
Within each time point are 5-7 files. As with the upper limb subgroup folders some of the extra files are output variables and graphs, generated via MATLAB.

*.swb is the Step Watch file and needs to be opened with Modus software.
*.csv a comma-separated values file of the exported Step Watch data.

[Note that this file storage/file naming system (especially for the upper limb) is a good example of how not to store data for future sharing. Please direct questions about the data to langc@wustl.edu]

Additional papers published so far using these accelerometry data include:

- Holleran CL, Bland MD, Reisman, DS, Ellis TD, Earhart GM, Lang CE (2020) Day-to-day variability of walking performance measures in individuals post-stroke and individuals with Parkinson disease. *Journal of Neurologic Physical Therapy*, 44:241-247.

Key to variables and coding:
Redcap_event_name is the timepoint
Affected_side and dominant_side: 1 = left, 2 = right, 3 = both
Sex: 0 = female, 1 = male
Ethnicity: 1 = non-Hispanic/Latino, 2 = Hispanic/Latino

Race: 1 = White, 2 = Black or African-American, 3 = Asian or Asian-American, 4 = American Indian or Alaska Native, 5 = Native Hawaiian or Other Pacific Islander

Time_poststrokeconsent is the days since stroke

Stroke_type: 1 = ischemic, 2 = Hemorrhagic, 3 = other, 4 = unknown

Stroke_location: 1 = cortical, 2 = subcortical, 3 = cortical & subcortical, 4 = posterior circulation or cerebellum, 5 = other

Number_strokes is the total number of strokes

Time_postpdconsent is the years since PD diagnosis

Work_before_dx: 1 = yes, 2 = no, 3 = unknown

Work_now: 0 = not working for paid employment, 1 = working < 20 hrs/wk, 2 = working part-time ≥ 20 hrs/wk, 3 = working full time, ≥ 37.5 hrs/wk

Education: 0 = up to 8^{th} grade, 1 = High school, 2 = vocational/technical school, 3 = some college, 4 = Bachelors degree, 5 = Master, Doctoral, or professional degree

Living_status: 1 = living along and independent with basic ADLs, 2 = living alone and needs assistance with basic ADLs, 3 = living with others and independent with basic ADLs, 4 = living with others and needs assistance with basic ADLs.

Pdmeds are the medications taken specificially for Parkinson Disease

Group: 1 = upper limb stroke, 2 = stroke walking, 3 = PD walking