Rapid Review of Physical Activity Measurement and Surveillance Tools and Toolkits

Creating a culture of Physical Activity in Sheffield

Prepared by Prof. Robert Copeland, Dr Anna Myers and Dr Simon Nichols on behalf of the NCSEM Sheffield and the Centre for Sport and Exercise Science at Sheffield Hallam University

To Martin Towe & Diarmaid Brown, Armagh City, Banbridge and Craigavon Borough Council.

November 2018

Table of contents

Abbreviations

Centre for Sport and Exercise Science (CSES)

Energy expenditure (EE)

Moderate-to-vigorous physical activity (MVPA)

National Centre for Sport and Exercise Medicine (NCSEM)

Sheffield Hallam university (SHU)

Physical Activity (PA)

1. Introduction

- The National Centre for Sport and Exercise Medicine - Sheffield (NCSEM) and the Centre for Sport and Exercise Science (CSES) at Sheffield Hallam University (SHU), were commissioned by Armagh City, Banbridge and Craigavon Borough Council to conduct a rapid review of physical activity (PA) measurement and surveillance tools and toolkits.

1.1 Purpose

- The purpose of this rapid review is to a) highlight what current literature (reviews and primary articles) tells us about the strengths and limitations of various methods of measuring PA and b) to identify available PA measurement toolkits, in order to establish the most appropriate PA measurement tool for a given scenario.

1.2 Case Scenario

- To illustrate the application of tool to monitor PA, this Rapid Review uses a case study of an urban park.
- In this context, we provide suggestions as to how the various tools could be used
- These suggestions are not rules but rather guides as to how PA measures can be used to build up a picture of PA behaviour in a given context and depending on the primary outcome of interest.

1.3 Definition of terms

- Throughout this rapid review the following definition will apply when referring to PA, PA measurement tools and PA measurement toolkits.
    - **PA**: Any bodily movement produced by the skeletal muscle that results in energy expenditure (EE).
    - **PA measurement tools**: Any tool (subjective or objective) that measures the quantity of PA an individual undertakes. These can be subjective (questionnaires, diaries, surveys and interviews) or objective (movement sensors [pedometers, accelerometers, inclinometers, multi-sensors and Global Positioning System (GPS) and biological markers (doubly labelled water, calorimetry and heart rate).
    - **PA measurement toolkits**: A PA measurement toolkit is a resource to enable researchers/practitioners/end users to identify appropriate methods for measuring PA and/or evaluating the impact of programmes designed to increase PA.
    - **Agreement**: How close two measurements (made using the same scale) on the same subject are.

2. Methods

- We employed a range of techniques to review the evidence base for monitoring tools and available toolkits in PA contexts and used the data to inform a summary narrative for the rapid topic review.
- Given the timescales and nature of the request we approached this as a traditional 'Rapid Review', which has been characterised in the literature as a type of accelerated systematic review but with no commonly accepted or validated methodology.
- To strengthen the approach we also underpinned this work with a scoping review method (Arksey and O'Malley, 2005, Davis et al., 2009).
- A scoping review is a way of collecting and organising important background information which can be used to map existing literature (Davis et al., 2009) and is particularly appropriate when rapid and broad insights on a topic area are required.
- The process involved the following 6 stages and include a review of the academic, grey and practitioner literature. We also searched the online and print literature as comprehensively as possible:
    1. Design and conceptual clarification of rapid review
    2. Purposeful evidence-based searching, and gathering of evidence
    3. Policy and practice review
    4. A critical appraisal of existing research which is "time-limited"
    5. Validation and discussion among experts to ensure completeness (in this case the NCSEM stakeholders)
    6. Preparation of interim assessment, and final report

- Literature was sourced using a variety of methods;
    - Retrieval of primary sources (empirical studies) using exhaustive searches of online databases.
    - Use of secondary sources including published systematic reviews and meta-analyses (e.g. Cochrane database).
    - Reviews of national reports and existing guidelines/toolkits.
    - Hand searches of appropriate journals. Where existing relevant review articles are identified, these will also be hand searched.
- In addition to the literature review, we took advantage of the NCSEM Sheffield partnership to ensure we have a complete picture of the extant evidence for monitoring tools and available toolkits in PA contexts.

3. Results

3.1 Subjective physical activity measurement tools

- Subjective self-report measures, such as questionnaires, diaries, surveys and interviews are often used when PA data is being collected at a population level because they are practical, low cost and less burdensome to participants compared with objective measurement methods.
- Subjective measures of PA lack accuracy and their ability to detect change in PA levels is therefore questionable.
- Subjective measures of PA might be best employed as a tool to categorise groups based on PA levels, or as an adjunct to objective measures to provide information about mode of PA.
- Overall, the papers (reviews and studies) summarised in Table 2 (Appendix A) report poor agreement between subjective measures of PA and a criterion measure of PA (usually accelerometer).
- Agreement between questionnaire-based measurement and objective measurements is better when estimating sedentary time, compared with time spent participating in MVPA.
- Generally, self-report questionnaire data overestimates MVPA and underestimates sedentary behaviour when compared with accelerometer derived estimates of these parameters.
- There is a lack of research into whether subjective PA measurement tools are sensitive to PA change over time when administered sequentially.
- Limited evidence suggests that subjective PA measurement tools have poor sensitivity to change. Furthermore, sensitivity to change becomes worse as the time between recalls increases.
- Based on current data, the International Physical Activity Questionnaire - Short Form (IPAQ-SF) (http://youthrex.com/wp-content/uploads/2017/06/IPAQ-TM.pdf) is perhaps the most appropriate outcome measure for clinical and research use, although it still only reports moderate correlation with accelerometry.

3.2 Direct Observation

- Although not a subjective measure, direct observation of PA is a non-device based option for monitoring PA (hence included here).
- Direct observation is particularly valuable when activity is restricted to a particular space (e.g., use of tennis courts) or focused on population for whom self-report is more challenging (e.g. young children).
- Direct observation can also be used to gather contextual information about PA behaviour (e.g. preferred location, time, and clothing) including details of the PA itself (e.g., type, personalized variations to activities).
- Direct observation can be time consuming and therefore costly, particularly if the observation period is long and the number of observations is high. This method does not provide an estimate of EE.

Case Scenario – Subjective measures

- In our case scenario of an Urban Park, subjective self-report measures of PA (e.g. IPAQ-SF, Single Item Questionnaire) are perhaps best used to categorise and understand the populations that are currently using the park.
- When administered to the same population, they could be used to determine if there has been a change in PA category (status) over time (combined with other measures such as postcode) as a result of an intervention.
- Subjective self-report measures of PA can be used to assess the impact of an intervention but they are less sensitive to change compared to objective measures.
- These questionnaires categorise individuals based on total volume of PA accumulated – typically within a period of up to 4 weeks (usually in the past week).
- These measures often over-report the amount of MVPA that people take part in compared to objective measures but are practical for large populations due to their low cost
- These questionnaires can also be translated into electronic format's making them accessible via smartphone or tablets to large volumes of people.
- Direct observation of PA could be particularly useful to determine the impact of a discrete intervention on volume of use e.g. the restoration/transformation of disused areas of the urban park such as tennis courts or playground areas.
- Direct observation of PA would also help to identify contextual information such as time of use and nature of use (e.g. are the tennis courts used for tennis or something else).

3.3 Objective physical activity measurement tools

- Objective measures of PA, such as movement sensors (pedometers, accelerometers, inclinometers, multi-sensors and Global Positioning System [GPS]) and biological markers (doubly labelled water, calorimetry and heart rate) are more accurate and reliable than subjective measures of PA and, eliminate the potential for recall and response bias.
- Objective measurement tools are more commonly used in smaller scale research studies rather than population level surveillance because they are expensive compared to questionnaires and can be burdensome to participants.
- In addition, device setup and data analysis can be time consuming and requires expertise.

3.3.1 Movement sensors

- Objective measurement of human movement includes; pedometers, accelerometers/inclinometers, multi-sensors and GPS.
- The following section considers the merits of each of these in turn.

3.3.1.1 Pedometers

Studies assessing the effectiveness of pedometers are summarised in Table 3 (Appendix A).

- A pedometer is a portable device which counts the number of steps an individual takes by detecting movement - usually worn at the hip.
- Older pedometers are mechanical and have a lever and clock to count steps.
- Newer pedometers still have a lever, but use an electrical system to record step count.
- More advanced pedometers are entirely electrical (piezoelectric) and use accelerometers.
- Generally, pedometers have very good agreement with criterion measures (direct observation of step count, accelerometer) when estimating step count.
- However, pedometers are not designed to detect specific PA intensity categories (e.g. time spent in sedentary, light, moderate and vigorous PA).
- Pedometers are therefore an inappropriate choice for users who require this type of data.
- Newer pedometer do provide an indication of activity intensity by calculating step cadence (steps per minute), but further research is required to ascertain the reliability and validity of pedometers to estimate PA intensity.
- Pedometers generally agree with other objective measures, however piezoelectric pedometers appear to be less error-prone than mechanical pedometers, especially among people who use walking aids.

3.3.1.2 Accelerometers/inclinometers

Studies assessing the effectiveness of accelerometers/inclinometers are summarised in **Table 4** ***(Appendix A).***

- Like pedometers, accelerometers are typically worn on the hip. However, unlike pedometers, accelerometers measure acceleration of the body.
- Information on acceleration of the body is then converted into various outputs relating to PA using proprietary algorithms.
- Different accelerometers provide different PA related outputs including steps, estimated EE and time spent in different intensities of PA.
- A limitation of some accelerometers is they cannot detect body posture and therefore cannot distinguish between sitting/reclining and standing nor between different modes of activity (e.g. cycling, running)
- A further limitation of accelerometers is their inability to provide contextual information related to the setting and type of activity. Accelerometers are also unable to accurately detect the intensity of cycling and activities involving the use of upper extremities; this also applies to arm-worn monitors.
- Inclinometers overcome this limitation by measuring the angle of the body (usually the thigh) using accelerometer-derived information about body position and acceleration which in turn is used to determine body posture (i.e. sitting/lying and standing).
- As with pedometers, accelerometers provide an accurate measure of steps across a wide range of walking speeds.
- Furthermore, they also provide an estimate of time spent in different intensities of activity. This is useful where accurate data on the duration *and* intensity of PA over a specific period of time, is of interest.
- However, accelerometers are less accurate when estimating EE and generally underestimate the EE associated with PA compared with indirect calorimetry, particularly at higher intensities.
- Nevertheless, accelerometers provide a more accurate measure of PA EE than subjective measures and their use is also more feasible than the doubly labelled water method (see section 3.5.1.2).
- Inclinometers are more accurate than accelerometers for posture detection (i.e. sitting vs. standing) and are therefore recommended when sedentary behaviour/sitting time is an outcome of interest.
- Accelerometers and inclinometers have several limitations, they are more expensive than pedometers and subjective measures of PA.
- They also require more time and expertise to initialise, download, process and interpret the data.
- These strengths and weaknesses should be considered when deciding on a PA measurement tool.

3.3.1.3 Multi-sensors

Studies assessing the effectiveness of multi-sensors to monitor PA are summarised Table 5 (Appendix A).

- Multi-sensors combine accelerometer technology and sensors that measure physiological outcomes associated with physical exertion.
- These include heat flux, galvanic skin response (electrical resistance of the skin) and heart rate.
- The information from the multiple sensors is then used in proprietary algorithms to estimate PA related outcomes such as EE, steps and time spent in different intensities of PA.
- Different multi-sensors provide different outputs.
- Based on the studies summarised in Table 5 (Appendix A), multi-sensor PA monitors provide a more accurate estimate of EE than accelerometers alone, but tend to underestimate EE during higher intensity activities compared to doubly labelled water.
- Multi-sensors have also been shown to misclassify the intensity of PA when compared with EE estimated from indirect calorimetry.
- Nevertheless, multi-sensors provide a better estimate of EE and PA than self-report estimates of activity intensity and are better suited to free-living conditions than indirect calorimeters.

3.3.1.4 Global positioning system (GPS)

Studies assessing the effectiveness of GPS trackers to monitor PA are summarised in Table 6 (Appendix A).

- GPS provides information on the location, direction, and speed of the individual carrying a GPS receiver (e.g. smart phone, sports watch). Many people now carry GPS enabled devices (e.g. smart phones) that can store large amounts of data.
- GPS estimates of PA might therefore offer a cheap and feasible method of determining population PA.
- Most studies combined GPS data with accelerometer data and to our knowledge, no study has demonstrated that GPS data alone is a reliable and valid measure of PA. GPS derived information shows promise for identifying where and when PA and sedentary behaviour takes place.
- This could inform PA and sedentary behaviour interventions and lead to a more targeted and focussed approach to increasing PA/reducing sedentary behaviour.
- Context specific information about PA and sedentary behaviour could also inform future developments to the built environment.
- A limitation to using GPS to monitor PA is that devices often fail to record indoor activity (particularly in concrete buildings) and activity conducted under heavy tree canopy and in dense urban areas.
- The use of GPS functionality within smartphone devices can also drain battery life and hence is often a limitation of this approach from a user perspective.

Case Scenario – Objective measures

- In our case scenario of an Urban Park, objective measures (e.g. pedometers) are perhaps best used to understand the impact of different interventions on specific populations that are currently using or are the focus of greater engagement in use of the park.
- The exception to this is GPS that has the potential to be used to monitor large numbers of people (This is explored further in section 3.4).
- The type and nature of the intervention and the costs available for the evaluation will determine which of the objective measures is most appropriate.
- Pedometers are a good choice for use with large numbers of people and where assessment of the total volume of PA is the target, or a walking based intervention is employed (e.g. signed/led walk route). They are less suitable when assessing changes in time spent in different intensities of PA.
- Accelerometers are the ‘gold standard’ in terms of assessing time spent in different intensities of activity and assessing changes in PA over time.
- They are particularly useful where accurate data on the duration *and* intensity of PA over a specific period of time is of interest.
- For example, accelerometers would be an appropriate choice when assessing the impact of an intervention that aims to increase MVPA (by a given amount) in a certain population – such as the least active.
- They are also helpful in determining differences between the activity cost of different interventions. For example – are children more active in a traditional playground (with swings, slides, climbing frames) compared to a woodlands area (with felled trees, open spaces and natural obstacle to climb on)?
- The same value will come from the use of multi-sensors but these are likely to be less well received by users and prohibitively expensive for anything other than formal research.
- The use of GPS as a standalone tool or as an adjunct to accelerometers would provide rich contextual information about the PA behaviour of users in parks.
- Standalone, GPS would provide location and movement data within the Urban park, but also to and from the park, which might be of benefit to urban and transport planners.
- Whilst GPS appears an attractive option, there are shortcomings that need to be considered. Physical activity conducted under heavy tree canopy and in dense urban areas is often not recorded via GPS and we found no studies that have demonstrated that GPS data alone provides a reliable and valid measure of PA.

3.4 Smartphone technology

- Whilst we have explored objective measures of PA, including sensors used in smart phone technology (e.g. GPS), a thorough review of commercially available apps on mobile smart devices was beyond the scope of this review.
- That said, we have provided brief comment here on the potential of smartphone technology to monitor PA based on the findings of a fairly recent systematic review by Bort-Roig et al. (2014) (see Table 8 for associated studies).
- We have also considered findings from Coughlin et al. (2016) on the use of smartphone apps in the promotion of PA (see Table 9 for associated studies).
- Few studies have considered the validity of phone-based assessment of PA but for those that have, they found that measurement agreement were average-to-excellent for different PA behaviours - with the mobile phone placed mainly in the waist-to-hip area.
- Recent studies suggest that activities such as sitting, standing, walking, and jogging can be recognised with relatively high accuracy using in-built tri-axial accelerometer, gyroscope and magnetic sensors, however, measurement accuracy was mainly assessed with small samples completing a limited set of standardized activity trials. The accuracy of accuracy of smartphone motion sensors under free-living conditions is less well understood.
- Current data also suggests that smartphone technology can accurately measure a range of PA behaviours.
- The variety of novel and engaging intervention strategies offered, as well as the user perceptions on their usefulness and viability means that smartphones may have an important role in PA promotion.
- Qualitative data show that participants of various ages and gender respond favourably to user-friendly apps that automatically track PA (e.g., steps taken) and feedback progress toward personalised PA goals. Participants also want apps to be flexible enough so that they can be used with several types of PA.
- Participants also prefer apps that coach and motivate them and provide tailored feedback toward personally set goals.
- It is important to be mindful that intervention effects reported in the extant literature are modest at best. Future studies need to utilize randomized controlled trial research designs, larger sample sizes, and longer study periods to better explore the PA measurement and intervention capabilities of smartphones.
- One further note of caution is that that battery life is likely to limit longer measurement periods of PA. In our own research in Sheffield, we have noted that users delete monitoring apps if battery life is compromised.

Case Scenario – Smartphones

- In our case scenario of an Urban Park, the use of a smartphone to assess the PA of users on face-value presents a number of opportunities;
    - Firstly, the smartphone could be used as a tool to collect self-reported information using one of the validated questionnaires that we highlighted in section 4.1
    - Secondly, many smartphones include the functionality to monitor step counts and activity counts (via in-built tri-axial accelerometers, gyroscopes and magnetic sensors) and therefore might negate the need for additional monitoring tools (e.g. accelerometers).
    - Thirdly, smartphones also include GPS and therefore hold the potential to not only track PA within the Urban Park but beyond it.
    - Applications on smartphones can be tailored to different audiences to increase engagement and there is growing evidence of the types of interventions that appear to work best in terms of encouraging use of these applications (e.g. apps that coach, provide tailored feedback toward personally set goals).
- There is a strong word of caution here however, as smartphone applications are not cheap to develop and require constant upgrades and amendments to software to keep them operable.
- The accuracy of smartphones to assess PA is also in it's infancy and despite some early promise as highlighted by the reviews from Bort-Roig et al. (2014) and Coughlin et al. (2016), the extant literature lags far behind that of other objective and indeed self-reported measures.

3.5 Biological measures

3.5.1.1 Heart rate monitors

Studies assessing the effectiveness of HR monitors to assess PA are summarised in Table 7***(Appendix A).***

- Heart rate monitors can be used to predict PA EE based on the linear relationship between HR and EE.
- The flex HR method, which involves calibration of an individual's HR-to-EE relationship, has been shown to be a reliable and valid measure of EE.
- The flex HR method for estimating EE has been extensively validated against doubly labelled water and studies show it provides an accurate and reliable estimates of EE.
- However, individual calibration of the relationship between HR and EE can be time consuming and requires specialist equipment (indirect calorimeter).
- Furthermore, HR estimates of EE can be affected by prescribed medication such as beta blockers.
- An increase in HR stimulated by the sympathetic nervous system when an individual is resting may also result in inaccurate estimates of PA.
- Therefore this method is not suited for monitoring population PA levels.

3.5.1.2 Doubly labelled water and indirect calorimetry

- Doubly labelled water is the gold-standard method of determining EE (Melanson et al., 1996). It involves ingesting a non-radioactive isotope.
- Although widely accepted as the gold standard measure of EE, the feasibility of utilising it in population surveillance research is limited due to its high cost and participant burden.
- Studies using doubly labelled water to assess free-living PA have therefore not been reviewed.
- Due to similar limitations, indirect calorimetry has not been included in this review.
- However, studies which have examined the validity of other PA measurement tools with doubly labelled water and indirect calorimetry as the criterion measure were included.

3.6 Physical activity measurement toolkits

Table 1 overleaf provides links to a number of PA measurement toolkits along with a brief description of each.

- To help develop protocols to assess PA, we included a summary of some of the currently available toolkits that claimed to help evaluate PA interventions.
- As such, toolkits were only included in this rapid review if they contained at least one method of evaluating PA levels.
- We have not undertaken a quality rating of these toolkits, nor have we prioritised one over the other in any formal way. That said, we are familiar with these toolkits and in our opinion they do provide valuable and appropriate support for monitoring and evaluating PA.

4. Conclusions

- The purpose of this rapid review was to;
    a) Highlight what current literature tells us about the strengths and limitations of various methods of measuring PA
    b) Identify available PA measurement toolkits
    c) Establish the most appropriate PA measurement tool for a given scenario
- The review identified that the monitoring and evaluation of PA is complicated and there is no single measure that can adequately assess all the facets of PA and in different populations and contexts.
- In this rapid review, we covered a number of commonly used approaches – including those that have a long history of being used in PA assessment (e.g. self-report questionnaires) and more recent and novel approaches (e.g. smartphones).
- Self-reported assessments including self or interviewer-administered questionnaires (and diaries), can collect mode or type of activity in large numbers of people at low cost. They are reliable but their validity is questionable and the majority tend to over-report MVPA when compared to objective measures of PA.
- In contrast, objective measures, such as accelerometers or pedometers, are not subject to self-report error, provide detailed insight into an individual's movement patterns including PA intensity. However, objective measurements also have several limitations such as cost, time to administer and analyse subsequent data sets. Some of these tools are also unable to accurately measure some types of PA (e.g. weight lifting, and water activities).
- Smartphone applications, including GPS functionality have much potential in terms of reaching and monitoring large numbers of people. However, the evidence for such approaches is limited and similar issues of cost and expertise also exist here.

- It is important to approach PA measurement with a clear understanding of what type of data is required and, what that data will be used for. Considerations when choosing a PA measurement tool include:
    - Budget/resources
    - Purpose of the assessment
    - Component of PA being measured (i.e. steps, time spent in different intensities of activity, EE, posture)
    - Population under study and participant burden
    - Reliability and validity of the PA measurement tool
- In most circumstances, a combination of tools is likely to provide the 'best' assessment of PA but there are examples (as highlighted in our case scenarios) where single measures can provide valuable insight and at little cost (e.g. a self-report measure to categorise and understand the populations that are currently using a park).

Table 1 Physical activity measurement toolkits.

Name	Year	Organisation	Link	Type	Notes/description
Evaluation Framework	2018	Sport England	https://evaluationframework.sportengland.org/	Website and embedded PDFs	This resource aims to help Sport England colleagues and partners to evaluate funding streams and projects effectively and get maximum value from measurement and evaluation (M&E).
Diet, Anthropometry and Physical Activity (DAPA) Measurement Toolkit	NA	MRC - University of Cambridge	http://dapa-toolkit.mrc.ac.uk/	Website	Free web-based resource to assist researchers and public health or public end-users to identify methods for the assessment of diet, anthropometry and PA.
MRC Population Health Sciences Measurement Toolkit	NA	MRC - University of Cambridge	http://www.mrc-epid.cam.ac.uk/research/resources/	Website	MRC Population Health Sciences Measurement Toolkit is currently under review and will replace the DAPA toolkit.
Standard Evaluation Framework for physical activity interventions	2012	National Obesity Observatory	http://www.getirelandactive.ie/Professionals/Built%20Environment/Resources/Evaluating-Physical-Activity-.pdf	PDF	1. How to identify appropriate physical activity outcomes for evaluating different types of intervention. 2. How to define suitable measures for different types of physical activity outcome.

Name	Year	Organisation	Link	Type	Notes/description
					3. How to approach the challenges of assessing and measuring physical activity and energy expenditure.
A Practical Guide to Measuring Physical Activity	2015	Journal Article	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3915355/	Journal Article	This commentary summarises the main methods of measuring PA as well as providing examples of their uses in research trials
Physical Activity Evaluation Handbook	2002	Centres for Disease Control and Prevention	https://www.cdc.gov/nccdphp/dnpa/physical/health_professionals/interventions/handbook.pdf	PDF	The handbook provides tools to evaluate PA programmes and aims to help providers demonstrate programme outcomes and continuously improve provision.
Measurement of Physical Activity and Sedentary Behaviour	NA	Alberta Centre for Active Living	https://www.centre4activeliving.ca/services/measurement-physical-activity/	Website	The website summarises information that is relevant to the measurement of PA and sedentary behaviour. This includes definitions, guidelines, considerations when measuring PA and sedentary behaviour and examples of measurement tools.
Physical Activity and Sport Evaluation Toolkit	2018	University of Derby	http://derby.openrepository.com/derby/handle/10545/622421	Spreadsheet	The evaluation toolkit was developed through a collaboration between the East Midlands County Sports Partnership and the University of Derby. The purpose of the toolkit is to support the identification of tools and methods to monitor and evaluate the effectiveness of PA interventions.

Name	Year	Organisation	Link	Type	Notes/description
Measures Registry User Guide: Individual Physical Activity	2017	National Collaborative Childhood Obesity Research	http://nccororgms.wpengine.com/tools-mruserguides/wp-content/uploads/sites/2/2017/NCCOR_MR_User_Guide_Individual_PA-FINAL.pdf	PDF	The guide focuses on enhancing use of measures and tools to assess PA. The guide covers a variety of issues relating to PA measurement including the complexities of measuring PA, terminology and selecting a measurement tool.
Measuring diet and physical activity in weight management interventions	2011	National Obesity Observatory	https://www.google.com/search?ei=-_BrW9ieG9C2aZP5mpAK&q=measuring+diet+and+physical+activity+in+weight+management+interventions&oq=measuring+diet+and+physical+activity+in+weight+management+interventions&gs_l=psy-ab.3..0i71k1l4.0.0.0.13219.0.0.0.0.0.0.0.0..0.0....0...1c..64.psy-ab..0.0.0....0.Ve8lFlk8dg0	PDF	A shortlist of practical and validated questionnaires for the assessing PA and diet, to support practitioners to evaluate weight management interventions. The paper reviews the scientific literature and highlights the strengths and limitations of each questionnaire.
Move More, Sit Less: A toolkit for evaluating physical activity programs in your workplace	2017	Heart Foundation	https://www.heartfoundation.org.au/images/uploads/publications/4729_HF_-_Move_More_Sit_Less_toolkit_FA_Web.pdf	PDF	The toolkit builds on the existing Healthy Workplace Guide: Ten steps to implement a workplace health program. The toolkit is designed to help workplaces evaluate PA interventions.

5. References

ANDERSON, I., MAITLAND, J., SHERWOOD, S., BARKHUUS, L., CHALMERS, M., HALL, M., BROWN, B. & MULLER, H. 2007. Shakra: tracking and sharing daily activity levels with unaugmented mobile phones. *Mobile networks and applications,* 12**,** 185-199.

ARKSEY, H. & O'MALLEY, L. 2005. Scoping studies: towards a methodological framework. *International journal of social research methodology,* 8**,** 19-32.

BARREIRA, T. V., KANG, M., CAPUTO, J. L., FARLEY, R. S. & RENFROW, M. S. 2009. Validation of the Actiheart monitor for the measurement of physical activity. *International Journal of Exercise Science,* 2**,** 7.

BELL, R., TENNANT, P. W., MCPARLIN, C., PEARCE, M. S., ADAMSON, A. J., RANKIN, J. & ROBSON, S. C. 2013. Measuring physical activity in pregnancy: a comparison of accelerometry and self-completion questionnaires in overweight and obese women. *European Journal of Obstetrics & Gynecology and Reproductive Biology,* 170**,** 90-95.

BERNTSEN, S., HAGEBERG, R., AANDSTAD, A., MOWINCKEL, P., ANDERSSEN, S. A., CARLSEN, K. & ANDERSEN, L. B. 2010. Validity of physical activity monitors in adults participating in free-living activities. *British Journal of Sports Medicine,* 44**,** 657-664.

BHAMMAR, D. M., SAWYER, B. J., TUCKER, W. J., LEE, J.-M. & GAESSER, G. A. 2016. Validity of SenseWear® Armband v5. 2 and v2. 2 for estimating energy expenditure. *Journal of sports sciences,* 34**,** 1830-1838.

BORT-ROIG, J., GILSON, N. D., PUIG-RIBERA, A., CONTRERAS, R. S. & TROST, S. G. 2014. Measuring and Influencing Physical Activity with Smartphone Technology: A Systematic Review. *Sports Medicine,* 44**,** 671-686.

CALABRÓ, M. A., LEE, J.-M., SAINT-MAURICE, P. F., YOO, H. & WELK, G. J. 2014. Validity of physical activity monitors for assessing lower intensity activity in adults. *International Journal of Behavioral Nutrition and Physical Activity,* 11**,** 119-128.

CARTER, M. C., BURLEY, V. J., NYKJAER, C. & CADE, J. E. 2013. Adherence to a smartphone application for weight loss compared to website and paper diary: pilot randomized controlled trial. *J Med Internet Res,* 15**,** e32.

CASEY, M., HAYES, P. S., GLYNN, F., G, O. L., HEANEY, D., MURPHY, A. W. & GLYNN, L. G. 2014. Patients' experiences of using a smartphone application to increase physical activity: the SMART MOVE qualitative study in primary care. *Br J Gen Pract,* 64**,** e500-8.

CELIS-MORALES, C. A., PEREZ-BRAVO, F., IBANEZ, L., SALAS, C., BAILEY, M. E. & GILL, J. M. 2012. Objective vs. self-reported physical activity and sedentary time: effects of measurement method on relationships with risk biomarkers. *PloS one,* 7**,** e36345.

CHU, A. H., NG, S. H., KOH, D. & MÜLLER-RIEMENSCHNEIDER, F. 2015. Reliability and validity of the self-and interviewer-administered versions of the Global Physical Activity Questionnaire (GPAQ). *PLoS One,* 10**,** e0136944.

COUGHLIN, S. S., WHITEHEAD, M., SHEATS, J. Q., MASTROMONICO, J. & SMITH, S. 2016. A review of smartphone applications for promoting physical activity. *Jacobs journal of community medicine,* 2.

DAVIS, K., DREY, N. & GOULD, D. 2009. What are scoping studies? A review of the nursing literature. *International journal of nursing studies,* 46**,** 1386-1400.

DONAIRE-GONZALEZ, D., DE NAZELLE, A., SETO, E., MENDEZ, M., NIEUWENHUIJSEN, M. J. & JERRETT, M. 2013. Comparison of physical activity measures using mobile phone-based CalFit and Actigraph. *Journal of medical Internet research,* 15.

DOWD, K. P., SZEKLICKI, R., MINETTO, M. A., MURPHY, M. H., POLITO, A., GHIGO, E., VAN DER PLOEG, H., EKELUND, U., MACIASZEK, J. & STEMPLEWSKI, R. 2018. A systematic literature review of reviews on techniques for physical activity measurement in adults: a DEDIPAC study. *International Journal of Behavioral Nutrition and Physical Activity,* 15**,** 15.

DOWNS, A., VAN HOOMISSEN, J., LAFRENZ, A. & JULKA, D. L. 2014. Accelerometer-measured versus self-reported physical activity in college students: Implications for research and practice. *Journal of American College Health,* 62**,** 204-212.

DRENOWATZ, C. & EISENMANN, J. C. 2011. Validation of the SenseWear Armband at high intensity exercise. *European journal of applied physiology,* 111**,** 883-887.

DUNCAN, M., VANDELANOTTE, C., KOLT, G. S., ROSENKRANZ, R. R., CAPERCHIONE, C. M., GEORGE, E. S., DING, H., HOOKER, C., KARUNANITHI, M., MAEDER, A. J., NOAKES, M., TAGUE, R., TAYLOR, P., VILJOEN, P. & MUMMERY, W. K. 2014. Effectiveness of a web- and mobile phone-based intervention to promote physical activity and healthy eating in middle-aged males: randomized controlled trial of the ManUp study. *J Med Internet Res,* 16**,** e136.

EVENSON, K. R., CHASAN-TABER, L., SYMONS DOWNS, D. & PEARCE, E. E. 2012. Review of self-reported physical activity assessments for pregnancy: summary of the evidence for validity and reliability. *Paediatric and perinatal epidemiology,* 26**,** 479-494.

FOLEY, L., MADDISON, R., OLDS, T. & RIDLEY, K. 2012. Self-report use-of-time tools for the assessment of physical activity and sedentary behaviour in young people: systematic review. *obesity reviews,* 13**,** 711-722.

FOWLES, J. R., O'BRIEN, M. W., WOJCIK, W. R., D'ENTREMONT, L. & SHIELDS, C. A. 2017. A pilot study: Validity and reliability of the CSEP− PATH PASB-Q and a new leisure time physical activity questionnaire to assess physical activity and sedentary behaviours. *Applied Physiology, Nutrition, and Metabolism,* 42**,** 677-680.

GAO, C., KONG, F. & TAN, J. Healthaware: Tackling obesity with health aware smart phone systems. Robotics and Biomimetics (ROBIO), 2009 IEEE International Conference on, 2009. Ieee, 1549-1554.

HARRISON, C. L., THOMPSON, R. G., TEEDE, H. J. & LOMBARD, C. B. 2011. Measuring physical activity during pregnancy. *International Journal of Behavioral Nutrition and Physical Activity,* 8**,** 19.

HART, T. L., AINSWORTH, B. E. & TUDOR-LOCKE, C. 2011. Objective and subjective measures of sedentary behavior and physical activity. *Medicine and science in sports and exercise,* 43**,** 449-456.

HE, Y. & LI, Y. 2013. Physical activity recognition utilizing the built-in kinematic sensors of a smartphone. *International Journal of Distributed Sensor Networks,* 9**,** 481580.

HERMAN HANSEN, B., BØRTNES, I., HILDEBRAND, M., HOLME, I., KOLLE, E. & ANDERSSEN, S. A. 2014. Validity of the ActiGraph GT1M during walking and cycling. *Journal of sports sciences,* 32**,** 510-516.

HOLLIDAY, K. M., HOWARD, A. G., EMCH, M., RODRÍGUEZ, D. A., ROSAMOND, W. D. & EVENSON, K. R. 2017. Deriving a GPS Monitoring Time Recommendation for Physical Activity Studies of Adults. *Medicine and science in sports and exercise,* 49**,** 939-947.

HOOS, T., ESPINOZA, N., MARSHALL, S. & ARREDONDO, E. M. 2012. Validity of the global physical activity questionnaire (GPAQ) in adult Latinas. *Journal of Physical Activity and Health,* 9**,** 698-705.

IGELSTRÖM, H., EMTNER, M., LINDBERG, E. & ÅSENLÖF, P. 2013. Level of agreement between methods for measuring moderate to vigorous physical activity and sedentary time in people with obstructive sleep apnea and obesity. *Physical therapy,* 93**,** 50-59.

JOHANNSEN, D. L., CALABRO, M. A., STEWART, J., FRANKE, W., ROOD, J. C. & WELK, G. J. 2010. Accuracy of armband monitors for measuring daily energy expenditure in healthy adults. *Medicine and science in sports and exercise,* 42**,** 2134-2140.

KETABDAR, H. & LYRA, M. System and methodology for using mobile phones in live remote monitoring of physical activities. Technology and Society (ISTAS), 2010 IEEE International Symposium on, 2010. IEEE, 350-356.

KHALIL, A. & GLAL, S. StepUp: A step counter mobile application to promote healthy lifestyle. Current Trends in Information Technology (CTIT), 2009 International Conference on the, 2009. IEEE, 1-5.

KIRWAN, M., DUNCAN, M. J., VANDELANOTTE, C. & MUMMERY, W. K. 2013. Design, development, and formative evaluation of a smartphone application for recording and monitoring physical activity levels: the 10,000 Steps "iStepLog". *Health Educ Behav,* 40**,** 140-51.

KOZEY-KEADLE, S., LIBERTINE, A., LYDEN, K., STAUDENMAYER, J. & FREEDSON, P. S. 2011. Validation of wearable monitors for assessing sedentary behavior. *Med Sci Sports Exerc,* 43**,** 1561-1567.

KRENN, P. J., TITZE, S., OJA, P., JONES, A. & OGILVIE, D. 2011. Use of global positioning systems to study physical activity and the environment: a systematic review. *American journal of preventive medicine,* 41**,** 508-515.

LEE, J. A., WILLIAMS, S. M., BROWN, D. D. & LAURSON, K. R. 2015. Concurrent validation of the Actigraph gt3x+, Polar Active accelerometer, Omron HJ-720 and Yamax Digiwalker SW-701 pedometer step counts in lab-based and free-living settings. *Journal of Sports Sciences,* 33**,** 991-1000.

LEE, M. H., KIM, J., JEE, S. H. & YOO, S. K. 2011. Integrated solution for physical activity monitoring based on mobile phone and PC. *Healthcare informatics research,* 17**,** 76-86.

MADDISON, R. & MHURCHU, C. N. 2009. Global positioning system: a new opportunity in physical activity measurement. *International Journal of Behavioral Nutrition and Physical Activity,* 6**,** 73.

MARTIN, C. K., MILLER, A. C., THOMAS, D. M., CHAMPAGNE, C. M., HAN, H. & CHURCH, T. 2015. Efficacy of SmartLoss, a smartphone-based weight loss intervention: results from a randomized controlled trial. *Obesity (Silver Spring),* 23**,** 935-42.

MATTILA, J., DING, H., MATTILA, E. & SARELA, A. Mobile tools for home-based cardiac rehabilitation based on heart rate and movement activity analysis. Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE, 2009. IEEE, 6448-6452.

MELANSON, E. L., FREEDSON, P. S. & BLAIR, S. 1996. Physical activity assessment: a review of methods. *Critical Reviews in Food Science & Nutrition,* 36**,** 385-396.

MIDDELWEERD, A., VAN DER LAAN, D. M., VAN STRALEN, M. M., MOLLEE, J. S., STUIJ, M., TE VELDE, S. J. & BRUG, J. 2015. What features do Dutch university students prefer in a smartphone application for promotion of physical activity? A qualitative approach. *Int J Behav Nutr Phys Act,* 12**,** 31.

MORRISON, L. G., HARGOOD, C., LIN, S. X., DENNISON, L., JOSEPH, J., HUGHES, S., MICHAELIDES, D. T., JOHNSTON, D., JOHNSTON, M., MICHIE, S., LITTLE, P., SMITH, P. W., WEAL, M. J. & YARDLEY, L. 2014. Understanding usage of a hybrid website and smartphone app for weight management: a mixed-methods study. *J Med Internet Res,* 16**,** e201.

MYERS, J., DUPAIN, M., VU, A., JAFFE, A., SMITH, K., FONDA, H. & DALMAN, R. 2014. Agreement between activity-monitoring devices during home rehabilitation: a substudy of the AAA STOP trial. *Journal of aging and physical activity,* 22**,** 87-95.

RABIN, C. & BOCK, B. 2011. Desired features of smartphone applications promoting physical activity. *Telemed J E Health,* 17**,** 801-3.

RACHELE, J. N., MCPHAIL, S. M., WASHINGTON, T. L. & CUDDIHY, T. F. 2012. Practical physical activity measurement in youth: a review of contemporary approaches. *World Journal of Pediatrics,* 8**,** 207-216.

RAMIREZ-MARRERO, F. A., MILES, J., JOYNER, M. J. & CURRY, T. B. 2014. Self-reported and objective physical activity in postgastric bypass surgery, obese and lean adults: association with body composition and cardiorespiratory fitness. *Journal of Physical Activity and Health,* 11**,** 145-151.

SPIERER, D. K., HAGINS, M., RUNDLE, A. & PAPPAS, E. 2011. A comparison of energy expenditure estimates from the Actiheart and Actical physical activity monitors during low intensity activities, walking, and jogging. *European journal of applied physiology,* 111**,** 659-667.

TURNER-MCGRIEVY, G. & TATE, D. 2011. Tweets, Apps, and Pods: Results of the 6-month Mobile Pounds Off Digitally (Mobile POD) randomized weight-loss intervention among adults. *J Med Internet Res,* 13**,** e120.

VANDELANOTTE, C., CAPERCHIONE, C. M., ELLISON, M., GEORGE, E. S., MAEDER, A., KOLT, G. S., DUNCAN, M. J., KARUNANITHI, M., NOAKES, M., HOOKER, C., VILJOEN, P. & MUMMERY, W. K. 2013. What kinds of website and mobile phone-delivered physical activity and nutrition interventions do middle-aged men want? *J Health Commun,* 18**,** 1070-83.

WEBBER, S. C., MAGILL, S. M., SCHAFER, J. L. & WILSON, K. C. 2014. GT3X+ accelerometer, Yamax pedometer, and SC-StepMX pedometer step count accuracy in community-dwelling older adults. *Journal of Aging and Physical Activity,* 22**,** 334-341.

WIETERS, K. M., KIM, J.-H. & LEE, C. 2012. Assessment of wearable global positioning system units for physical activity research. *Journal of Physical Activity and Health,* 9**,** 913-923.

WU, W., DASGUPTA, S., RAMIREZ, E. E., PETERSON, C. & NORMAN, G. J. 2012. Classification accuracies of physical activities using smartphone motion sensors. *Journal of medical Internet research,* 14.
ZHONG, S., WANG, L., BERNARDOS, A. M. & SONG, M. 2010. An accurate and adaptive pedometer integrated in mobile health application.

Appendix A

Table 2 Summary of studies (original research and reviews) examining the methodological effectiveness of subjective physical activity measurement tools.

Author	Population	Link (URL)	Measure(s)	Comparator	Findings
Dowd et al. (2018) - systematic review of reviews	Adults	https://ijbnpa.biomedcentral.com/articles/10.1186/s12966-017-0636-2	Various	Various	Self-reported measurements of PA were more variable than objective measures. Test-retest reliability of self-reported measurements were wide. There was a trend for reduced levels of test-retest reliability as the duration of recall increased. Responsiveness to change in PA levels was the least reported property of PA measurement tools and requires further investigation.
Fowles et al. (2017)	Adults	http://www.nrcresearchpress.com/doi/abs/10.1139/apnm-2016-0412#.W2MR_E2oti4	Physical Activity and Sedentary Behaviour Questionnaire (PASB-Q) and the modified Leisure-Time Physical Activity Questionnaire (mLTPA-Q)	Accelerometer (ActiGraph GT3X)	Objectively measured moderate-to-vigorous physical activity (MVPA) was moderately correlated with the PASB-Q's physical activity vital sign (http://www.exerciseismedicine.org/assets/page_documents/The%20Physical%20Activity%20Vital%20Sign%20without%20Strength_2015_07_09_PDF.pdf) Absolute

Author	Population	Link (URL)	Measure(s)	Comparator	Findings
					agreement between self-reported and objective measurements of MVPA were good (little over or under estimation) Self-reported sedentary time was greatly underestimated in the PASB-Q compared with the objective measure (6.4 ± 3.5 vs 12.2 ± 1.2 h/day) and there were no correlations (between self-reported and objectively measured sedentary time.
Silsbury et al. (2015)	Adults	https://bmjopen.bmj.com/content/5/9/e008430	Various (e.g.IPAQ-SF, Single Item, Godlin-Shephard)	Various	Ambiguity in PA terminology, patient reporting of PA, and the variable nature of activity across the seasons and 7 days, makes daily activity difficult to assess using self-report questionnaires The optimum SRPAQ has not been reported although the IPAQ-SF appears the most appropriate outcome measure for clinical and research use, as it has excellent reliability and moderate correlation with accelerometry.

Author	Population	Link (URL)	Measure(s)	Comparator	Findings
Chu et al. (2015)	Adults and students	http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0136944	Global Physical Activity Questionnaire (GPAQ) - Self- and Interviewer-Administered	Accelerometer (ActiGraph wGT3X-BT)	Moderate correlations were found between the GPAQ and accelerometer derived estimates of PA at all PA intensities. In terms of validity, there was no difference in the correlation between the two modes of administration.. Relative to the GPAQ, the accelerometer measured lower daily total MVPA, vigorous-intensity activity and moderate-intensity activity.
Downs et al. (2014)	College students	https://www.tandfonline.com/doi/abs/10.1080/07448481.2013.877018	International Physical Activity-Short form (IPAQ-SF)	Accelerometer (ActiGraph GT3X+)	Estimates of time spent engaged in MVPA were significantly higher when measured via self-report compared with accelerometer derived estimates. Self-report and accelerometer derived MVPA were not correlated.
España-Romero et al. (2014)	Older British Adults	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3916297/	EPIC Physical Activity Questionnaire (EPAQ2)	Combined heart rate and movement sensors	Substantial differences observed between the EPAQ2 and the criterion method which highlights the challenges of assessing PA accurately in populations and puts current prevalence estimates and dose-response relationships based on self-report into question.

Author	Population	Link (URL)	Measure(s)	Comparator	Findings
					Differences between methods were comparable with other studies where these PA-subcomponents were assessed by means of PAQs and objective methods. Study suggests that the EPAQ2 has properties for ranking adults in their 60 s according to PAEE, sedentary time, light PA and MVPA, similar to its use in other populations and similar to other instruments.
Bell et al. (2013)	Overweight/obese pregnant women	https://www.sciencedirect.com/science/article/pii/S0301211513002492?via%3Dihub	Recent PA Questionnaire (RPAQ), Australian Women's Activity survey (AWAS)	Accelerometer (Actigraph GT1M)	Questionnaires over-estimated MVPA. There was no correlation between accelerometry and either questionnaire when measuring MVPA. There was also substantial disagreement in classification of those achieving at least 30 min of MVPA. Both questionnaire overestimated MVPA compared with accelerometer data. The median difference between the accelerometer and questionnaire estimates of MVPA were 101 min/day and 42 min/day for the AWAS and RPAQ, respectively.

Author	Population	Link (URL)	Measure(s)	Comparator	Findings
Ramirez-Marrero et al. (2014)	Post gastric-bypass surgery, obese and lean adults	https://journals.humankinetics.com/doi/pdf/10.1123/jpah.2012-0048	IPAQ-SF	Actigraph GT1M	Lean individuals accrued more MVPA than obese individuals and gastric bypass patients. All groups self-reported significantly more MVPA compared with accelerometer estimates. Obese participants showed a greater overestimation of MVPA compared to lean and gastric bypass patients. Obesity status may therefore influence peoples perception of how long and how hard PA is.
Myers et al. (2014)	Older adult patients with abdominal aortic aneurysm disease	https://journals.humankinetics.com/doi/abs/10.1123/japa.2012-0133	Hourly activity logs, interview to complete 3day activity recall questionnaire (3-DPAR)	Accelerometer (Actigraph GT1M), pedometer (Omron 720-ITC), Heart Ratemonitor (Polar F6)	Accelerometry provides 'reasonable' estimates of EE in patients participating in a home rehabilitation program. Mean energy expenditure (kcals/day) was 1687±458 (Heart Rate; HR), 2068±529 (accelerometer) and 1974±491 (3DPAR). Sinificant differences in EE were found between HR and accelerometry, HR and 3DPAR but not accelerometry and 3DPAR. Agreement between a accelerometry and 3DPAR was greatest..

Author	Population	Link (URL)	Measure(s)	Comparator	Findings
					The 3DPAR was used as criterion measure despite relying on self-report data and predictive equations too estimate energy expenditure.
Igelström et al. (2013)	Obstructive sleep apnoea and obesity	https://www.ncbi.nlm.nih.gov/pubmed/22956426	IPAQ, logbook	Accelerometer (SenseWear pro 3 armband)	Agreement between the measurement methods was limited. Little agreement between measures – on average the IPAQ overestimated MVPA by 47 minutes and logbook bike 32 minutes compared with accelerometer estimates. These PA measurement tools cannot be used interchangeably.
Helmerhorst et al. (2012) – systematic review	Lifecourse (youth and adults)	https://ijbnpa.biomedcentral.com/articles/10.1186/1479-5868-9-103	Various	Various	Although the majority of PAQs appear to have acceptable reliability, the validity is moderate at best. Importantly, newly developed PAQs do not seem to perform any better than existing instruments in terms of reliability and validity.

Author	Population	Link (URL)	Measure(s)	Comparator	Findings
van Poppel e al. (2012) – systematic review of physical activity questionniares for adults	Adults	https://link.springer.com/article/10.2165%2F11531930-000000000-00000	Various	Various	Fifty-one questionnaires were tested for reliability with only a few of sufficient construct validity and reliability. There is a clear lack of standardization of PA questionnaires, resulting in many variations of questionnaires. No questionnaire or type of questionnaire for assessing PA was superior and therefore could not be strongly recommended above others.
Evenson et al. (2012) - systematic review	Pregnant women	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3419488/	Pregnancy Physical Activity Questionnaire; IPAQ; 7DPAR; occupational questionnaire; modified Kaiser Physical Activity Survey; Norwegian Mother and Child Cohort Study Survey; Leisure-Time Exercise Questionnaire; third	Pedometers & accelerometers	Agreement between questionnaires and objective measurements ranged from slight to fair agreement. Comparisons to other self-reported measurements ranged from substantial to almost perfect agreement. Five studies (42%) assessed test-retest reliability of the questionnaires. Reliabilityranged from substantial to almost perfect agreement. The four studies that used diaries were all assessed for validity

Author	Population	Link (URL)	Measure(s)	Comparator	Findings
			Pregnancy infection and Nutrition (PIN3) Study physical activity and pregnancy questionnaire; STORK physical activity and pregnancy questionnaire		against objective measurements.Results rangedfrom slight to substantial agreement.
Rachele et al. (2012) - narrative review	Youth	https://link.springer.com/article/10.1007%2Fs12519-012-0359-z	IPAQ, PA diary, PDPAR	$VO_{2\ MAX}$, TEE, accelerometer, pedometer	Correlations between criterion (objective) measures of PA, IPAQ, PA diary and PDPAR were eeak to moderate. Accuracy of recall is questionable particularly if the period of recall is lengthy. Youth PA patterns are more variable than adults. Adolescents and children are less likely to make accurate recall estimates due to development differences. Where possible, objective measures of PA should be used in young populations. If PA recall methods are used, data should be interpreted with caution

Author	Population	Link (URL)	Measure(s)	Comparator	Findings
Foley et al. (2012) - systematic review	Youth	https://onlinelibrary.wiley.com/doi/epdf/10.1111/j.1467-789X.2012.00993.x	Use-of-time tools: PDPAR, 3DPAR, physical activity interview, Multimedia Activity Recall for Children and Adolescents (MARCA), Computerised Activity Recall (CAR), Activitygram, Multimedia Activity Recall for Children and Adolescents (MARCA)	Accelerometer, pedometer, HR, doubly-labelled water	Weak to moderate reliability and validity. Use-of-time tools have indicated moderate reliability and validity for the assessment of PA and energy expenditure. Generally, correlation coefficients against validation methods were in the range of 0.30–0.40 (moderate strength), although they ranged widely from 0.16 to 0.88. Weaker coefficients seen in younger participants.
Hoos et al. (2012)	Adult Latinas	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3743722/	Global PA Questionnaire (GPAQ)	Accelerometer (Actigraph GT1M)	The GPAQ was sensitive to change in vigorous leisure time PA and total PA Correlations between PA (intensity ranged from sedentary to vigorous) estimated by the GPAQ and the accelerometer ranged from weak to moderate with no correlation for some intensities (i.e. no correlation between accelerometer and questionnaire

Author	Population	Link (URL)	Measure(s)	Comparator	Findings
					estimates of moderate intensity PA at baseline or 6 months). GPAQ may be a useful tool to detect change in vigorous PA as a result of an intervention but it cannot be used interchangeably with accelerometer estimates of PA.
Celis-Morales et al. (2012)	European and Chilean adults	http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0036345	IPAQ	Accelerometer (ActiTrainer Actigraph)	IPAQ led to sig over-reporting of PA and under-reporting of sedentary behaviour compared with accelerometer derived estimates. Concordance coefficient for accelerometer-derived vs IPAQ-reported activity measures for sedentary behaviour was moderate-to-strong (0.52), but weaker for PA indices (≤0.22 for all measures).
Hart et al. (2011)	Adults	https://insights.ovid.com/pubmed?pmid=20631642	Bouchard Activity Record (BAR)	Accelerometers (ActiGraph GT1M and activPAL)	Significant difference (ActiGraph vs. activPAL, ActiGraph vs. BAR) for summary time spent in sedentary behaviour and time spent walking (ActiGraph vs. activPAL, ActiGraph vs. BAR). Mean agreement ranged from 54% (ActiGraph and activPAL,

Author	Population	Link (URL)	Measure(s)	Comparator	Findings
					walking) to 86.7% (ActiGraph and BAR, MVPA).
Harrison et al. (2011)	Pregnant women (26-28 weeks gestation)	https://ijbnpa.biomedcentral.com/articles/10.1186/1479-5868-8-19	International Physical Activity Questionnaire (IPAQ)	Accelerometer (ActiGraph GT1M)	Accelerometer and IPAQ estimates of total, light and moderate Metabolic Equivalent minutes/day (MET min^{-1} day^{-1}) were not significantly correlated and there was poor absolute agreement. Relative to the accelerometer, the IPAQ under predicted daily total METs (105.76 ± 259.13 min^{-1} day^{-1}) and light METs (255.55 ± 128.41 min^{-1} day^{-1}) and over predicted moderate METs (-112.25 ± 166.41 min^{-1} day^{-1}).

Table 3 Summary of studies (original research and reviews) examining the methodological effectiveness of pedometers for measuring PA.

Author	Population	Link (URL)	Measure	Comparator	Findings
Dowd et al. (2018) - systematic review of reviews	Adults	https://ijbnpa.biomedcentral.com/articles/10.1186/s12966-017-0636-2	Various	Various	The test-retest reliability of pedometer determined steps in a laboratory setting was high across the majority of speeds, but the reliability appeared to weaken at higher speeds. For pedometers, a minimum of 2-4 days of measurement was required to provide a reliable estimate of steps in older adults, while 2-5 days of measurement was required in adults.
Lee et al. (2015)	College age males and females	https://www.ncbi.nlm.nih.gov/pubmed/25517396	Pedometer (Omron HJ-720 T)	Direct observation for treadmill walking and pedometer (Yamax Digiwalker SW-701) for free-living conditions	The Omron pedometer had excellent agreement with direct observation during treadmill walking at a variety of speeds (3.2 - 6.4 km/h). During free-living conditions the association between Omron and Yamax estimated step count was strong. A limitation of this study is that the criterion measure used to estimate steps under free-living conditions performed worse than the Omron under laboratory

Author	Population	Link (URL)	Measure	Comparator	Findings
					conditions when compared with direct observation of step count.
Webber et al. (2014)	Community-dwelling older adults (some with walking aids and some without)	https://journals.humankinetics.com/doi/abs/10.1123/JAPA.2013-0002	Mechanical pedometer (Yamax SW200), piezoelectric pedometer (SC-stepMX)	Direct observation	Participants walked 100m wearing the devices. A limitation of this study is the validity of the pedometers was not assessed under free-living conditions. Lowest error value observed with the SC-stepMX. No significant differences among monitors for those who walked without aids ($p = .063$). However, individuals who used walking aids exhibited slower gait speeds (M - 0.83 m/s, $SD = 0.2$) than non-walking aid users ($M = 1.21$ m/s, $SD = 0.2$, $p < .001$), and for them the SC-StepMX demonstrated a significantly lower percentage of error than the other device. These results support using a piezoelectric pedometer for measuring steps in older adults who use walking aids and who walk slowly
Rachele et al. (2012) -	Youth	https://link.springer.com/content/pdf/10.100	Pedometers (various)	Doubly-labelled water	PA EE calculated from pedometers has a moderate correlation with EE calculated

Author	Population	Link (URL)	Measure	Comparator	Findings
narrative review		[7%2Fs12519-012-0359-z.pdf]()			from the doubly labelled water method and has demonstrated responsiveness to changes in physical activity amongst youth samples. The primary limitation of pedometers is they are unable to record the magnitude of the movement which limits their ability to distinguish between walking, jogging and running. However, newer models provide an estimate of intensity by calculating cadence (steps per minutes), but further research is required to ascertain the reliability and validity of pedometers to estimate PA intensity.
Harrison et al. (2011)	Pregnant women (26-28 weeks gestation)	https://ijbnpa.biomedcentral.com/articles/10.1186/1479-5868-8-19	Pedometer (Yamax)	Accelerometer (ActiGraph GT1M)	Significant correlation between accelerometer and pedometer for estimated daily steps and good absolute agreement with low systematic error (mean difference: 505 ± 1498 steps/day). Compared with the accelerometer, the pedometer appears to provide a reliable

Author	Population	Link (URL)	Measure	Comparator	Findings
					estimate of physical activity (steps) in pregnancy.

Table 4 Summary of studies (original research and reviews) examining the methodological effectiveness of accelerometers and inclinometers for measuring PA.

Author	Population	Link (URL)	Measure	Comparator	Findings
Dowd et al. (2018) - systematic review of reviews	Adults	https://ijbnpa.biomedcentral.com/articles/10.1186/s12966-017-0636-2	Various	Various	Accelerometer had a high level of criterion validity for estimating step count. Although variability was lower for accelerometers, a substantial proportion of studies underestimated energy expenditure compared to DLW when proprietary algorithms or count-to-activity thresholds were employed. Moderate to strong test-retest reliability was observed for activity monitors in free-living environments. However, the reliability of accelerometers attenuated as the duration between measurements increased.

Author	Population	Link (URL)	Measure	Comparator	Findings
					For accelerometers, two days of measurement are recommended for a reliable estimate of steps per day, accelerometer counts per day and intermittent MVPA per day measured, 3 days for a reliable estimate of total PA and time spent in MVPA and 6 days are required for a reliable estimate of continuous 10 minute bouts of MVPA.
Bell et al. (2013)	Overweight/obese pregnant women	https://www.ncbi.nlm.nih.gov/pubmed/23849310	Accelerometer (criterion measure in this study)	RPAQ	Accelerometer feasible and acceptable. Objective methods should be used where possible in studies measuring physical activity in pregnancy.
Calabró et al. (2014)	Adults	https://ijbnpa.biomedcentral.com/articles/10.1186/s12966-014-0119-7	Inclinometer (ActivPAL - uni-axial)	Indirect calorimetry	Significantly underestimate EE during low intensity activities compared with IC by 22.2% during 60 minutes of semi-structured activities
Herman Hansen et al. (2014)	Adults	https://www.tandfonline.com/doi/abs/10.1080/026	Accelerometer (ActiGraph GT1M)	Indirect calorimetry	The ActiGraph GT1M is a valid tool for assessing walking across a wide range of speeds and gradients. However, there is no relationship between activity counts

Author	Population	Link (URL)	Measure	Comparator	Findings
		40414.2013.844347			and energy expenditure during cycling and physical activity is underestimated by 73% during cycling compared to walking
Rachele et al. (2012) - narrative review	Youth	https://link.springer.com/article/10.1007/s12519-012-0359-z	Accelerometers (various)	Various (including doubly-labelled water and indirect calorimetry).	Moderate to strong association between accelerometer estimates of PA and criterion measure (i.e. DLW and indirect calorimetry). Limitations of accelerometers include inability to account for the increased energy cost associated with walking up stairs or an incline, accurately measure activities such as cycling, lifting, or carrying objects, and differentiate well between sitting and standing. Furthermore, there is a lack of standardization in how data is collected, processed and analysed. This lack of standardization limits comparison between studies.

Author	Population	Link (URL)	Measure	Comparator	Findings
Kozey-Keadle et al. (2011)	Overweight, inactive office workers	http://www.umass.edu/physicalactivity/newsite/publications/Sarah%20Keadle/papers/4.pdf	Inclinometer (ActivPAL- (model not specified))	Direct observation	Compared with direct observation, the ActivPAL underestimated sedentary time by 7.7 minutes (2.8%) on average and was more accurate than the ActiGraph (GT3X) which underestimated sedentary time by 16.9 minutes (4.9%). Furthermore, the AP was sensitive to reductions in sedentary time resulting from advice to reduce sitting and increase standing, whereas the ActiGraph and multiple SB questionnaires were not.

Table 5 Summary of studies (original research and reviews) examining the methodological effectiveness of multi-sensors for measuring PA.

Author	Population	Link (URL)	Measure	Comparator	Findings
Bhammar et al. (2016)	Adults	https://www.tandfonline.com/doi/abs/10.1080/02640414.2016.1140220	SenseWear Armband Mini	Indirect calorimetry	The SenseWear armband misclassified activity intensity, generally underestimating time spent in light activities and overestimating time spent in moderate activities. This was due to the consistent overestimation of light activities such as sweeping and loading/unloading boxes.
Calabró et al. (2014)	Adults	https://ijbnpa.biomedcentral.com/articles/10.1186/s12966-014-0119-7	SenseWear Armband (Pro3 and mini) and Actiheart	Indirect calorimetry (portable)	Both models of the SenseWear Armband and the Actiheart provided accurate estimates of EE during light and moderate intensity activities of daily living. The multi-sensor monitors appear to have advantages compared to the standard accelerometers (ActiGraph and activPAL). The SenseWear Mini provided more accurate estimates of total EE during light to moderate intensity semi-structured activities compared to other activity monitors (slightly overestimated EE by 1%).

Author	Population	Link (URL)	Measure	Comparator	Findings
Drenowatz and Eisenmann (2011)	Endurance trained adults	https://link.springer.com/article/10.1007/s00421-010-1695-0	SenseWear Armband (model not specified)	Indirect calorimetry	While providing accurate results for energy expenditure during low-to-moderate intensity physical activities, the SenseWear armband does not provide accurate estimates of energy expenditure at high intensity levels. The threshold for accurate measurements seems to be around an intensity of ten METs.
Johannsen et al. (2010)	Adults	https://europepmc.org/abstract/med/20386334	SenseWear Armband Pro3 and mini	Doubly labelled water (DLW)	Both activity monitors showed good agreement with doubly labelled water measured total energy expenditure, the SenseWear Armband Pro and mini under estimating total energy expenditure by 112 kcal/d and 22 kcal/d, respectively. The SenseWear Armband mini provided estimates that were not significantly different to doubly labelled water and the two measures had an intraclass correlation of 0.85. Both models showed a greater underestimation of energy expenditure at higher total energy expenditure.

Author	Population	Link (URL)	Measure	Comparator	Findings
Spierer et al. (2011)	Adults	https://link.springer.com/article/10.1007/s00421-010-1672-7	Actiheart	Indirect calorimetry	The Actigraph tended to underestimate energy, sometimes by substantial amounts. The Actiheart provided significantly better estimates than the Actical for the activities in which acceleration of the pelvis is not closely related to energy expenditure (card playing, sweeping, lifting weights).
Berntsen et al. (2010)		https://bjsm.bmj.com/content/44/9/657	SenseWear Armband (Pro2)	Indirect calorimetry (portable)	The Sensewear Armband overestimated time in MVPA and underestimated total EE.
Barreira et al. (2009)	College students	https://digitalcommons.wku.edu/ijes/vol2/iss1/7/	Actiheart	Indirect calorimetry, heart rate monitor and electrocardiogram (ECG)	Actiheart heart rate (HR) was similar to HR measured by ECG at all workloads. Actiheart heart rate had strong correlation with HR from the Polar HRM (r = .93) under laboratory conditions, however, there was an overestimation of HR by the Actiheart monitor under free-living conditions. The Actiheart measure of energy expenditure was strongly correlated with measures of indirect calorimetry measures of energy

Author	Population	Link (URL)	Measure	Comparator	Findings
					expenditure, however, it underestimated energy expenditure at the highest workload under laboratory conditions. Overall, the Actiheart was valid at measuring and categorizing intensities of physical activity.

Table 6 Summary of studies (original research and reviews) examining the methodological effectiveness of GPS for measuring PA.

Author	Population	Link (URL)	Measure	Comparator	Findings
Holliday et al. (2017)	Adults	https://europepmc.org/articles/pmc5392135	Accelerometer (ActiGraph GT1M) and GPS (Qstarz BT-Q1000X) worn on a single belt	NA	This study suggests that 12 days of surveillance are needed for a reliable estimate of both MVPA and vigorous physical activity bout minutes in fitness facilities, footpaths, parks, roads, and schools. Minutes of PA in the home environment and commercial locations may be best assessed through other means given the lengthy estimated monitoring time required.
Wieters et al. (2012)	Adults (only 4 participants)	https://journals.humankinetics.com/doi/pdf/10.1123/jpah.9.7.913	Four GPS models	Actual route walked, variations based on position of unit on user's body and variations against a known geodetic point	The Garmin Forerunner appeared to be most accurate (especially in terms of accurate data collected along a known route and data points compared with a geodetic point).
Krenn et al. (2011) - systematic review	NA	https://www.sciencedirect.com/science/article/pii/S0749379711005460#!	GPS (various)	Various	GPS is a promising tool for improving understanding of the spatial context of physical activity. The current findings suggest that the choice of an appropriate device and efforts to maximize participant adherence are key to improving data quality, especially over longer study periods.

Author	Population	Link (URL)	Measure	Comparator	Findings
Maddison and Mhurchu (2009) - narrative review	NA	https://ijbnpa.biomedcentral.com/articles/10.1186/1479-5868-6-73	GPS (various)	Various	Controlled studies focused on validating GPS systems against chronometry for walking, running and cycling speeds. Studies show GPS provides an accurate measure of walking, running and cycling speed. Under free-living conditions GPS has been shown to provide accurate and reliable measures of location. Free-living studies tend to integrate GPS data with accelerometer data. These studies have shown GPS information is precise enough to determine where physical activity and sedentary behaviours are performed. GPS can also be used to map active travel routes to understand how the built environment affects physical activity. Combining GPS and accelerometer data also shows promise for determining the mode of physical activity and a study showed this method accurately classified activities 91% of the time. However, no studies have shown that GPS alone is a reliable and valid measure of physical activity.

Table 7 Summary of studies (original research and reviews) examining the methodological effectiveness of heart rate monitors for measuring PA.

Author	Population	Link (URL)	Measure	Comparator	Findings
Rachele et al. (2012) - narrative review	Youth	https://link.springer.com/article/10.1007/s12519-012-0359-z	HR monitoring, combined with accelerometry (to measure energy expenditure)	Energy expenditure measured by whole-room calorimetry	Acceptable method (but increased burden of wearing more than one device must be considered). Age, body size, emotional stress, aerobic fitness, and body composition influence the relationship between heart rate and EE. HR lags behind change in PA and remains elevated after PA has stopped which may mask sporadic changes in PA.
Myers et al. (2014)	Older adult patients with abdominal aortic aneurysm disease	https://journals.humankinetics.com/doi/abs/10.1123/japa.2012-0133	HR monitor	Hourly activity logs, 3DPAR, accelerometer, pedometer,	Mean energy expenditure kcals/day) was 1687±458 (HR), 2068±529 (accelerometer) and 1974±491 (3DPAR). Sig differences between HR and accelerometry and HR and 3DPAR but not between accelerometry and 3DPAR. Highest agreement was between accelerometry and 3DPAR (coefficient of variation 0.86).

Table 8 Summary of studies reporting the accuracy of smartphone PA measurement ***(adapted from Bort-Roig et al. (2014))***.

Author/country	Measurement technology	Position	Algorithm	Behaviours measured	Key findings (% accuracy)
Anderson et al. (2007), UK	App used patterns of fluctuation in mobile signal strength and number of cell phone tower locations	Normal Smart Phone use	Artificial Neural Network Hidden Markov Models	Stationary, walking, or travelling (car, bus, or train)	ANN: stationary (83 %), walking (87 %), and travelling (73 %) HMM: stationary (92 %), walking (80 %), and travelling (74 %)
Donaire-Gonzalez et al. (2013), Spain	In-built accelerometer	On a belt attached to the waist	Freedson's MET prediction algorithm Phone-based vertical axis g-force converted to ActiGraph counts/min via linear regression	Time spent in sedentary, light, moderate, and vigorous PA, and EE (METs) vs. estimates from the ActiGraph GT3X	Mean difference between ActiGraph GT3X and the app was 2.24 % (95 % CI 0.76–3.72) for the duration of active time (>1.5 METs) and 0.07 METs (95 % CI 0.04–0.1) for PA intensity. Measures of vigorous PA showed a tendency to underestimate the duration in vigorous PA
Gao et al. (2009), USA	In-built accelerometer	Smart Phone on the chest, waist at the front, and at the side of the hip, upper pocket	Artificial Neural Network on raw tri-axial accelerometer signal with band pass filtering	Walking or running	Held still in front of the chest when walking (100 %) and running (100 %) Front waist walking (98–100 %) and running (98 %); side waist walking (96–100 %) and running (94–97 %) Upper pocket while walking (100–105 %) and running (97–98 %)

Author/country	Measurement technology	Position	Algorithm	Behaviours measured	Key findings (% accuracy)
He and Li (2013) ,China	In-built accelerometer, gyroscope, and magnetic field sensor	Placed inside an adjustable band attached to the chest	A binary hierarchical classifier system was used to recognize 14 activities	Static activities (sitting, lying, standing), transitions, dynamic activities (walking, upstairs, downstairs, running, jumping), and falls	Static activities (98 %), transitions (94 %), dynamic activities (91 %)
Ketabdar and Lyra (2010), Germany	In-built accelerometer	Trouser pocket	Gaussian Mixture Model	Stationary, walking, or running	Stationary (96 %), walking (93 %), and running (93 %)
Khalil and Glal (2009), UAE	In-built accelerometer	Upper pocket, lower pocket and side pocket, hand-held, bag	Dynamic Peak Detection Algorithm	Walking	Upper pocket (88–99 %), lower pocket 86 %, side pocket (84–92 %), hand-held (97–100 %), and bag (92–100 %)
Mattila et al. (2009), Finland	External accelerometer and heart rate monitor with input uploaded to the app via Bluetooth	Integrated accelerometer and electrode unit worn on chest	Peak Detection Algorithm after Fast Fourier Transformation of acceleration signal Regression-based prediction of energy expenditure	Walking Energy expenditure	Step rate detection was very accurate at step rate of 90 steps per minute or more EE (METs) estimation was fairly accurate at walking speeds (0.996, $p < 0.01$), unreliable results when running
Lee et al. (2011), South Korea	In-built accelerometer	Waist, trouser pocket	Fuzzy C means clustering algorithm	Lying, sitting, standing,	Lying (100 %), sitting (96 %), standing (98 %),

Author/country	Measurement technology	Position	Algorithm	Behaviours measured	Key findings (% accuracy)
				walking, running, or falling	walking (98 %), running (100 %), and falling (99 %) Overall activity classification (98 %), waist (99.6 %) and trouser pocket (99.1 %)
Zhong et al. (2010), China	External accelerometer with input uploaded to the app via Bluetooth	Foot	Dynamic Peak Detection Algorithm	Walking, running, climbing stairs, or gait transitions	Walking (100 %), running (96 %), up-stairs (98 %), and gait transitions (95 %)
Wu et al. (2012) ,USA	In-built accelerometer and gyroscope	Front shorts pocket, armband	Multiple algorithms evaluated. Decision tree, Artificial Neural Network, naive Bayes, logistic regression, k-nearest neighbour	Walking, climbing up and down stairs, jogging, sitting	The k-nearest neighbour classifier achieved high accuracies for walking at different paces (90–94 %), jogging (91 %), and sitting (100 %). Activity recognition was lowest for climbing stairs (52–79 %)

Table 9 Summary of studies (Qualitative and Randomized Controlled Trials) of Smartphone Applications for Promoting PA ***(adapted from Coughlin et al. (2016))***.

Study	**Sample**	**Design**	**Results**	**Limitations**
Casey et al. (2014)	12 participants (mean age 42 yrs, 75% female) in 3 primary care centres in Ireland	Semi-structured interviews	Four themes emerged from the analysis: transforming relationships with exercise, persuasive technology tools, usability, and the cascade effect. The app appeared to facilitate a sequential and synergistic process of positive change, which occurred in the relationship between the participants and their exercise behaviour.	Non-randomized design, small sample size
Rabin and Bock (2011)	15 sedentary adults in Rhode Island	Formative study	The users have preferences with regard to PA app features related to PA, including provision of automatic tracking of PA (e.g., steps taken and calories burned), tracking of progress toward PA goals, and integrating a music feature. Participants also preferred that apps be flexible enough to be used with several types of PA and have user-friendly interfaces.	Non-randomized design, small sample size
Middelweerd et al. (2015)	30 Dutch students aged 18–25 yrs	Focus groups	Participants most often used social networking apps (e.g., Facebook or Twitter), communication apps (e.g., WhatsApp), and content apps (e.g., news reports or weather forecasts). They preferred a simple and structured layout and a companion website for detailed information about their accomplishments and progress. They preferred	Non-randomized design

			apps that coached and motivated them and provided tailored feedback toward personally set goals.	
Vandelanotte et al. (2013)	30 middle-aged men in Queensland, Australia	Focus groups	The men supported the use of the Internet to improve and self-monitor PA and dietary behaviours provided the interventions were quick and easy to use. Participants preferred smartphones over regular mobile phones.	Non-randomized design
Morrison et al. (2014)	13 adults (6 male, 7 female, median age 27 yrs) in Southampton, United Kingdom	Telephone interviews. The participants had an Android smartphone	Access to the app was associated with an increase in participants' awareness of their PA goals (P=0.03). Participants used the POWeR website for similar amounts of time during the weeks when POWeR Tracker was (mean 29 minutes, SD 31 minutes) and was not available (mean 27 minutes, SD 33 minutes).	Non-randomized design, small sample size
Kirwan et al. (2013)	12 Australian adults (90% white, 10% Asian), of whom 6 were women (mean female, 49% African-American) in	Qualitative and quantitative study with premodification testing to identify usability intensive counselling intervention, 2) intensive counselling plus smartphone intervention, 3) a less intensive counselling plus smartphone intervention,	Four usability themes emerged from the data related to design, feedback, navigation, and terminology. Design improvements to the app resulted in a reduction in the problems experienced and a monitoring smartphone group tended to lose more weight than other groups (5.4 kg and 3.3 kg, respectively). Of those who completed the 6-month follow-up, 64% of participants in the intensive counselling plus self-monitoring smartphone group and 40% in the less intensive counselling plus self-monitoring smartphone	Non-randomized design, small sample size

	Baltimore, MD who used an iPhone or Android smartphone	and 4) smartphone intervention only. The outcome measures included weight, body mass index (BMI), waist circumference, and self-reported PA and dietary intake	group achieved greater than or equal to 5% decrease in their body weight. In contrast, only 25% in the counselling only group and 20% in the self-monitoring smartphone only group achieved at least a 5% weight loss. Females were more likely to lose weight compared to males ($P = 0.005$).	
Carter et al. (2013)	128 overweight volunteers (77% female, 91% white, mean age 42 yrs) in Leeds, United Kingdom	6-month randomized controlled trial (pilot study) comparing smartphone app, website intervention, and paper diaries. The outcome measures were change in weight, body mass index, and body fat. The smartphone app uses an Android system.	Mean weight change at 6 months was −4.6 kg (95% CI −6.2 to −3.0) in the smartphone app group, −2.9 kg (95% CI −4.7 to −1.1) in the diary group, and −1.3 kg (95% CI −2.7 to 0.1) in the website group. Change in BMI at 6 months was −1.6 kg/m^2 (95% CI −2.2 to −1.1) in the smartphone group, −1.0 kg/m^2 (95% CI −1.6 to −0.4) in the diary group, and −0.5 kg/m^2 (95% CI −0.9 to 0.0) in the website group. Change in body fat was −1.3% (95% CI −1.7 to −0.8) in the smartphone group, −0.9% (95% CI −1.5 to −0.4) in the diary group, and −0.5% (95% CI −0.9 to 0.0) in the website group.	
Duncan et al. (2014)	301 adult men in Queensland, Australia ages 35 to 54 yrs	9-month randomized trial comparing mobile phone-based intervention to print-based intervention. The outcome measures were	Participants reported an increased number of minutes and sessions of PA at 3 months and 9 months. The participants in the IT-based arm were less likely to report that 30 minutes of PA per day improves health and more likely to report that vigorous intensity PA 3 times per week is essential.	

		total minutes of PA and total sessions of PA		
Turner-McGrievy and Tate (2011)	96 overweight adults (75% women, 20% African American, ages 25 to 45 yrs) in Raleigh-Durham, NC	Randomized trial comparing a combination of podcasting, mobile support communication, and mobile diet monitoring. The outcome measures included number of days each participant monitored their weight and weight loss. The participants had an Internet-capable mobile devices: iPhone, iPod Touch, BlackBerry, or an Android-based phone.	Adjusting for randomized group and demographics, users of the physical activity app self-monitored exercise more frequently over the 6-month study (2.6±0.5 days/week) and reported greater intentional physical activity (196.4±45.9 kcal/day) than non-app users (1.2±0.5 days/week physical activity self-monitoring, $p<0.01$; 100.9±45.1 kcal/day intentional physical activity, $p=0.02$). At 6 months, users of the physical activity app also had a lower BMI (31.5±0.5 kg/m2) than non-users (32.5±0.5 kg/m2; $p=0.02$).	
Martin et al. (2015)	40 adults in Baton Rouge, LA (25 < BMI < 35 kg/m2, 82.5% female, mean age 44.4 yrs)	12-week randomized controlled trial comparing smartphone app to health education control group. The outcome measure was weight loss.	Weight loss was significantly larger in the smartphone app group	Small sample size