Using accelerometers to measure physical activity in large-scale epidemiological studies: issues and challenges

I-Min Lee, Eric J Shiroma

ABSTRACT
Background The current guidelines for aerobic activity require that adults carry out ≥150 min/week of moderate-intensity physical activity, with a large body of epidemiological evidence showing this level of activity to decrease the incidence of many chronic diseases. Less is known about whether light-intensity activities also have such benefits, and whether sedentary behaviour is an independent predictor of increased risks of these chronic diseases, as imprecise assessments of these behaviours and cross-sectional study designs have limited knowledge to date.

Methods Recent technological advances in assessment methods have made the use of movement sensors, such as the accelerometer, feasible for use in longitudinal, large-scale epidemiological studies. Several such studies are collecting sensor-assessed, objective measures of physical activity with the aim of relating these to the development of clinical endpoints. This is a relatively new area of research; thus, in this article, we use the Women’s Health Study (WHS) as a case study to illustrate the challenges related to data collection, data processing and analyses of the vast amount of data collected.

Results The WHS plans to collect 7 days of accelerometer-assessed physical activity and sedentary behaviour in ~18 000 women aged ≥62 years. Several logistical challenges exist in collecting data; nonetheless, as of 31 August 2013, 11 590 women have already provided some data. In addition, the WHS experience on data reduction and data analyses can help inform other similar large-scale epidemiological studies.

Conclusions Important data on the health effects of light-intensity activity and sedentary behaviour will emerge from large-scale epidemiological studies collecting objective assessments of these behaviours.

INTRODUCTION
Physical activity is medicine, as this special issue of the BJSM declares. A large body of epidemiological literature, accumulated over more than 60 years, shows that individuals who are physically active have a better cardiometabolic risk profile, lower rates of major non-communicable diseases, better physical and mental function, and live longer, compared with those who are inactive.1–3 It is, therefore, unfortunate that approximately one-third of the world’s population does not get sufficient physical activity to meet the current recommendations.4 Indeed, this high prevalence of insufficient activity has been estimated to cause as many deaths worldwide each year as smoking5; although smoking increases the risk of dying for the individual more than does inactivity, the prevalence of smoking worldwide (about one-quarter) is less than that of inactivity,6 resulting in comparable harm attributable to each behaviour at the population level.

Although we have learnt much about the health benefits of physical activity, there remain major gaps in knowledge. Some of these gaps have resulted from imprecise assessments of physical activity in research studies. In recent years, technological advances have allowed for better assessments via movement sensors such as the accelerometer, which has become increasingly feasible for use in research from the perspective of the participant (minimal inconvenience), as well as the investigator (reasonable cost).7 In this article, we address two specific gaps in knowledge related to physical activity and health and discuss how accelerometers can add information, as well as challenges related to employing these devices in large-scale epidemiological studies, using the Women’s Health Study (WHS) as an example.
its recommendations on the findings from an extensive review of existing literature by an expert panel.\textsuperscript{1} This panel emphasised data from studies that examined the relation between physical activity and the risks of developing hard clinical endpoints, such as all-cause mortality, cardiovascular disease, type 2 diabetes and cancer, reviewing the data from individual studies in detail. The expert panel did consider data from studies of physical activity in relation to cardiometabolic risk factors (eg, blood pressure, lipid parameters, measures of glucose and insulin processing); however, these data were considered supporting information.

To date, information on how physical activity influences the risks of developing hard clinical endpoints among generally healthy adults have, in large part, come from observational epidemiological studies which used self-reports of physical activity by study participants.\textsuperscript{1} This has occurred because such studies typically require thousands, and even hundreds of thousands of participants to possess sufficient statistical power for investigating the associations of physical activity with the incidence of hard clinical endpoints (eg, major non-communicable diseases), and self-report is a cost-effective method of assessment. Self-reports of physical activity are more reliable and valid for activities of moderate-to-vigorous intensity than light-intensity activities, which tend to be poorly reported.\textsuperscript{9,10} Therefore, observational epidemiological studies that rely on self-reports have tended to limit their assessments to activities of moderate and vigorous intensity only. As a consequence, current recommendations do not prescribe light-intensity physical activity because few data are available, and not because existing data indicate no benefit. Indeed, the US guidelines acknowledge that “we don’t understand enough about whether doing more baseline activity results in health benefits.”\textsuperscript{9,10} International physical activity guidelines have also employed a similar reasoning for their recommendations.\textsuperscript{3}

**SEDENTARY BEHAVIOUR: AN INDEPENDENT RISK FACTOR?**

A growing body of epidemiological literature over the past 5–10 years has described the associations between sedentary behaviour and increased risks of all-cause mortality and cardiometabolic diseases,\textsuperscript{11} with plausible biological mechanisms (eg, lipoprotein lipase regulation) proposed from animal studies.\textsuperscript{12} In several studies, investigators reported on the persistence of these associations even among persons active enough to meet physical activity recommendations.\textsuperscript{13–20} For example, in a recent study of almost one-quarter million US older adults (50–71 years) free from cardiovascular disease and cancer, more time spent on TV viewing was associated with higher all-cause mortality rates, regardless of physical activity level. Among persons reporting >7 h a week of moderate-to-vigorous intensity activity, the risk of dying during follow-up was 50% higher among those declaring seven or more hours of TV viewing a day, compared with <1 h a day.\textsuperscript{18}

One limitation of these studies on sedentary behaviour has been, as with studies of physical activity, reliance on self-reports since this method of assessment is most feasible for studies with large numbers of participants. In a review of 48 longitudinal studies investigating sedentary behaviour in relation to subsequent health outcomes, 45 used participants’ self-reports, primarily of TV and screen viewing time.\textsuperscript{21} It is unclear how well self-reported TV/screen time reflects total sedentary time. A study comparing self-reported TV viewing time with accelerometer-assessed sedentary time among a representative sample of US adults in the National Health and Nutrition Examination Survey yielded a correlation of only 0.22, indicating a low agreement.\textsuperscript{22} Imprecision in self-reports also may partly explain the findings such as these: in a nationally representative sample of adults from the Health Survey for England, cross-sectional analyses showed consistent deleterious associations of self-reported sitting time with a panel of cardiometabolic risk factors; however, when accelerometer measures of sedentary behaviour were analysed, little relation was observed.\textsuperscript{23} Thus, one could argue that it may not be the sedentary behaviour per se but, rather, associated unhealthy behaviours—such as snacking while watching TV—that accounts for the observed adverse associations.

There do exist data from several studies that used accelerometers to measure the sedentary behaviour, showing unfavourable associations with risk factors (eg, the Australian AusDiab study,\textsuperscript{24} the Proactiv UK trial,\textsuperscript{25} the US National Health and Nutrition Examination Survey\textsuperscript{26}). However, these studies are predominately cross-sectional in design; thus, one cannot ascertain the direction of association (ie, does being sedentary result in poor health? Or, does being in poor health predispose one towards being sedentary?). In addition, observed improvements in risk factors do not always result in reduced rates of disease, as demonstrated by the Look AHEAD trial, a randomised controlled trial comparing an intensive lifestyle intervention with a traditional diabetes support and education programme among patients with overweight/obese type 2 diabetes.\textsuperscript{27} In recently published results, the intensive intervention clearly improved cardiometabolic risk factors; however, there was no difference in rates of cardiovascular events between groups after a median follow-up of 9.6 years, when the trial was stopped for futility.

Therefore, there is a need for longitudinal studies of sedentary behaviour employing objective assessments in relation to hard clinical endpoints, to complement and clarify the existing data that may be limited by one or more of the following factors: self-reports of sedentary behaviour, cross-sectional design, ascertainment of only risk factors but not disease incidence.

**USING ACCELEROMETERS IN LARGE-SCALE EPIDEMIOLOGICAL STUDIES**

One movement sensor capable of measuring light-intensity physical activity and sedentary behaviour objectively is the accelerometer. As its name suggests, the device measures accelerations. The first commercial use of the accelerometer occurred in the 1920s, where they were used in bridges (to record vibrations) and aircraft (to measure accelerations).\textsuperscript{28} It weighed about a pound and cost $420 in the 1930s (some $6000 today). Technological advances have made the device smaller and lighter (<1 oz), and less expensive (eg, the Actigraph GT3X+ model, used in the case study of the WHS below, costs $200–$250) today. Thus, the accelerometer has now become feasible for use in large-scale studies, such as the examples shown in table 1.

While accelerometers provide objective measurements, it is important to note some limitations. They primarily measure locomotor activity when worn over the hip (typical placement position), and so miss upper body movement. Furthermore, they cannot distinguish whether a person is carrying any weight (eg, walking carrying a heavy bag expends more energy vs walking with no load). Accelerometers do not inform on body posture; thus, they cannot distinguish between sitting and standing still. However, for some populations, these are not major limitations; for example, in the WHS below that includes older women, walking is the most common physical activity. Since participants are older, accelerometer indications of very low levels of activity over long periods during waking hours likely indicate sitting, and not standing with little movement.
TABLE 1  Examples of cohort studies of cardiometabolic diseases and cancer using accelerometers to assess physical activity and sedentary behaviour

<table>
<thead>
<tr>
<th>Study, country</th>
<th>Approximate N, age</th>
<th>Years and method of data collection</th>
<th>Device</th>
</tr>
</thead>
<tbody>
<tr>
<td>European Prospective Investigation of Cancer (EPIC)—Norfolk Study, UK*</td>
<td>3900 men and women, 60–80 years</td>
<td>2008–2011; by mail</td>
<td>Actigraph GT1M</td>
</tr>
<tr>
<td>Reasons for Geographic and Racial Differences in Stroke (REGARDS) Study, USA29</td>
<td>9400 men and women, 56+ years</td>
<td>2009–2012; by mail</td>
<td>Actical</td>
</tr>
<tr>
<td>British Regional Heart Study, UK†</td>
<td>2500 men and women, 70–90 years</td>
<td>2010; by mail</td>
<td>Actigraph GT3X</td>
</tr>
<tr>
<td>Maastricht Study, Netherlands‡</td>
<td>10 000 men and women, 40–75 years</td>
<td>2010 until sample reaches 10 000; in person</td>
<td>ActiPal</td>
</tr>
<tr>
<td>Women’s Health Study (WHS), USA30</td>
<td>18 000 women, 62+ years</td>
<td>2011–2014; by mail</td>
<td>Actigraph GT3X+</td>
</tr>
<tr>
<td>Women’s Health Initiative Study (WHI), USA31</td>
<td>8000 women, 72+ years</td>
<td>2012–2013; in person and by mail</td>
<td>Actigraph GT3X+</td>
</tr>
</tbody>
</table>

†Jeffers, B, personal communication, 2013.
‡Koster, A, personal communication, 2013.

CASE STUDY: ISSUES AND CHALLENGES
Women’s Health Study

The WHS is a completed randomised trial (1992–2004) testing aspirin and vitamin E for preventing cardiovascular disease and cancer among 39 876 healthy women, 245 years, living throughout the USA.32–34 When the trial ended, women were invited to continue in an observational study. Of those alive, 89% or 33 681 women consented, reporting on their health habits and medical history annually on questionnaires. In 2011, data collection began for an ancillary study, in which the main aim is to examine the relations of accelerometer-measured physical activity and sedentary behaviour with health outcomes. Women are asked to wear an accelerometer (ActiGraph GT3X+, ActiGraph Corporation, Pensacola, Florida, USA) on their hip during all waking hours for 7 days. Data collection will be completed at the end of 2014. Below, we discuss our experience and show some preliminary data.

Logistical challenges

When data collection for the WHS accelerometer study began, approximately 30 000 women were alive; thus, the most feasible method for collecting data was via mail. As participants had already been in the main study for an average of almost 17 years, we believed it was possible to get good response rates and compliance with instructions. Typically, in the main study, we obtain ~80% response to the first mailing of study questionnaires, another ~10% to the second, and a further ~5% to the third, for a total response of ~95%. However, because of funding cuts, we chose initially to send only a single invitation to participate in the accelerometer study (and will consider a second request to achieve our goal of ~18 000 women). Invitations are sent on a rolling basis because of the limited number of accelerometers available.

When a participant agrees to participate, we mail her an accelerometer, a log sheet to record dates/times that the device is put on and taken off (further discussed below), and instructions. On receiving this packet, some women change their mind about participation, or request postponement to a more convenient time. Two issues result: first, in such a large study, it is challenging to keep track of the stage at which each individual participant is. In addition, other substudies are occurring in the WHS (eg, repeated yearly assessments in a subgroup; comparing wrist-worn and hip-worn devices in another subgroup), which further complicate the tracking. Second, with dwindling participation at each stage, sample size can decrease substantially.

As of 31 August 2013, 23 266 women have been invited; 19 432 (84%) have responded, of whom 15 314 (79% of respondents) have agreed to participate (16% declined and 5% were not eligible because they cannot walk outside without assistance). Packets have been mailed to 14 073 women, of whom 11 590 have worn and returned their devices, 725 have requested participation at a later date (we assume some of these will eventually decline), and 1087 have changed their mind about participation (the remaining 671 women were not due as of 31 August 2013 to return their device). Of the ~30 000 total women, we expect 1320 (4.4%) to be non-eligible, and project participation by ~18 000 (~63%) of the remaining 28 680. This experience of dwindling participation is not unique; for example, in the mail-based Reasons for Geographic and Racial Differences in Stroke (REGARDS) study (table 1), 48% (10 863/22 195) of invited participants returned a device with some data.29

Participants do lose or fail to return their accelerometers; the rate of loss in the WHS is 2.1%. While this seems negligible, 2.1% of the ~20 000 whom we anticipate sending accelerometers translates to ~420 units, at a substantial cost. In the REGARDS study, investigators reported an 8% loss rate.29

Challenges related to data reduction

The technological capability for data capture using accelerometers has outpaced the current knowledge on how to reduce and process these data. Furthermore, while best practices and standards have been proposed35–37 no consensus has been reached. Thus, in the WHS, we keep multiple data files, processed using different conventions, for each participant: consequently, we anticipate storing a large, ~20 terabytes, amount of data when data collection ends.

To illustrate the challenges in this area, we discuss two important decisions that have to be made. The first relates to determining actual wear time of the device by participants, and not mistaking movements in the mail process as participant physical activity. We ask women to record, daily, dates and times of wear on a log; ideally, these would not be required as they are onerous for participant and investigator, and missing data may be an issue. Automated algorithms are available (eg, the commonly used Troiano et al38 and Choi et al39 algorithms) for wear-time determination; however, these were developed for protocols where participants received the device directly from investigators (thus, ‘mail noise’ is not applicable). We hope to use the WHS experience to develop guidelines for other large-scale studies using a mail protocol. Based on preliminary data, we have found that using logs only to record whether the device was ‘worn’ or ‘not worn’ for the day, but not to indicate times (minimising participant burden and likelihood of missing
Challenges related to data analyses

As with data reduction, the current analytic methods for accelerometer-assessed data appear rudimentary compared with the technology available for data collection. Most analyses to date integrate acceleration data from the vertical axis only into counts per user-specified time period (eg, counts per minute, cpm) and utilise cut-points for classifying time spent in sedentary behaviour or physical activity (higher cpm reflects higher intensity activities), which are then related to health outcomes. However, there is no consensus on what the ‘correct’ cut-points might be, with several proposed in the literature based on calibration studies primarily carried out under laboratory settings and not in free-living conditions. Furthermore, cut-points developed among middle-aged persons may not be applicable for older populations. And, the current prediction techniques that use cut-points to estimate energy expenditure are limited, with the occurrence of underestimation and overestimations.

Newer accelerometer models (including the device used in the WHS) collect data not only from the vertical (up-and-down movements), but also from the anteroposterior (back-to-front) and lateral (side-to-side) axes. There are hardly any data for cut-points developed using triaxial data. A recent work indicates a strong correlation between counts from the vertical axis and counts combined from all three axes using vector magnitude; thus, the vector magnitude data may not provide much additional information beyond vertical axis counts.

Finally, cut-point-based analyses make only very limited use of the wealth of data available. Efforts are ongoing to take a better advantage of the rich data, such as pattern recognition or machine learning data analytic approaches to identify specific activities carried out.

Some preliminary data

Data collection in the WHS is anticipated to occur between 2011 and 2014. While the entire sample of eligible and willing women will be needed to provide sufficient statistical power for examining accelerometer-assessed sedentary behaviour and physical activity in relation to incidence of chronic diseases and all-cause mortality (primary aims of the study), a smaller sample size is adequate to address secondary aims, such as describing patterns of sedentary behaviour and physical activity.

Thus, we decided a priori to create an interim, closed dataset that comprised all women who returned their accelerometer by 31 March 2013 (n=8373; approximately half the total anticipated sample of ~18 000). Women were of mean age 71 years, 20% were obese, and 4% were current smokers. They complied well with instructions: 95% wore their device for ≥10 h each day on ≥4 days (conventionally regarded as ‘valid’ data). Figure 1 shows their physical activity levels compared with a comparably aged US national sample (data are presented using commonly used cut-points for defining different intensities of physical activity). A detailed description of patterns of sedentary behaviour among women will be published elsewhere.

CONCLUSIONS

That we face the challenges discussed is a testament to how far the field of physical activity epidemiology has advanced. In the early days of research in this field, pioneers—led by Professors Jerry Morris and Ralph Paffenbarger, Jr—sought to convince others of the health benefits of physical activity. Today, with this...
clearly established, investigators have moved to answering other questions such as those related to the ‘dose’ needed. In particular, given the low levels of physical activity around the world, there is a great deal of interest on the low end of the physical activity spectrum, that is, light-intensity activities and sedentary behaviour. Imprecise assessments of these behaviours and cross-sectional study designs have limited our knowledge. However, recent technological advances have made the use of movement sensors, such as the accelerometer, feasible for use in longitudinal, large-scale epidemiological studies. Presently, several such studies, including the WHS that we used as a case study, are collecting sensor-assessed, objective measures of physical activity with the aim of relating these to the development of hard clinical endpoints. While challenges exist, especially those related to processing and analysing the vast amount of data collected, over the next 5–10 years much interesting data will emerge.

What are the new findings?

- Recent technological advances have made the use of movement sensors, such as the accelerometer, feasible for use in longitudinal, large-scale epidemiological studies that intend to investigate the associations of objectively measured physical activity or sedentary behaviour with hard clinical outcomes.
- Logistical challenges exist in collecting objectively measured physical activity or sedentary behaviour data from large-scale epidemiological studies, which result in a reduced number of participants available for investigation.
- While large amounts of data can be collected using accelerometers, procedures to reduce and process these data are not well developed; thus, best practices and standards for accelerometer data reduction and processing are needed.

How might it impact on clinical practice in the near future?

The current physical activity guidelines recommend activities of at least moderate intensity, but there are no guidelines targeting activities of lower intensity or sedentary behaviour. Imprecise assessments of these behaviours and cross-sectional study designs have limited knowledge; however, recent technological advances have allowed for more precise assessments that are feasible in longitudinal, large-scale epidemiological studies. Several ongoing studies will provide information that can inform the development of guidelines related to light-intensity activity and sedentary behaviour.

Acknowledgements The authors are grateful to Dr Patty Freedson and Dr Steven Hooker, as well as to the staff of the Women’s Health Study, particularly Ara Sarkissian, Bonnie Church, Colby Smith and Jane Jones, for their assistance with the study.

Contributors I-ML designed and conducted the study, conducted the data analyses and drafted the article. EJS oversaw data collection, assisted with data analyses and provided critical review to the article. I-ML and EJS had full access to the all of the data and take full responsibility for the integrity of the data and the accuracy of the data analysis.

Funding This study is supported by grant CA154647 from the US National Institutes of Health.

Competing interests None.

Ethics approval Brigham and Women’s Hospital’s institutional review board committee.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES
