Assessment of physical activity in epidemiological studies: Are questionnaires obsolete in the era of accelerometry?

Erfassung körperlicher Aktivität in epidemiologischen Studien: Sind Fragebögen überholt im Zeitalter der Akzelerometrie?

Abstract

Background: The rapid development in technology promotes the increasing use of electronic activity monitors to assess physical activity in large-scale epidemiological studies. Our aim was to explore and discuss both subjective and objective methods assessing physical activity.

Methods: Based on a thorough literature search, major strengths and limitations of questionnaires and electronic activity monitors in assessing physical activity are elaborated and discussed, taking into consideration physical activity in all its complexity. Important research questions and the aim and scope of physical activity assessment for next-generation research are defined.

Results: Questionnaires can provide details and background information of physical activity, including type of activity, and can identify the activity behaviour patterns that underlie measurable endpoints such as energy expenditure, cardiorespiratory fitness, muscular strength, or certain biomarkers. They can differentiate well between settings and enable retrospective assessment. Electronic activity monitors, on the other hand, are rapidly becoming better at assessing energy expenditure and are good at quantifying the amount and intensity of physical activity and sedentary behaviour. Both methods, however, also have their weaknesses. While questionnaires may not be accurate, due to recall bias or incomplete assessment of all activity domains, electronic activity monitors are not able to provide information about setting, exact type and mode of activity, and cannot recognize and reliably assess resistance exercise or activities with or without carrying weights.

Conclusions: Since physical activity is multidimensional and complex, no single method is suitable to capture all aspects and domains. Both methods have their strengths and limitations and do not compete with each other, but should be seen as complementary tools that assess distinct aspects of physical activity. Studies should therefore rely on both methods to enable improved public health recommendations for this complex lifestyle factor.

Keywords: questionnaire, accelerometer, epidemiology, reliability, validity, physical activity

Zusammenfassung


Methoden: Basierend auf einer umfangreichen Literaturrecherche und unter Berücksichtigung der Komplexität von körperlicher Aktivität wurden die wichtigsten Stärken und Schwächen von Fragebögen und elektronischen Aktivitätsmessgeräten ausgearbeitet und erörtert. Des Weiteren
wurden wichtige Forschungsfragen sowie Ziele und Umfang zukünftiger Aktivitätserfassungen definiert. 


**Schlussfolgerung:** Körperliche Aktivität ist ein mehrdimensionales und komplexes Verhalten. Es gibt keine Methode, die alleinstehend alle Aspekte und Bereiche körperlicher Aktivität hinreichend erfasst. Beide Methoden haben ihre Stärken und Schwächen und erfassen verschiedene Aspekte von körperlicher Aktivität. Somit konkurrieren sie nicht miteinander, sondern sind als komplementäre Erhebungsinstrumente zu sehen. Zukünftige Studien sollten sich bei der Erfassung dieses komplexen Lebensstilfaktors auf beide Methoden stützen, um verbesserte gesundheitsfördernde Empfehlung zu ermöglichen.

**Schlüsselwörter:** Fragebogen, Akzelerometer, Epidemiologie, Reliabilität, Validität, körperliche Aktivität

**Introduction**

Questionnaires have been the undisputed instrument in assessing physical activity (PA) for several decades [1], [2], [3], [4]. They have provided the basis for most of the currently available knowledge on the association between PA and numerous diseases (e.g. cardiovascular diseases, diabetes, colon cancer, post-menopausal breast cancer, osteoporosis, and dementia). Further methods used for the assessment of PA include doubly labelled water, direct/indirect calorimetry, accelerometer (1–3 axial), heart rate monitoring (HRM), combined heart rate and accelerometer devices, pedometers, PA diaries/logs, and direct observations.

Numerous improvements in technology over the last decade have enhanced the applicability of electronic activity monitors (EAMs) and increased their implementation in large-scale epidemiological studies [5], [6], [7], [8], [9], [10], [11], [12], [13]. EAMs include common accelerometer devices that are able to measure the acceleration of a person as well as multi-functional devices which additionally measure heart rate. Due to the limitations inherent in using questionnaires, such as recall bias and self-reporting [14], [15], [16], their application has been increasingly questioned in the recent past.

The aim of this paper is to discuss the major strengths and limitations of EAMs versus questionnaires in assessing PA in (large-scale) epidemiological studies of free-living populations. Important future research questions will be highlighted and it will be argued that the aim and scope of PA assessment need to be redefined for next-generation research. This will lead to the discussion on how PA might be best assessed in order to tap the full potential of this complex lifestyle factor in improving individual and public health.

**Methods**

Literature searches were conducted in PubMed. Search terms included “physical activity assessment”, “accelerometer/accelerometry”, “questionnaire”, “subjective/objective measurement methods”, “motion sensor”. These were additionally combined with “validity AND ...” or “reliability AND ...”. Furthermore, a manual literature search, based on the ‘snowball’ principle, was conducted in individual literature lists. From the identified literature strength and limitations, and practical and feasibility issues of the different PA assessment methods were extracted, summarized and critically discussed.
General aspects of physical activity assessment

Physical activity is a complex and multidimensional behaviour that can be subdivided into a number of dimensions, such as frequency, duration, intensity, and type of activity, as well as various domains (e.g. occupation, household, locomotion, leisure time, or sports). In the past decades and especially in the current era of accelerometry, it has, however, become common to focus more on overall or average time spent physically active [17, 18, 19]. This makes it difficult to differentiate between vigorous exercise once a week followed by sedentariness the rest of the week and regular moderate PA such as cycling to work. It is furthermore difficult to assess seasonal differences or irregular behaviour such as occasional vigorous running or mountain biking. Yet different PA patterns might have different health effects, even if the resulting overall energy expenditure (EE) is similar. It is still unclear which type, intensity, and temporal pattern of activity is most beneficial to health [18], [20], [21]. Thus, the scientific basis for current recommendations (e.g. the WHO’s “Global Recommendations on Physical Activity for Health” in 2010 or the CDC’s “Be Active Your Way: A Guide for Adults” in 2008) is limited. Systematic investigations of health benefits from different activity patterns are lacking. In most prospective observational studies people were generally asked about the total or average amount of PA performed over a specific time period (e.g. EPIC study, Nurses’ Health Study, MONICA Project). This impedes a precise differentiation between activities conducted in a single, long session versus those conducted in multiple short sessions spread over the day or throughout the week. It is unclear whether there is a relationship between the duration of structured exercise sessions and fitness responses, when total volume is held constant, especially for PA with vigorous intensity [21]. The effect of irregular or acute periods of exhaustive activity may differ from the effect of constant training or an overall active lifestyle, which is why aim and scope of PA assessment must be redefined. Without detailed investigation on the effects of different PA patterns, it will remain difficult to give concrete public health recommendations.

Electronic activity monitors

The following section will present the main strengths and limitations of EAMs, such as accelerometers, and a combination of accelerometers and HRMs. Accelerometers measure the acceleration of a person in up to three axes. Signals are summarized as activity counts or vector magnitude units and can also be transformed into an estimate of EE by using one of a variety of algorithms [22]. A combination of such devices with electrocardiography electrodes offers the additional information of heart rate. An overview of the main facts and key attributes of both methods are summarised in Table 1 and Table 2.

Accelerometers

Strengths

Accelerometers play an increasingly important role in the assessment of PA, especially in documenting the pattern of light-, moderate-, and vigorous-intensity activity throughout the day [3]. Today’s devices provide estimates of EE as well as the amount, frequency, intensity, and duration of body movements. The collection of these data causes minimal restrictions to the person [2], [23] and allows precisely timed measurements [14], [24]. A number of devices are also able to assess some activities that have proved to be difficult to measure by questionnaires. This has been shown regarding walking-EE in free-living subjects [25], [26], [27] and regarding a number of activities including sitting, standing, housework, and walking on a treadmill [28]. Some newer devices (e.g. the ActiGraph GT3X) allow the measurement of further specific activities with the help of body position, such as sitting, standing, lying down, or non-wear time [29].

For epidemiology, the ability to measure routine, moderate activities is most important [24], as well as to differentiate between „sedentary behaviour“ and activity conducted with light, moderate, or vigorous intensity. Accelerometers can not only assess free-living activities, which can cause health benefits similar to structured exercise, but also differentiate between activities such as walking or running and provide information on varying intensity levels. A number of newer devices are also able to distinguish sedentary activity from not wearing the monitor. An additional advantage is that there is a body of literature on the validity and reliability of accelerometers using the cantilever beam technology [14]. A current study by Van Remoortel et al. [22], comparing six state-of-the-art piezoelectric accelerometers to indirect calorimetry, showed good results regarding their validity for everyday tasks such as walking, stair climbing, sweeping the floor, etc. However, only correlations were calculated (minute-by-minute correlations: r=0.73 – 0.82; mean correlations: r=0.45 – 0.76). Measures of agreement (Bland Altman Analyses) were not possible, as not all devices could convert their outcome into EE. Additionally, the study was conducted in a laboratory setting, which reduces its generalisability.

Due to numerous improvements over the past years accelerometers have become cheaper, smaller, lighter, with improved battery performance as well as memory capacity, and are now quite user-friendly. They have also improved in evaluating data due to enhanced statistical methods such as better algorithms [30], artificial neural networks [31], [32], [33], the Gaussian Process [34], and branched equation modelling [35]. These are good reasons for their use in large-scale observational cohort studies. There are, however, also a number of limitations.
<table>
<thead>
<tr>
<th>Aspects</th>
<th>Electronic activity monitors (EAM)</th>
<th>Questionnaires</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy expenditure (EE)</td>
<td>• assessment improved over the past decade</td>
<td>• EE assessment of a whole day is laborious and difficult</td>
</tr>
<tr>
<td>Activity type</td>
<td>• difficulties in distinguishing different types of physical activity (PA) and of sports</td>
<td>• provide more detailed information about the type of activity and enable a more detailed investigation of PA</td>
</tr>
<tr>
<td></td>
<td>• cannot reliably differentiate activities with or without carrying weights</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• enable measurement of the intensity and frequency of activity in real time</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• differentiation between sitting, lying, and standing is occasionally possible</td>
<td></td>
</tr>
<tr>
<td>Activity patterns</td>
<td>• impossible to measure complex activity patterns</td>
<td>• allow the gathering of detailed information about activity behaviour</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• can discover and distinguish complex long-term activity patterns</td>
</tr>
<tr>
<td>Settings</td>
<td>• not measurable</td>
<td>• can assess the environment participants are active in and how it influences their activity level</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• enable differentiation between domains</td>
</tr>
<tr>
<td>Practical issues</td>
<td>• easily measure solid data</td>
<td>• easy to use</td>
</tr>
<tr>
<td></td>
<td>• difficult to interpret data correctly without any further information (e.g. activity log)</td>
<td>• enable plausibility checks</td>
</tr>
<tr>
<td></td>
<td>• standards do not yet exist. It is difficult to compare old and newly assessed data and impossible to guarantee comparability within or between models</td>
<td>• investigations of validity are limited (no gold standard)</td>
</tr>
<tr>
<td></td>
<td>• not all EAMs can recognize if the device was worn by a participant or carried around in a backpack (unreliable data)</td>
<td>• use of computerized questionnaires will ease their use and help to collect data faster, with less effort and more accuracy</td>
</tr>
<tr>
<td>Time window</td>
<td>• unable to assess past-year or lifetime PA</td>
<td>• enable the retrospective assessment of PA, or lifetime PA, as well as the influence of important or profound events such as illness, injuries, etc.</td>
</tr>
<tr>
<td></td>
<td>• good in assessing real-time activity; can easily measure the frequency and duration of activity over a total week</td>
<td></td>
</tr>
<tr>
<td>Inactivity</td>
<td>• good in assessing inactivity, although data can get mixed up with non-wearing time, low activity or activities such as sleeping</td>
<td>• often inaccurate due to recall bias</td>
</tr>
<tr>
<td></td>
<td>• improvement in this aspect is necessary</td>
<td>• improvement in this aspect is necessary</td>
</tr>
</tbody>
</table>

Table 1: Overview of the main facts regarding aspects of physical activity assessment.
Table 2: Key attributes of the individual assessment methods

<table>
<thead>
<tr>
<th>Assessment method</th>
<th>Questionnaire</th>
<th>Accelerometry device</th>
<th>Combined heart rate and movement sensor</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Assessment of…</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>energy expenditure</td>
<td>X</td>
<td>++</td>
<td>+++</td>
</tr>
<tr>
<td>type of activity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. endurance training</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
</tr>
<tr>
<td>2. resistance training</td>
<td>+++</td>
<td>X</td>
<td>+</td>
</tr>
<tr>
<td>physical activity pattern</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. for a short period</td>
<td>++</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>2. for a long period</td>
<td>++</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>past activity</td>
<td>+++</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>inactivity</td>
<td>X</td>
<td>+++</td>
<td>+++</td>
</tr>
<tr>
<td>settings</td>
<td>+++</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td><strong>Practical issues</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>objective</td>
<td>X</td>
<td>+++</td>
<td>+++</td>
</tr>
<tr>
<td>non-reactive*</td>
<td>+++</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>easy to use</td>
<td>+++</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>feasible for large-scale studies</td>
<td>+++</td>
<td>++</td>
<td>+</td>
</tr>
</tbody>
</table>

X=poor; ++=acceptable; +++=excellent  
* relates to the influence a method has on the individual behaviour due to the measurement process

Limitations

One crucial limitation of current accelerometers is their inability to assess PA for a longer period of time (several months up to a year) due to technical and practical reasons. This prohibits the assessment of long-term patterns, seasonal changes, or irregular (e.g. weather-dependent) activities, and the selected wearing period may not be representative for the usual activity. Most importantly, it is still not possible to truly detect contextual information which includes insight about the type of activity, such as soccer and gymnastics, as well as information about the behavioural setting within which participants engage in a given behaviour (e.g., at home or work, sitting in a car) [36]. Although some devices claim to have very good algorithms for predicting PA already, they are also still not able to assess certain basic or everyday behaviour, such as walking on an incline, carrying loads, resistance exercise, cycling, upper body movement [1], [5], [6], [14], [22], [23], [24], [25], or swimming [9], which can result in an underestimation of EE.

The preferred location of the sensor on the body depends on the participant, the device, and the research question. With devices for the waist, wrist, ankle or thigh, there are many possibilities to choose from. It must also be taken into consideration that participants have varying movement patterns. Whereas children are very active, running and jumping a lot, older people perform more sedentary activities. For this reason children might be better off with a waist accelerometer, whereas PA in older people might be underestimated using such a location/device. However, the most common location is the waist [1], [14]. This complicates the assessment for large-scale studies, which often include a wide range of age groups (e.g. the German National Cohort with an age range from 20–69). According to Strath et al. [37], careful consideration is needed when certain populations, such as children, older adults or participants with functional limitations, are to be observed by objective monitoring methods. These methods are not “one-size-fits-all”. Although a number of new devices no longer need to be tested and calibrated, for many devices this procedure is still unavoidable to ensure proper functioning, especially when using a cantilever beam [14]. Even when devices are maintained and handled properly, real-life problems such as software/hardware malfunctioning or external circumstances (e.g. extreme temperatures) might influence the device in a way that alters data or even makes them useless [38], [39].

Another major limitation of EAMs is the frequently observed fact that participants are influenced by the awareness that they are being monitored. This can lead to a more socially acceptable behaviour, such as being more active than usual [25], [40], [41], [42]. Regarding the evaluation of the collected data there are major differences within or between models and the algorithms used. This complicates the comparison of results from different devices and may lead to unreliable data. Furthermore, industry standards are still lacking to-date [1], [5], [16], [37]. Although newer devices as well as new software are quite easy to use, it still requires some skill and experience to process and interpret the large amount of raw data correctly.

Cut points, as well as counts, are still quite arbitrary units. For sedentary activities cut points range from 100 to 800 counts per minute and variability is even larger for moderate intensity activities, ranging from 1,900 to 8,200 counts per minute [1], [43], [44], [45], [46]. Additionally, it is still unclear whether the same cut points that were derived from healthy adults can be applied to populations with functional limitations [37] and are applicable for all age groups [5].
New data processing techniques allow the collection of raw data and could enable researchers to successfully identify activity types [47]. However, such methods are still limited and there is a huge difference between the identification of laboratory-based or free-living activities. Applying laboratory-based equations or thresholds to everyday life settings can lead to systematic errors [48], [49], [50]. Although a large number of algorithms are already available, it is still difficult to determine which one is most precise [31].

Another point that must be considered, especially in large-scale epidemiological studies, is the comparability of data over a longer period of time. Cohort studies are usually conducted over decades. However, unless a study uses only one specific model (and only one model version) over this period of time, there is no guarantee that the derived data is comparable. According to Intille et al. [51], few measurement devices in use today will be on the market 10 years from now (in an identical form). It therefore cannot be guaranteed that data collected by newer devices will be comparable with previously collected data. However, the support of old data is very important for long-term studies. Finally, there are also higher costs and logistical demands associated with implementing objective measures in large-scale studies [5], [36]. Devices have to be prepared by researchers (charged, handed out, collected, etc.) and participants have to meet the researcher, get instructions, keep a diary, and remember to take it off if it is not waterproof.

**Combination of accelerometer and heart rate monitor**

**Strengths**

A combination of accelerometer and HRM provides more accurate data on PA. This is due to a combination of the advantages of each device, which negates some of the disadvantages of each method used alone, as well as the fact that measurement errors are not correlated [1], [35], [52]. Specific activities such as cycling, walking up or down-hill, carrying weights, upper-body work, etc., which cannot be assessed by accelerometer alone, become measurable with the addition of an HRM [53], [54]. While accelerometers have difficulties recognizing non-wearing time, HRMs are able to verify whether the device is worn by detecting a heartbeat. On the other hand, while HRMs have difficulties in measuring the EE at lower intensities, accelerometers are most accurate at that level of activity [1].

Combined devices are reliable and valid for walking and running, which could be verified for the Actiheart device in adults and children. Other free-living activities are, however, still difficult to measure [53], [55]. A study conducted by Crouter et al. [56] evaluated activity energy expenditure (AEE) of the Actiheart device during 18 structured activities (washing dishes, stationary cycling, etc.) and found similar estimates of AEE compared to a portable metabolic system (Cosmed K4b2) measuring oxygen consumption (VO2).

**Limitations**

For a number of reasons, combined devices are not feasible for large-scale studies. First of all, they are usually in a higher price range (e.g. the Actiheart costs about 1000, – €) [58]. As electrodes must be attached to the body, study participants often suffer from skin reactions or refuse to wear the pads for a longer time, even though some study centres use a special user-friendly type instead of the standard self-adhesive electrocardiogram pads. Depending on the device, it can also take some effort to attach and calibrate the device correctly. Where it is not possible to implement a single-piece device such as the Actiheart, participants must bear the encumbrance of wearing two devices. Despite the advantage of having raw data, it is a challenge to analyse these large data sets [1], limiting its use to smaller studies [56]. The difficulty of cleaning and interpreting the data in the right way might require expert guidance in this measurement method and enough computing power to carry out the calculations in reasonable time.

It should also be taken into consideration that the use of combination devices in free-living activities often produces different kinds of measurement noises, which are still quite difficult to avoid [34]. For some devices (e.g. Actiheart), a waterproof sticking plaster can be used to reduce noise during swimming [59]. However, study participants must keep this in mind, which again constitutes an annoyance. Combined devices may influence the activity behaviour of a participant even more than a single accelerometer. Especially if the device is located at the wrist or unpleasant to wear (e.g. due to itching), awareness of being monitored might be high. Additionally, discomfort due to the device may increase non-wearing time.

**Questionnaires**

In the following section, the main strengths and limitations of questionnaires will be presented. In contrast to EAMs, questionnaires can assess characteristics of PA that took place in the past [60]. Depending on focus, length, and detail, they can also assess various information on physical activity/inactivity in the present and do not have to focus only on EE. An overview of the main facts and a summary of the key attributes of this method can be found in Table 1 and Table 2.
Strengths

A general and well-known advantage of PA assessment by questionnaire is the cost-effectiveness, user-friendliness, and suitability for large-scale epidemiological studies [2], [16], [61], [62]. Data collection is less time-consuming than in other measurement methods, as it takes study participants only minutes to fill out the questionnaire, compared to objective measurements methods, which need time to be fixed and maintained and then must be worn for at least 7 days [16]. There is also no need for expensive equipment and its maintenance.

A major advantage of questionnaires is the ability to assess (long-term) behaviour patterns and to differentiate between types of PA [16] which up until now is not possible with the objective measurement methods mentioned above. They also enable a profounder qualitative assessment of physical activity as well as of inactivity, and can distinguish between important domains, such as, e.g. occupation, household, locomotion, leisure time, or sports [1], [19], [62], [63], [64], [65], [66]. They furthermore allow us to shed light on past activities and motives to be physically active/inactive, both not assessable with electronic devices.

Asking different kinds of questions provides a more detailed picture of the participants’ PA, including, besides activity patterns, also the behavioural setting within which participants engage in a given behaviour, environmental factors influencing activity, as well as their motivation and preferences. The results enable researchers to give more concrete public health recommendations and to identify important targets for health interventions, both necessary in the fight against inactivity.

Asking direct questions also helps to verify the data participants deliver. Missing or implausible data can be queried and corrected, and plausibility checks can be conducted using questions of control, which is not possible with accelerometers.

With the rise of computer-based questionnaires, this method can now be administered in a more user-friendly manner. According to Warren et al. [1], an electronic questionnaire can be standardised but remains at the same time flexible enough to offer explanatory material, prompts, error corrections, menus, branches and skips. Compared to objective measurement methods, the input is corrected in real-time, thus eliminating coding errors and offering the possibility of immediate scoring, reporting, and interpretation of results. Growing computer networks and the availability of the internet facilitate the participation of numerous individuals at the same time with minimum additional costs.

Limitations

The major limitation of the questionnaire is that the collected information is in general subjective and prone to recall bias [14], [15], [63]. Answers might be influenced by social desirability or cultural factors [46], [63], [67], [68]. Many questionnaires are designed for a wide age group, with the risk of inaccuracy when used in a younger or older population, particularly due to the underestimation of the performance of light and moderate intensity activities [5], [40], [69], [70]. PA assessment of younger and older groups in general is not easy: children usually have irregular activity patterns, which results in short rather than more sustained periods of activity [1], [71]; older people might be influenced by fluctuations in health status and mood, depression, anxiety or cognitive ability [15], or by problems with memory and cognition [72], [73].

A further disadvantage of questionnaires is their low reliability and validity measuring EE. According to a recently published review by Helmerhorst et al. [74], only a small number of physical activity questionnaires (PAQs) showed acceptable to good results for both reliability and validity when compared with objective criterion measurement of PA. Despite numerous validation studies regarding PAQs for activity EE estimation, it remains difficult to draw any firm conclusion about validity of existing PAQs [75]. Westerterp [2] lists a number of studies which show low correlation, with systematic underestimation, overestimation, or agreement at a group level with considerable error on an individual level, when compared to doubly labelled water.

Additionally, it is difficult to transfer the results of validity studies between different types of population, ethnic groups, or other geographical regions [1], [74]. It would thus be necessary to validate each questionnaire for different subgroups of the population in as many age groups as possible.

Discussion

We have elaborated strength and limitations of questionnaires and electronic devices for the assessment of PA. Questionnaires are good in providing details and context of behaviour patterns, including type of activity, and can differentiate well between settings. Electronic devices, on the other hand, are rapidly becoming better at assessing EE. They are good at quantifying the amount and intensity of PA and the amount of sedentary behaviour. Both methods have their strengths and limitations, as summarized in Table 2. Questionnaires are thus not obsolete but should be considered as a complimentary tool to EAMs.

In the past, assessment of PA was mainly focussed on estimation of EE. Yet, it must be recognized that PA is a behaviour, which in turn results in many different effects on the body. For this reason, the aim and scope of PA assessment need to be redefined for future research. Since PA is multidimensional and complex, and given the derived characteristics of the different assessments, no single method is suitable to capture all dimensions and domains. There is no perfect assessment method that covers all aspects of PA, as can be concluded from Table 1, which gives an overview of the main facts regard-
ing aspects of physical activity assessment. Depending on the required information, the respective method should be implemented.

In the following sections, the major strengths and limitations of both assessment methods will be discussed, separately for the individual aspects of PA. A summary is given in Table 2.

Regarding EE resulting from a day’s total PA, EAMs may in principal provide good estimates, without the need of questionnaires for this specific aspect. However, it is difficult to give concrete public health recommendations based solely on EE. Epidemiological studies will thus have to consider the combination of both methods in order to provide new insights into the benefits of physical activity and to answer open questions in the future [3], [64], [76]. This concept has been proposed to the German National Cohort and is also used in a number of other large-scale projects [7], [8], [9], [10], [11], [13], resulting in a combined assessment of physical activity by accelerometry devices and questionnaires.

In the past, numerous studies provided evidence that PA is associated with the reduced risk of contracting many diseases. However, advising the public to eat healthy and be physically active is not helpful. Concrete evidence-based recommendations are needed. To date, existing PA guidelines are still based on weak evidence. Future innovative epidemiological research on disease-related outcomes must investigate the effects of more specific types of activity, including aspects such as resistance versus aerobic exercise or regular versus irregular moderate activity patterns and even different sleeping patterns and qualities. As evidence arises that stress may have negative impacts on health and aging, and a stress-buffering effect of PA has been observed, it might also be of interest whether PA is experienced as a recreational activity or a stressful duty [77], [78], [79]. Hence, it is essential to gather details on activity behaviour, an area where questionnaires do play an important role. EAMs are not able to provide information about the setting, exact type and mode of activity, and cannot recognize and reliably assess resistance exercise or activities with or without carrying weights.

Individual physical activity patterns need to be investigated. Sports science recommends regular PA over a longer time period to increase (or maintain) physical fitness. Analogous to dietary research, which observed that weight cycling (i.e. losing weight by fasting or special diets and thereafter regaining body weight) might be worse than staying at a certain overweight, fluctuating PA patterns need to be investigated. A typical example is physical inactivity for half of the year followed by strenuous workouts in the summer months. However, up to date epidemiological research did not consider temporal patterns of PA or vigorous activities, but typically assessed average or composite hours of PA per week [6], [18], [80]. Thus, it is an important challenge for future PAQs to assess those temporal PA patterns. EAM cannot fulfill this task, because devices are worn only for a limited time period, typically 7–10 days. Hence, they are not able to consider how active a person has been the rest of the year or how active s/he is in general. Assessing a whole year with EAMs is not feasible. Furthermore, EAMs do not deliver unbiased data either, due to their influence on the behaviour of participants. At least during the first few days, people are aware that they are being monitored and might alter their behaviour by being more active as usual.

An obvious target of questionnaires is the retrospective assessment of PA in the past, including lifetime PA or PA in adolescence. Such a time window is not assessable by electronic devices. Retrospective assessment of PA is especially of interest in case-control studies, where actual PA is influenced by the disease status. Depending on disease aetiology, assessment of lifetime or adolescence PA might also be of interest as an influencing factor in other study designs.

Physical inactivity is increasingly considered to be a burden as it is responsible for about 5 million deaths worldwide each year [81]. Although, EAMs are good at assessing inactivity, activity data can get mixed up with non-wearing time. Questionnaires, however, may also be not accurate, due to memory and recall bias, or incomplete assessment of all activity domains. In addition, sedentary behaviour which is defined as ‘any waking behaviour characterized by an EE≤1.5 METs while in a sitting or reclining position’ has been identified as a potential risk factor independent of physical inactivity [82]. Some EAMs are good at the quantification of sedentary time. Questionnaires might be less accurate in assessing the total time spent sitting, but may provide more details such as the distinction between sitting without physical activity (e.g. watching TV) or sitting with upper body activity (e.g. as clerk at the cash desk).

Another aspect, not yet measureable with motion sensors, is the role of the environment on individual behaviour and the settings participants engage in activity. Questionnaires might be less accurate in assessing the total time spent sitting, but may provide more details such as the distinction between sitting without physical activity (e.g. watching TV) or sitting with upper body activity (e.g. as clerk at the cash desk).

Practical issues such as standards and definitions play an important role in guaranteeing comparability between instruments. With improving technology, EAMs will change rapidly. This development must be considered in planning longitudinal studies. It is essential that data assessed with current devices remain comparable to data assessed with emerging technology. Regarding questionnaires, definitions must be clear. Often there is a different understanding of certain terms, as not all types of activity are perceived equally as a sport (e.g. Tai Chi or Yoga). All this can be achieved by a) improved collaborations between researchers as well as between researchers and compan-
ies producing EAMs and b) establishing a common framework with consistent interpretation when measuring physical activity [83]. However, one problem remains. As there is no feasible gold standard available for determining PA, EE, or PA patterns, the investigation of the reliability and validity of EAMs or questionnaires is still limited. EAMs might improve in practicability and will assess more information, which will open their use to large-scale studies [5]. A further issue is the currently unpredictable development of future prices for EAMs, which will have prime impact on their usage in large-scale studies. Furthermore, utilisability will only be achieved when developers, researchers, and users better cooperate in facilitating the development of future devices and enable their real-time use in field studies. Questionnaires, on the other hand, will be enhanced by electronic devices (such as computers, tablets, and smartphones) and may be administered via the internet. This will enable reaching more participants and gaining more detailed information. Future methods should help disentangle the complex relationships between activity behaviour and its consequences for the body. This can be achieved by identifying the activity patterns that underlie measured levels of EE, cardiorespiratory fitness, muscular strength, or certain biomarkers. All this can best be achieved with a combination of EAMs, questionnaires, and analytical techniques. The advantages and especially the limitations of existing methods must be carefully considered in order to find out which measure best fits a given research question.

**Conclusion**

Despite the rapid development of technology, there are several aspects of PA which cannot be measured by EAMs, but can be covered by questionnaires. Both methods have strengths and limitations. They do not compete with each other, but rather are complementary tools that assess distinct aspects of PA. Especially large cohort studies should rely on both methods in comprehensive investigations of this complex lifestyle factor, which may finally enable investigators to give concrete public health recommendations.

**Notes**

**Abbreviations**

PA: physical activity; HRM: heart rate monitoring; EAMs: electronic activity monitors; EE: energy expenditure; AEE: activity energy expenditure; PAQs: physical activity questionnaires

**Competing interests**

The authors declare that they have no competing interests.

**Authors’ contributions**

BB performed the literature research and drafted the manuscript. KS and MS have been involved drafting the manuscript and revising it critically for important intellectual content. All authors read and approved the final manuscript.

**References**


4. Lagieros YT. Physical activity—the more we measure, the more we know how to measure. Eur J Epidemiol. 2009;24(3):119-22. DOI: 10.1007/s10654-009-9316-0


Corresponding author:
Prof. Dr. Karen Steindorf
Unit of Physical Activity and Cancer, Division of Preventive Oncology (G110), German Cancer Research Center (DKFZ) and National Center for Tumor Diseases (NCT), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany, Phone: +49 6221 42-2351, Fax: +49 6221 42-2229
k.steindorf@dkfz.de

Please cite as

This article is freely available from

Published: 2014-07-18

Copyright
©2014 Brühmann et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by-nc-nd/3.0/deed.en). You are free: to Share — to copy, distribute and transmit the work, provided the original author and source are credited.