

Description of files for raw and derived data for the observational validation study

Contact/correspondence for data set: Arne Mueller (arne.mueller@novartis.com)

Access/URLs:

- Independent validation study
 - Derived data and metadata: <https://doi.org/10.5281/zenodo.2841297>
- Interventional clinical trial
 - Derived and metadata: <https://doi.org/10.5281/zenodo.2846013>
- Both studies
 - Full dataset: <https://mueller-et-al-2019.s3.amazonaws.com/index.html>

Title of study: An observational study of gait parameters in old men and women with slow walking speed

Summary

Objectives: The primary objective of this study was to investigate if a portable accelerometer device (actibelt) could reliably quantify the walking speed of slow-walking older women and men across different surfaces.

Study design: Interested and eligible participants were examined for their usual walking speed (4 m walking distance) using this screening method, and those patients with mean walking speeds <0.8 m/s were evaluated for further inclusion and exclusion criteria. Participants who fulfilled the inclusion criteria were invited to participate in the study and complete the given course three times (at least twice). The participant walked across the defined parcours (a minimum of 70m) at his usual walking speed while wearing the actibelt®. The screening and the completion of the course could take place on the same day or on different days.

Target population and main criteria for inclusion: The study included all patients (≥ 60 years old) who were in outpatient or inpatient care at the Department of General, Accident and Reconstructive Surgery or another LMU clinic, who met the inclusion criteria and consent to participate in the examinations. Participants had to have a normal walking speed between 0.4 and 0.79 m/s, inclusive, to be eligible for the study. The exclusion criteria were intended to ensure that patients whose walking speed and natural movement are restricted by trauma or other relevant medical conditions are not included. Since the

primary study purpose is a cross-sectional observational study with the primary goal of data collection, there is no control group.

Number of patients/subjects: It was expected that 25 patients would participate to the study.

Parcours description

The parcours was a predefined walking course that, to some extent, reflected real life environment. The parcours length was approximately 250 meters and included the climb and/or descent of several sets of stairs with about 50 steps in total (up and down).

If a participant could not manage the stairs, an elevator was available and it was up to the participant to decide whether to use the staircase or the elevator. Partial use of the elevator, either up or down, was also accepted. The parcours length was around 210 meters if the stairs were excluded.

The parcours stretched inside as well as outside the building and passed over even floor surfaces, carpets and hard outdoor concrete tiles.

The parcours started with a hallway section, followed by a staircase or elevator, then led through the ground floor lobby to an outdoor section in front of the hospital building. The subject returned by walking up a ramp for disabled and then the same route back to the starting point.

Parts of the parcours could be skipped if the participant or investigator believed it was in the best interest of the participant, for example when there was heavy rain, wind gusts, the patient was exhausted, or just a general risk that the subject could not complete the rest of the parcours.

Description of the parcours sections:

Section codes	Detailed description
Hallway forward Hallway back	An approximately 40 meter straight corridor with even hardwood floor.
Stairs down Elevator down Stairs up Elevator up	This section included a landing area in front of the stairs and elevators, as well as the staircase. The stairs had 23 steps separated by two plateaus.
Lobby forward Lobby back	The area stretching from the bottom of the staircase, alternatively from the elevator doors, through a lobby with even floor tiles and then further through the two automated sliding entrance doors to the outside part of the parcours. Carpets covered the floor of some parts of this section.
Outside forward Outside back	An approximately 55 meters long outdoors section on reasonably flat concrete tiles and short sections of cobblestone surface. In the outside forward section, there were also 5 steps down from the hospital entrance level to the pavement.
Ramp	An approximately 12 meter long ramp for disabled, including a small plateau that divided the ramp into two parts and where the patient had to turn 180 degrees to continue on the second part. The walking surface consisted mostly of the ramp's metal grate but also short sections of cobblestones.

The parcours sections described by above were annotated, meaning the time stamps of when the patient entered and left the different sections were determined from the video footage.

In addition to the sections, time stamps of different so called “events” were annotated. Events are sequences in the video footage for which the corresponding actibelt data have to be omitted before data analyses. A typical example is when the wheel operator seems to move with a different speed than the patient, thus not reflecting the patient’s true speed. Other examples are if the patient had to pause for a while, or the five steps of the outside forward section where the wheel was lifted off the ground. Description of annotated events are listed in the table below.

Description of the various events:

Event codes	Detailed description
outside stairs	The sequence when the subject walks down the five steps in the outside forward section.
wheel behind	A sequence where the wheel falls behind or catches up and therefore has a lower or higher speed than the subject.
pause	A sequence when the subject and wheel are pausing for at least 10 seconds.
blind	A sequence when no steps can be identified at all in the footage, e.g. due to poor video quality, total loss of focus or the camera does not keep the feet or legs in frame, i.e. the annotator is essentially “blind”.
other	Any other sequence worthwhile noting, which interrupts the flow of walking or compromises the validity of the data. For such an event, a specifying comment was added to the annotation file.

Note that the time stamp for the end of a section was normally also the start of the next, i.e. the annotated time points were the same. However, the time stamps of two adjacent sections were different if an event occurred in between.

A common example is when a wheel behind event started some seconds before the patient reached the end of a section and the wheel operator only managed to align the speed again after the patient had entered the next section. The wheel behind event would thus span across the boundary between two sections. For such situations, the time stamp for the end of the first section would be when the wheel behind event started and the start of the next section when the event ended. In fact, the wheel behind event did not even have to be annotated because data analysis only included actibelt data from section start to section end, i.e. the data during the wheel behind event are by default already omitted.

If a subject aborted the parcours attempt in the middle of a section, the time point when the subject stopped was used as the section end.

It also happened that a subject turned around in the middle of a section to return to the starting point, in which case the section end was annotated to when the subject started to turn around, and the next section, for the opposite direction, to when the subject had finished turning and started to walk back. The turn itself was considered an event but as it occurred between two sections, it did not have to be annotated.

The heel strike and toe-off events of all steps (both left and right) and for all patients were also annotated to allow a more detailed analysis of the algorithm’s step detection.

Determination of the synchronisation offset

The patient-worn and wheel-mounted actibelts recorded data at a frequency of 100 Hz and the smartphone recorded video at either 120 or 30 frames per second. Because the three devices used separate intrinsic clocks when recording the data, the data streams had to be precisely aligned to allow accurate analysis. Every parcours attempt therefore started with a synchronisation procedure.

The site staff simply used the smartphone to film when they tapped the two actibelt recording boxes held together twice with a normal pen. The tapping event created a very distinct and identical accelerometer signature in the two actibelt data flows, allowing one of the streams to be shifted out by an offset until the signatures were perfectly superimposed. The offset was determined and applied to the downstream analysis.

As the tapping event was filmed, the offset for the video file could also be determined and used for annotation of the data sets.

Description of how subject and wheel speed was compared

Per subject and run gait speed is calculated. With the step and speed algorithm for the subject steps are detected and speed is estimated per detected step. The measuring wheel and accelerometer measures distance and time per quarter wheel revolution from which speed per second epochs is calculated – this the reference speed. The wheel is pushed by the operator walking next to the subject with approximately the same walking speed.

Each subject walks a parcours up to three times. The parcours is split into 7 sections of different surface and environment. Per subject run step speed and wheel speed are median aggregate per section and compared graphically and statistically to analyze speed estimate bias and error

Description and justification of excluded subject runs

Some video annotation had to be adjusted (wrong annotations or annotation timings) – these are corrections and not exclusions. Four subject runs were excluded:

- 287-1: no sync event annotated (devices cannot be synced without)
- 929-3 & 136-1/2: very soft/flat heel strike resulting in no (or very few) detected steps, and therefore no step speed estimation possible (see figure below). Note, these subjects were using rollators.

When subjects pause for a while or there was an obvious discrepancy between wheel operator and subject speed these periods were excluded from averaging subject/run/section annotation. Periods with steps or wheel speeds of less than 0.25 m/sec were excluded as this is outside the range of the step-speed model and may not be reliably measured. Steps within sections had to be in continuous walking bouts of at least 10 steps to focus the analysis on “purposeful” walking. This avoid frequent stopping/starting of walks which usually result in discrepancies between wheel operator and subject.

Derived data

The data described in this section is derived from the raw data and video annotations, checked and cleaned.

annotation.csv: Section/parcours and event annotation (see above for details)

1. SubjectId: subject Id
2. RunId: run id
3. Annotation: annotation (see text above for possible annotations)
4. AnnotationType: either 'Section' or 'Event'
5. Comment: optional comment for the annotation
6. start: timestamp of start of the annotation (UTC)
7. end: timestamp of end of the annotation

steps.csv: one row per step as reported by the step algorithm

1. subjectId: subject Id
2. runId: run Id
3. running: is this a running step (logical)
4. speed: step speed in m/sec
5. duration: duration of the step in sec
6. distance: distance of the step in meters
7. side: left or right step (logical)
8. bout: the bout identifier this step belongs to (in a continuous series of steps)
9. freq: windowed step frequency
10. walk.ratio: windowed walking ratio
11. start: timestamp of the heel strike (UTC)
12. end: start + duration of the step (e.g. from heel strike of one foot to heel strike of the other foot)

wheel.csv: calculated wheel speed per seconds (reference speed)

1. SubjectId: subject id
2. RunId: run id
3. speed: average wheel speed during this second
4. start: start of a recording chunk (one second apart from the next chunk)
5. mid: mid timestamp of the chunk (usually 0.5 sec)

sectionSpeed.csv: speed summaries per subject, run and section

1. SubjectId: subject id
2. RunId: run id
3. Annotation: section annotation
4. subj.speed.sd: speed standard deviation per row
5. wheel.speed.sd: standard deviation of the reference speed for this subject, run and section
6. subj.speed: median speed for the subject as calculated by the algorithm
7. wheel.speed: median reference speed calculated by the measurement wheel (within this seconds)
8. steps: number of steps in this section
9. rollator: did this subject use rollator for this run (logical)?

Raw data and related to raw data

sync.csv: This table reports estimated clock-offsets between the two sensors and the video camera and can be used to correct timestamps for synchronization. This file also links to the raw data recordings in the HDF5 file.

1. SubjectId: subject id
2. RunId: run id
3. wheel.tap.idx: the index (0.01 sec) at which the “tapping” event was detect in the wheel sensor
4. subject.delta: subject tapping offset to the wheel as index (or units of 0.01 sec), can be used to sync clock skew between wheel and subject sensors
5. subject.path: path/key to the raw data in the HDF file
6. wheel.path: path/key to the raw data in the HDF file
7. subject.start: start time stamp of subject sensor in UTC
8. subject.end: end time stamp of subject sensor in UTC
9. subject.duration: duration of subject recording in seconds
10. wheel.start: start time stamp of wheel sensor in UTC
11. wheel.end: end time stamp of wheel sensor in UTC
12. wheel.duration: duration of wheel recording in seconds
13. subject.belt.id: serial number of subject sensors
14. wheel.belt.id: serial number of wheel sensor
15. video.delta: video offset in seconds to the tapping event on the wheel sensor (used to adjust the time in the video to the time on the wheel sensor)
16. video.start: the tapping event in in the video (in sec. from start of the video)

Note, to get a corrected (clock synchronized) time stamp for the subject sensor add the following offset in seconds: $Wheel_sensor_start - Subject_sensor_start + seconds(subject.delta/100)$

raw.h5: Raw acceleration data

This file contains all acceleation raw data with 100 Hz frequency. The sensors (when worn correctly) record x, y and z values corresponding to vertical, lateral and back-forward acceleration. The 4th column (“KSS”) can be ignored, it indicates if the belt buckle was a given time closed.

To get acceleration gravity units the x, y or z value needs to be divided by 336. The KSS indicates whether the belt buckle was closed (1) or open (0). Note, the order in which the data is returned depends on the programming language you use, e.g. in R the dataset is returned as a matrix of 4 columns and n rows, in phython it is n- columns and 4 rows. In the HDF5file itself, h5ls will list it as a matrix with 4 rows and n columns.

The path to access the raw data within the file is: wheel|subject, subjectId-runId, filename. These paths per file are available in the sync.csv file.

video-annotations.tar.gz: an archive with original per subject run parcours and step video annotation. The numeric values in these files are seconds relative to video start and are not synchronized with the accelerometers. The ‘sync’ event is the moment when the operator tapped the two accelerator sensors (held tight together) with a pen. The ‘*_steps.txt’ annotations annotate each step’s heel-strike and toe-off event with a quality estimate. Note the data in this archive is unprocessed and un-checked.

