ESSnet Smart Surveys

Grant Agreement Number: 899365 - 2019-DE-SmartStat

Link to our CROS website

Workpackage 2

Smart Survey Pilots

Deliverable 2.8: Functional and technical descriptions of tools – Physical activity (WP2.3)

Version 1.0, 2-12-2021

Prepared by:

Annemieke Luiten (CBS, Netherlands)
a.luiten@cbs.nl

Workpackage Leader:

Barry Schouten (CBS, Netherlands)
Jg.schouten@cbs.nl
telephone : +31 70 3374905

SUMMARY: WP2 performs four diverse pilots to inform WP3 on the specifications of a smart survey platform in the European Statistical System (ESS). This deliverable describes how the outcomes of WP2.3 (health) are fed back to WP3. The Inventory of Innovative Tools and Sources is described, using CSPA recommendations. A CSPA compliance template is filled in and added as appendix. The tools and sources inventory is supplemented by a methodological level, a logistics level and a legal-ethical-policy level. On top of these levels, a mapping of the WP2.3 pilots to the WP3 POCS and smart survey features is described. It is stressed that empirical motivation for specifications in this deliverable are given in deliverable 2.1 c.

1. Introduction

In WP2.3 we investigate possibilities to use sensor systems that inform on health issues. In this WP this is operationalized as the search for a sensor system for measuring motion and sedentary behaviour. Sensor measurement, in combination with machine learning algorithms may form a high quality method that gives precise information on people's physical activity and calories usage. WP2.3 was designed as a series of small scale pilots with a relatively large focus on Phase 0 (tool selection), and Phases 1 (conceptualisation and data collection instruments) and 2 (methodology). Phase 2' (IT and infrastructure) are touched upon, but are no main focus of the project. A description is given of the IT and architectural environment as needed for the chosen device. Other solutions are possible though, with additional or other requirements.

After an extensive research period (described in van Hoek et al 2021), Phase 0 ended with the selection of the activPAL3[1] as the tool of choice for all pilots. The activPAL3 ™(PAL Technologies Ltd., Glasgow, UK) is a small and slim device that directly measures the postural aspect of sedentary and active behaviour.

Other measurement systems of physical activity systems than the one chosen for these pilots are imaginable. It is for example possible to use smart watches or people's own fitbits. We have chosen not to use these for this ESSnet, but are researching the possibilities of these kind of devices. See Kraakman, Luiten and Toepoel (2021) for a feasibility study.

Deliverable 2.7 (Schouten, 2020) describes the elements of a smart survey tool, being the IT frontend, IT backend, Methodology and Logistics. Not all of the elements described in the Deliverable 2.7 are relevant for the smart survey using the activPAL, and there are on the other hand additional elements: the logistics of loading, preparing, sending out, receiving back and cleaning of the activPALs.

In the remainder of this deliverable, these four elements will be described for surveys using the activPAL. In Section 5 the legal-ethical-policy features of measuring physical activity in this way are describes. Sector 6 shortly mentions how country specific application could influence comparability of outcomes. Finally, in section 7, the mapping of these findings on the WP3 POCs is summarized.

For the IT frontend and backend, the existing Eurostat initiatives to create and maintain an inventory following CSPA (Common Statistical Production Architecture) are mostly sufficient. The inventory, as

[1] There is also an ActivPal4 version, with additional battery power, and more fine grained measurement. However, for our purposes the more economically prized activPAL3 was deemed sufficient.

described at https://webgate.ec.europa.eu/fpfis/wikis/display/ISTLCS/INVENTORY is filled in for the IT front and backend for the activPAL tool. See Annex A for the inventory questionnaire.

2. IT frontend

Requesting participants to wear an activPAL will most probably take place within the context of a Health or Living Condition Survey. Respondents will be asked if they are willing to undertake the additional task of wearing the accelerometer for a number of days. Hence, case management is primarily a task of the NSI's standard survey administration. For respondents consenting to the additional task, a new branch has to be developed towards the management of the activPALs.

Apart from the four elements described above, there are two additional element to the IT front end: the activPAL device itself and the PAL software. The software is needed to program the activPAL for use (assign an ID, assign a measurement period), to download the collected data, and to analyse the data.

2.1 The activPAL

An accelerometer measures the number of accelerations (g) in a given time period. Modern accelerometers like the activPAL measure on the three axes up/down, forward/backward/ and sideways. A gyroscope measures the relative position to the earth gravity. There is always 1 g towards the centre of the earth acting on the accelerometer. This can be used to estimate the angle of the accelerometer towards the line of gravity.

The activPAL is mounted on the upper thigh with a medicinal patch. ActivPAL proprietary algorithms inform about thigh position and acceleration that are used to determine body posture (i.e., sitting/lying and upright), stepping, and stepping speed (cadence), from which energy expenditure is inferred indirectly. The activPAL generates raw acceleration data that makes it possible to develop independent algorithms.

The accelerometer data are stored on the device. After the device is returned to the NSI, the data are extracted by linking the device to a computer with appropriate software. In principle it would be possible to have respondents install the appropriate software, download the data themselves, and transfer the data that they approve of to the NSI, but in practice this is of course not feasible.

An international consortium of research institutes, universities and health organisations also use the activPAL (or other thigh worn accelerometers with the same measurement qualities). This ProPASS consortium (the Prospective Physical Activity, Sitting and Sleep consortium) is working together towards the goal of the development of methods for processing, harmonizing, and pooling data and also towards harmonized methods for collecting data with thigh worn accelerometers.

Harmonized use of measurement systems and the use of (open source) algorithms to determine physical activity imply the use of similar devices by all NSIs. Choice for an accelerometer on another position of the body disqualifies algorithms developed for thigh worn accelerometers.

2.2. Setting up the activPAL for use

The software is downloaded from www.palt.com/users. New users need to register at the portal.

For the pilots, the software was installed on laptops that were not connected to the NSI mainframes. For the analysis of data, the software is preferably also available in the environment where additional analyses (e.g., with Python or SAS or SPSS) are performed.

There are two versions of the software tools available at the portal:

- **PAL Software version 8** (for both Microsoft Windows and Apple Macs)
- **PAL Software version 7** (the classic edition)

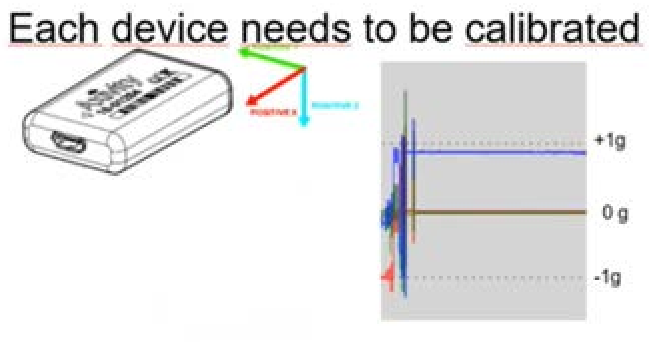
New users with new activPALs should select version 8. The software is split into 3 components all of which can be installed on Windows and OSX platforms:

- **PALconnect**
    - o supports the setup and downloading of the activPALs
    - o supports the seven port activDOCK for parallel setup and download of up to seven activPALs at a time
- **PALanalysis**
    - o visualisation tool providing views of the data
- **PALbatch**
    - o Allows batch processing of multiple data files generating pdf and csv (spreadsheet) outputs

Once a complete set of proprietary software is developed, independent of activPAL, PALanalysis and PALbatch are strictly speaking no longer necessary. However, PAL analysis generates attractive visualisations of the data that can be used for incentive purposes. The latter should ideally also be developed in house, for example in company (NSI) style, and in the language of the respondent. For this ESSnet we have used the standard PAL visualisations, with an explanation of its meaning in Dutch (for the NL and BE pilots) and German (for the GE pilot).

2.3. Calibration

Each device needs to be calibrated such that g is actually 1, and not slightly below, like in the picture below[2].



It is possible, but time consuming, to do that for each device beforehand. It is also possible, and much easier, to do that by using a calibration algorithm developed by van Hees et al (2014).

2.4 Diary

The inventory of elements mentions diaries as an element in a smart surveys tool. In Phase 0, (paper) diaries were used in which respondents wrote down the starting and ending time of a number of activities: starting to sleep, waking up, travelling (and with which mode), working, doing

[2] Figure derived from Palm & Hettiarachchi (2020).

sports (and which one), commuting (and how). The information was deemed necessary to inform the machine learning algorithms that we were building. The diary information also lends context to the activities (were activities performed as a result of work, or were they leisure, was biking electronically supported, etc.). However, subsequent data cleaning and analysis showed that most diaries were too imprecise to be of much use. For subsequent pilots we no longer used the diaries. If we feel we would like to have some context for the PA that we are measuring, other forms will need to be found. Having diaries also has methodological implications: the response burden is considerably higher if we ask respondents to fill in a diary, with consequences for response rates. It is our feeling that we need to further develop the algorithms in experimental settings, with the help of more sophisticated diaries, but once going live, the necessity of a diary should no longer exist.

2.5 Helpdesk support

If the supporting materials, explaining how the activPAL should be mounted and how long to wear it, are sufficiently clear, not much additional help is needed by respondents. The standard NSI helpdesk should be able to handle the few questions still remaining.

3. IT backend

IT backend describes case management, data processing, the web server, data models, the user database, linkage to existing data, etc. Not all of these aspects are relevant for WP2.3 in a smart statistics sense. Once the substantive variables are generated from the raw activPAL data, the IT backend is the same as for non-smart surveys. In this section, the processing of the activPAL data is described, the algorithms necessary for data cleaning and the algorithms to generate the substantial variables.

3.1. Data processing

The core output of the PAL analysis is a so called .PAL datx file, that is unreadable without the appropriate PAL software. For every participant (or more accurately for every activPAL), one datx file with a volume of about 9 MB is generated. From this core file, one or several data files and visualizations per respondent can be generated. This core file needs to be saved; all other generated filed can be deleted after analysis, as they can be generated again at any time.

The activPAL software lets the user choose a number of data formats. NSIs will need the raw data for the deployment of their proprietary algorithms. A raw data file with 7 days' worth of accelerometer data exists of about 450 MB of data per participant, in the form of a .csv file. Table 1 shows an example of the raw data that are generated.

Table 1. Raw data activPAL output of one second of sleeping (left) and one second of movement (right).

Time	Uncompressed sample index	X	Y	Z	Time	Uncompressed sample index	X	Y	Z
44003.99922	4	128	179	176	44004.31246	541283	60	129	122
44003.99922	5	128	178	176	44004.31246	541284	60	129	122
44003.99922	6	128	178	176	44004.31246	541285	60	129	122
44003.99922	7	128	178	176	44004.31246	541286	60	129	122
44003.99922	8	128	178	176	44004.31246	541287	60	129	122
44003.99922	9	128	178	176	44004.31246	541288	60	129	122
44003.99922	10	128	178	176	44004.31246	541289	60	120	120
44003.99922	11	128	178	176	44004.31246	541290	64	114	114
44003.99922	12	128	178	176	44004.31246	541291	58	123	122
44003.99922	13	128	178	176	44004.31246	541292	61	127	120
44003.99922	14	129	178	176	44004.31246	541293	63	125	119
44003.99922	15	128	178	176	44004.31246	541294	62	124	119
44003.99922	16	128	178	176	44004.31246	541295	61	124	120
44003.99922	17	128	178	176	44004.31246	541296	57	124	120
44003.99922	18	128	179	176	44004.31246	541297	61	118	119
44003.99922	19	128	178	176	44004.31246	541298	63	115	116
44003.99922	20	128	178	176	44004.31246	541299	64	113	114

Note. The time stamp has a floating point format that can be recalculated into a more interpretable format.

3.2. Plausibility checks and validation rules

The data need to be checked for non-wear: is the device worn for the requisite number of days? Is it worn for the minimum number of days? The calculation of ‘adherence to physical activity norms’ needs to be adapted to the number of wear hours. In addition, a check for alignment needs to determine if calculations need to be adapted to the position of the device on the body (perfectly vertical or otherwise). The activPAL algorithm checks and corrects for these errors. There exist open source algorithms to this end as well (Palm & Hettiarachchi, 2020).

3.3. Algorithms to subtract the substantive information

The raw data are input for the algorithms that determine kind and intensity of physical activity. See 4.2 for a description of the methodology of the algorithm development.

Once the algorithms have done their job, the 450 MB of information is condensed into a substantially smaller file, only containing the substantive variables. The latter file is only 4kB per respondent. Table 2 shows an (excerpt of) an example of the kind of condensed information that is generated by the PAL algorithm. Further analysis should condense this even more, to a summary of the total time spent on physical activity and sedentary behaviour per week. ActivPAL also offers this final condensation, but it is advisable to study the quality of data per respondent per day, in order to understand the quality of the final summary.

Table 2. Condensed PA summary output activPAL for one respondent; one row per day

Filenan Serial Num	Total Time (m)	Total Num Steps	Num Cycling Steps	Upright Time(m)	Standing Time(m)	Total Stepping Time(m)	Cycling Time(m)	Sitting Time(m)	Seated Transport Time(m)	Primary Lying Time(m)	Activity Score(MET.h)	Time Spent In Sitting Bouts >30m	Time Spent In Sitting Bouts >60m	Num Steps (duration >1m <=5m)	Num Steps (duration >5m <=10m)	Num Steps (duration >10m <=20m)	Num Steps (duration >20m)	Stepping Time(m) (Cadence >=75)
IOR01-IAP970211	1440	8960	460	652.075	527.503	124.572	6.46333	198.95	73.4967	518.28	34.8657	0	0	212	0	0	0	41.5717
IOR01-IAP970211	1440	9686	888	664.595	532.997	131.598	10.68	276.047	67.79	499.358	35.1592	91.8617	0	644	0	0	0	49.7283
IOR01-IAP970211	1440	7058	0	323.675	241.192	82.4833	0	376.668	9.97833	322.89	33.3579	195.052	139.168	646	0	0	2550	45.245
IOR01-IAP970211	1440	8422	256	597.657	488.788	108.868	3.17	177.767	72.5733	664.577	34.5354	0	0	460	0	0	0	47.9
IOR01-IAP970211	1440	9668	90	579.332	461.022	118.31	1.26833	338.387	71.975	522.282	34.9396	158.75	123.897	376	0	0	0	59.8983
IOR01-IAP970211	1440	7624	292	505.502	412.487	93.015	3.55667	439.693	70.9417	423.985	34.0169	220.8	118.905	952	0	0	0	51.1133
IOR01-IAP970211	1438.9	744	0	95.0817	83.2033	11.8783	0	519.268	0	824.553	30.4835	377.557	135.328	0	0	0	0	1.64167

Finally, the condensed summary data need to be combined to other information that the respondent provided through questionnaire or administrative data. The ID link given to the activPAL needs to be such that this final linking step is possible. It is only at this step that the data generated by the respondent becomes personal information. Prior to this point, the information does not contain any privacy sensitive information.

4. Methodology

The methodology level, as described in deliverable 2.7, consists of:

- User interface of frontend
- Data collection strategy, with the additional components of recruitment materials and incentives
- Plausibility checks and validation rules
- Machine Learning development

The user interface is not relevant for this application. The device does not interact with the respondent, nor gives any information back. The device needs to be returned to the NSI in order to extract the data that are stored on the device.

4. 1. Data collection strategy

The data collection strategy exists of the following elements: the choice whom among the population to measure with accelerometers (every sample person, the potentially least active ones, the ones with a moderate level of activity where adherence to PA guidelines may be strenuous). Other topics in this section are the choice of recruitment vehicle, recruitment materials and incentives.

4.1.1. Will everyone be measured with a device?

Most probably, more people will be willing to answer questions on physical activity, than there are people willing to wear an activity tracker. In addition, there may be differential willingness among groups within society. Two pilots within this ESSnet showed for example that people who are more physically active are also more willing to wear an accelerometer. Most of these were so active that a survey questionnaire captures their physical activity patterns adequately. These findings may implicate that statistics on physical activity of the population will need to be constructed from different sources. For some sample units accelerometer data will be available, but for others we may need to fall back on survey data. The question who to offer the device and who to offer a questionnaire, is an ongoing topic of research, but will potentially influence comparability across countries if different solutions are chosen.

4.1.2. How to offer the device to the respondent

Various scenarios of getting the activPAL to the respondent and various scenarios of getting them back again have been used in the ESSnet pilots and also in additional pilots and experiments that have been performed by CBS. The various scenarios have implications for the willingness of respondents to participate in this kind of survey research. More on the scenarios and implications for response rates will be described in deliverable 2.1, on substantive outcomes. For here, a short description suffices.

1. Sample units are approached with the express purpose of gaining participation in a research project on physical activity, to be measured by activPALs. Those who are willing receive the activPAL and are also asked to provide additional information on health and general physical activity.
2. The respondent in the (health) survey is asked, e.g., at the end of the survey, if they would be willing to wear the activPAL for one week. The device is sent in the mail, with an instruction how to mount and wear it. A return envelope is provided, and the respondent sends the device back.
3. If the request is made during a CAPI interview, the interviewer could be provided with a number of devices. She will be able to instruct the respondent, and hand over the device. The last scenario is probably the most response inducing technique, but could not be performed during the ESSnet period.

4.1.3. Recruitment materials

- Depending on the format chosen, see 4.1.2, the recruitment takes place via an invitation letter, a question in the questionnaire, or the interviewer. Regardless, the text of the recruitment warrents careful consideration. People are asked to do something that is alien to them. Therefore, the invitation needs to be clear, and address worries that people have expressed in pilots. How this invitation can best be formulated is a topic of ongoing research. This topic will be detailed in deliverable 2.1.
- Additionally, people need to receive a careful instruction on how to handle the device. The description is extensively written down for various research projects and also by activPAL. An instruction video exists on the web ([Putting on an ActivPAL 3 Activity Monitor - YouTube]), that should be adopted for the NSI's circumstances and language.

4.1.4. Incentives

Wearing an activPAL for a week is more burdensome than filling in a questionnaire. Hence, providing an incentive to participants is probably a prerequisite for attaining sufficient response rates. Statistics Netherlands has experimented with two forms of incentives in the WP2.3 pilots:

- conditional monetary incentives of various values,

- the promise of feedback on physical activity. This feedback existed of an overview of physical activity, as generated by the activPAL. In addition, an explanation was given how to interpret the information.

If, and how much incentive is offered will likely be different for different countries. The promise of feedback on physical activity did nothing for participation. If anything, there was a slight negative tendency of the promise. The results of the incentive experiments will be described more fully in deliverable 2.1.

4.2. Machine Learning development

An NSI using accelerometers like activPAL or another type, will not want to depend on the proprietary algorithms of the device manufacturer. That means that independent algorithms need to be developed. This is not an easy task. In order to understand physical activity and to deduce whether people adhere to physical activity guidelines (the substantive variable measured with the physical activity questionnaire in the Health Survey), the algorithm needs to be able to recognize the kind of activity performed, plus the intensity of the activity. The latter is easier than the former. An international consortium using the activPAL and other thigh worn accelerometers (ProPASS, the Prospective Physical Activity, Sitting and Sleep consortium) is working together to this aim. Their goal is the development of methods for processing, harmonizing, and pooling data and also towards harmonized methods for collecting data with thigh worn accelerometers. Algorithms have been developed for recognizing sleep and bicycling. The algorithms developed by ProPASS are open source. During Phase 0, Statistics Netherlands has dedicated considerable effort towards the goal of recognizing activities through machine learning. It appeared to be relatively straightforward to recognize the activities performed in a laboratory under supervision (standing, sitting, walking, running, cycling), where an accurate labelling was performed of timing and activities. However, the model generalized poorly to a 'free living' situation, where most activities were not labelled in the diary the participants filled in, and on top of that, inaccurately labelled. Van Hoek et al (2021) describe the steps taken by CBS and others to develop open source algorithms. The time and recourses within this ESSnet sub-project were not sufficient to fully develop those, but first steps have been taken.

The machine learning algorithm will need to be fully developed prior to going 'live' in official fieldwork, also for relevant subgroups (men, women, children, elderly, physically active, physically sedentary). During fieldwork, there is no communication between the device and the IT backend. No 'learning' can take place during fieldwork, for example by feedback from respondents on inferred activity. Other devices are perhaps suitable for a more interactive interplay with the backend, e.g., fitness apps on people's smartphones. However, a careful balancing of user friendliness, usability and quality needs to determine if the potential quality loss of using pre-developed machine learning algorithms outweighs the potential problems of other solutions.

5. Legal-ethical-policy

Deliverable 2.7 describes that smart surveys need a new framework of legal-ethical-and policy decisions. The bullet points below come from deliverable 2.7. For each, the relevance for WP2.3 is described.

- The data collection may use the computing and storage options of personal devices. In WP2.3 storage takes place on devices that are given to the respondent, not on personal devices.
- Part of the data collection may be passive. In WP2.3 all data collection is passive, although obviously, the respondent takes the active role of accepting the accelerometer and sticking it on their leg.
- Part of the data collection processing may be shifted to the personal devices rather than in-house at the institute. In WP2.3 we did not make use of this possibility; however, it is well within the realm of possibilities to use personal devices for the measurement of accelerometry. In that case, a number of decisions that are written down in this document need to be reconsidered.
- Part of the collected data may be of specialist content of which part of the respondents has no knowledge, even under informed consent. For WP2.3 we do not consider this to be a relevant dimension. The topic under study is relatively straightforward, i.e., the number of movements in three dimensions.
- In order to be able to participate, one may have to possess a specific device or to accept that one needs to use a device that is provided. This is the case for WP2.3, although we have discussed the position that data collection will still need a survey element for some groups in society.

6. Comparability over countries

Comparability of (smart) statistics over countries is crucial. Deliverable 2.7 makes a distinction between generic features, that are shared by all NSI's and that are considered to be MUST haves. For measuring physical activity by sensors, two generic aspects are relevant: the *position* of the device (not so much the device itself) and the algorithms used. Comparability can potentially also be threatened if countries deploy the device on varying subgroups within the population.
There are also features that a majority of countries prefer, but that not strongly effect measurement, the SHOULD haves. For example, differences in fieldwork strategy. This will most likely apply to measurement with accelerometers as well. Countries may differ in their fieldwork strategy (the use of

interviewers for example, or the use and amount of incentives). In the case of measurement with accelerometers it is however to be envisaged that some of these aspects *will* impact measurement, in cases where people who are not measured with a device will answer a questionnaire instead.
Finally, country-specific minority features concern those that are preferred by a few NSI's and do not affect comparability. They are considered COULD haves. For this particular case, no features like these have been described (yet).

7. WP3 mapping

Table 1 describes how the information described in this document maps to the areas of specific interest where WP3 will develop PoCs for modular prototype elements: Active-passive trade-offs, machine learning methodology, privacy-preservation, metadata, incentives and front-end configuration

TABLE 1: WP3 POC elements and pilots

	Health
Active-passive	Passive measurement of activity
Machine learning	Transformation of sensor data
Privacy preservation	No privacy sensitive data on sensor. Privacy becomes an issue once these data are linked to survey information and / or background variables. Standard PPT are available for that
Metadata	Sensor data and derivations
Incentives	Promised monetary incentives and personal statistics
Frontend configuration	Not in this pilot

In Deliverable 2.7, an overview is given of the (smart) features employed by smart surveys:

1. Device intelligence: It can use the intelligence (computing, storage) of the device on which it runs
2. Internal sensors: It can employ the sensors that are available in the device
3. External sensors: It can communicate through the device with other sensors close by
4. Public online data: It can go online and extract generally available data
5. Personal online data: It can go online and request access to existing external personal data
6. Linkage consent: It can ask consent to link external personal data already in possession of the survey institute

Table 2 describes how these elements work out for WP2.3.

TABLE 2: Smart features and pilots

	Health
Device intelligence	Accelerometer data are stored in the device
Internal sensors	YES
External sensors	NO. The data are transmitted by USB cables to computer storage
Public online data	Not in this pilot, but feasible with other devices than the one used in WP2.3
Personal online data	Not in this pilot, but feasible with other devices than the one used in WP2.3

Big data linkage	Not in this pilot, but the physical activity data will need to be linked to health survey data or register data on health consumption

8. References

Hees VT van, Fang Z, Langford J, Assah F, Mohammad A, da Silva IC, Trenell MI, White T, Wareham NJ, Brage S. Autocalibration of accelerometer data for free-living physical activity assessment using local gravity and temperature: an evaluation on four continents. J Appl Physiol (1985). 2014 Oct 1;117(7):738-44. doi: 10.1152/japplphysiol.00421.2014. Epub 2014 Aug 7. PMID: 25103964; PMCID: PMC4187052.

Hoek, S. van, Windmeijer, D., Luiten, A., Schouten, B., & Bolte, J. (2022). Comparing activity trackers to investigate physical activity: balancing quality, costs and usability. Statistics Netherlands working paper.

Kraakman, R. Luiten, A., and Toepoel, V. (2022). The Validity of Consumer-Level Activity Trackers in Research on Physical Activity: a Comparison with Self-Report and a Professional Accelerometer on Representation and Measurement. Under review Survey Research Methods.

Palm, P. & Hettiarachchi, P. 2021. ProPASS accelerometer data processing software. Presentation for the 3rd annual ProPASS meeting, June 2020. Available at [3rd Annual ProPASS Meeting - Dr Peter Palm & Dr Pasan Hettiarachchi - YouTube]

Schouten, B. (2020). Deliverable 2.7: Template for functional and technical descriptions of tools. [Available at confluence]