Cochrane, MA, Watson, PM, Timpson, H, Haycox, A, Collins, B, Jones, L, Martin, A and Graves, LEF

Systematic review of the methods used in economic evaluations of targeted physical activity and sedentary behaviour interventions

Systematic review of the methods used in economic evaluations of targeted physical activity and sedentary behaviour interventions

Cochrane, M.[a], Watson, P,M.[a], Timpson, H.[b], Haycox, A.[c], Collins, B.[d], Jones, L.[c], Martin, A.[d,e], Graves, L.E.F.[a]

[a]Physical Activity Exchange, Research Institute for Sport and Exercise Sciences, Liverpool John Moores University, 5 Primrose Hill, Liverpool, L3 2EX, UK

[b]Public Health Institute, Faculty of Health, Education and Community, Liverpool John Moores University, 3rd Floor, Exchange Station, Tithebarn Street, Liverpool, L2 2QP, UK

[c]Management School, University of Liverpool, Chatham Street, Liverpool, L69 7ZH, UK

[d]Department of Public Health and Policy, University of Liverpool, Waterhouse Building, Block B, 2nd Floor Liverpool, L69 3BX, UK

[e]HCD Economics, The Innovation Centre, Keckwick Lane, Daresbury, Warrington, WA4 4FS, UK

Abstract

The burden of noncommunicable diseases (NCD) on health systems worldwide is substantial. Physical inactivity and sedentary behaviour are major risk factors for NCD. Previous attempts to understand the value for money of preventative interventions targeting physically inactive individuals have proved to be challenging due to key methodological challenges associated with the conduct of economic evaluations in public health. A systematic review was carried out across six databases (Medline, SPORTSDiscus, EconLit, PsychINFO, NHS EED, HTA) along with supplementary searches. The review examines how economic evaluations published between 2009-March 2017 have addressed methodological challenges with the aim of bringing to light examples of good practice for future studies. Fifteen economic evaluations from four high-income countries were retrieved; there is a dearth of studies targeting sedentary behaviour as an independent risk factor from physical activity. Comparability of studies from the healthcare and societal perspectives were limited due to analysts' choice in cost categories, valuation technique and time horizon differing substantially. The scarcity of and inconsistencies across economic evaluations for these two behaviours have exposed a mismatch between calls for more preventative action to tackle NCD and the lack of information available on how resources may be optimally allocated in practice. Consequently, this paper offers a table of recommendations on how future studies can be improved.

Keywords

Systematic Review; Economic Evaluation; Physical Activity; Sedentary Behaviour; Equity; Public Health; Cost effectiveness analysis; Cost utility analysis

Introduction

Background

The burden of noncommunicable disease (NCD) on health systems is substantial. Worldwide NCD is the main cause of death and disability (WHO, 2018a). Physical inactivity is a major risk factor for NCD and the fourth leading cause of death globally. There is therefore an urgent need to invest in preventative interventions, such as those targeting individuals who do not meet the international guidelines of 150 minutes of moderate physical activity per week (Kohl et al., 2012). Furthermore, sedentary behaviour, defined as any waking behaviour where an individual is in a sitting, reclining or lying posture, has been identified as a risk factor for NCD and all-cause mortality independent of achieving the recommended physical activity guidelines. The level of physical activity found to attenuate the risks associated with sedentary behaviour is 60 minutes of moderate physical activity per day, which equates to 420 minutes per week (Ekelund et al., 2016). As over a third (35%) of females and a quarter (26%) of males in high-income countries do not presently meet the recommended weekly guidelines, a daily target of 60 minutes is unlikely to be attained (WHO, 2018b). Inaction to invest in preventative interventions tackling detrimental levels of physical inactivity and sedentary behaviour is expected to lead to greater levels of NCD and inequity, productivity losses and a continued overwhelming demand for costly curative health services (OECD, 2015).

As public resources are scarce, economic evaluations are important to prevent both national and local policymakers from disinvesting in highly cost-effective physical activity and sedentary behaviour interventions. Economic evaluations are also needed as not all public health interventions represent good value for money (Owen et al., 2017). Compared to population-level interventions, physical activity and sedentary behaviour interventions targeting individuals who are not meeting the recommended international physical activity guidelines are more likely to be: (a) funded by local-level commissioners; (b) evaluated by researchers. This is likely to be due to the challenge of measuring outcomes in the general population. For this reason, this review focuses on economic evaluations of targeted interventions such as exercise referral schemes, brief advice in primary care and exercise sessions.

Despite recommendations for economic evaluations to become routine within public health interventions (Kelly et al., 2005) cost-effectiveness information on physical activity and sedentary behaviour interventions remains scarce (Abu-Omar et al., 2017). One reason for this lack of analysis may be due to the lack of guidance and multidisciplinary efforts to inform analysts on how to conduct economic evaluations in the field of public health (Davis et al., 2014). Economic evaluations of public health interventions are subject to four key methodological challenges identified and described in former reviews (Alayli-Goebbels et al., 2014; Hill et al., 2017; Weatherly et al., 2009) as: attribution of effects; measuring and valuing outcomes; identifying intersectoral costs and consequences; and incorporating equity. These four challenges are explain in the subsequent sections.

Challenge 1: Attribution of effects

Randomised controlled trials (RCTs) are the gold standard for evaluating the effectiveness of an intervention. RCTs alone are however insufficient to inform long-term investment decisions in health systems aiming to be sustainable. This is because conducting experimental studies such as RCTs over many years or decades is likely to be resource intensive from both the research funder and participant's perspective. Attrition from the trial and insufficient funding is inevitable. Yet, the greatest health outcomes and cost savings attributable to physical activity and sedentary behaviour interventions do not typically manifest until decades after an intervention has taken place. Due to this long pay-back time (Wanless, 2004), it is recommended economic evaluations link up trial-derived intermediate or surrogate outcomes with additional sources of evidence (e.g. observational studies) (Ramsey et al., 2015).

Challenge 2: Measuring and valuing outcomes

Previous physical activity studies have used different outcomes, or have classified the same type of outcomes in different ways, which makes it challenging to meaningfully use cost-effectiveness results and compare interventions (Abu-Omar et al., 2017). This is likely to be because physical activity and sedentary behaviour interventions are associated with a broad range of outcomes, many of which are not captured in evaluations that conduct just one type of valuation analysis. Furthermore, many broader important and relevant outcomes such as improved wellbeing or someone's ability to return to work are difficult to assign a monetary value, as they do not have a market price (Weatherly et al., 2014).

Challenge 3: Identifying intersectoral costs and consequences

Many physical activity and sedentary behaviour interventions take place outside of the healthcare setting, necessitating a time and equipment commitment from intervention participants and providers (which has an opportunity cost). Moreover, physical activity and sedentary behaviour interventions are complex, impacting on various sectors simultaneously (Dahlgren & Whitehead, 1991). Therefore, it is important to consider the impact of these interventions on other stakeholders including public sector agencies beyond the health sector, private individuals and the voluntary sector (Weatherly et al., 2014; Weatherly et al., 2009). Yet, as there is no universal definition for each perspective type, the costs and consequences deemed relevant for inclusion in the analysis is primarily analyst-dependent (Husereau et al., 2013).

Challenge 4: Incorporating equity

A key objective in public health is to reduce inequity, meaning inequalities that are avoidable, but have not yet been avoided and are therefore unfair (Marmot & Allen, 2014). By contrast, a key objective in economic evaluation is to maximise efficiency across the whole population (Weatherly et al., 2014). If authors fail to acknowledge equity by not adapting their existing economic analysis approach, it is not transparent which socio-economic group have gained or lost out due to a resource allocation decision. Until the recent publication by Cookson et al. (Cookson et al., 2017) recommendations on how to incorporate equity have been limited within international and national guidelines for economic evaluation (Husereau et al., 2013; NICE, 2014; Ramsey et al., 2015; Sanders et al., 2016). Approaches for incorporating equity into the analysis described by Cookson et al. (2017) include: equity impact analysis, equity constraint analysis and equity weighting analysis.

Aim

In an attempt to learn how the four challenges outlined above have been addressed in practice, this systematic review aims to provide an overview of the methods used in economic evaluations of physical activity and sedentary behaviour interventions since 2009. Alayli-Goebbels et al. (2014) and Weatherly et al. (2009) reviewed the methods reported in economic evaluations of a range of public health areas including 17 and 26 physical activity economic evaluations published up to 2005 and 2009, respectively, but the reviews found little insight from the empirical evidence. Economic evaluation is a rapidly developing field especially with the growth of decision-analytic modelling and the economic evaluation reporting standards (Drummond et al., 2015; Ramsey et al., 2015). Accordingly there is a strong rationale to provide an update on methods carried out since 2009.

Methods

Information sources and search strategy

A comprehensive search took place across six electronic databases that host reports from the medical and economic field (Medline via Ovid; SPORTSDiscus, EconLit and PsycINFO via EBSCO*Host*; NHS EED and HTA via the Cochrane Library). The database NHS EED stores records up to April 2015, thus searches in this database went up to 2015 only. Additional, supplementary searching was performed: key websites were searched for studies that included specific free text terms: 'physical activity', 'sedentary behaviour', 'economic' and 'cost'; reference lists of two relevant systematic reviews (Gc et al., 2016; Wu et al., 2011) were hand searched; and protocols that met the majority of the eligibility criteria were used to search for completed studies via online searching and contacting the authors. An example of the full electronic search strategy for Medline is provided in Appendix A [INSERT LINK TO ONLINE FILE A, B, C, D & E]. This search was replicated for all databases, with amendments made as appropriate to align terms with individual database index terms.

Study selection

The protocol for this review can be retrieved from the PROSPERO database for registered systematic reviews (registration number CRD42017074382). Full economic evaluations of interventions targeting individuals aged 16 years or over, who are defined as being physically inactive or sedentary, were eligible for inclusion in the review. Population level interventions were excluded as well as protocols. Eligible studies needed to capture physical activity or sedentary behaviour at two or more time points to observe if a change in behaviour has occurred. Comparators could be any alternative intervention including no intervention. Interventions and comparators targeting multiple behaviours such as physical activity and diet were excluded unless the multiple behaviours were physical activity and sedentary behaviour. Both trial and model based economic evaluations were eligible. Letters to editors and conference briefings were excluded. Both published and unpublished 'grey' literature were included. Abstracts where the full text could not be retrieved were excluded. Only English language studies were included due to the restricted language skills of the reviewers available. Eligibility criteria was applied during both screening phases. The present systematic review identifies and discusses studies published from January 2009 to March 2017. In addition, a rapid systematic scoping search was performed in Medline to understand whether new studies had been published in this area from March 2017 to January 2019. Details on methods of the scoping search are not discussed below, rather they are presented in Appendix B [INSERT LINK TO ONLINE FILE A, B, C, D & E].

Screening

During the title and abstract screening phase two reviewers (first author, seventh author) screened 10% (n=612/ 6,123) of the studies and there was a disagreement rate of 2.94% (n=18). Reviewers discussed the disagreements and resolved them without the need to seek the expertise of a third reviewer. Reviewer one (first author) went on to screen the rest of the studies, informed by the disagreement discussions. Similarly, during the full text screening phase reviewer two (seventh author) screened 10% (n=15/ 153) of the studies. There was disagreement for 33.33% (n=5) of the studies. The reviewers discussed the disagreements and again a consensus was met without the need for a third reviewer. Figure 1 shows an overview of the study selection process.

Data extraction

A data extraction form was developed based on the items featured on the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) checklist (Husereau et al., 2013). The form was piloted independently by two reviewers (first author, seventh author) on two (10%) randomly selected studies. Following discussions the form was shortened, and items relevant to the four methodological challenges, and key study characteristics were retained. Following the piloting stage, the first reviewer extracted data for the remaining studies. A template of the final data extraction form is provided in Appendix C [INSERT LINK TO ONLINE FILE A, B, C, D & E]. It was not necessary to request additional information from the study authors.

Quality assessment

Drummond's 10-item checklist was selected as it is one of the most widely used quality assessment tools (Drummond et al., 2015). A component approach was used when applying the checklist in Appendix D [INSERT LINK TO ONLINE FILE A, B, C, D & E]. This approach is advocated in the PRISMA statement and entails assessing each item individually rather than generating a summary score (Liberati et al., 2009). Two reviewers (first author, seventh author) independently conducted the quality assessment for 10% (n=2/ 15) of the included studies. Disagreement was limited to item 6 (Item 6: Were costs and consequences valued credibly?) on the checklist, examples in Drummond et al. (2015) were consulted to overcome these disagreements. Practical application of item 10 (Item 10: Did the presentation and discussion of study results include all issues of concern to the users?) was challenging due to the limited guidance, thus findings from this question are less informative. Alayli-Goebbels et al. (2014) also experienced this barrier in an earlier version of the checklist.

Method of analysis

The published narrative synthesis framework by Popay et al. (2006) guided the analysis to ensure a transparent and systematic approach was performed. The narrative synthesis in this review goes beyond describing how authors have addressed each of the four challenges by attempting to explain why specific approaches have been chosen. The analysis was an iterative process. A priori analysis involved tabulating the data and producing bar charts on key study characteristics: study design, time horizon, valuation technique, study perspective and explicit/ implicit equity analysis. The same study characteristics were focused on in the two former methodological reviews (Alayli-Goebbels et al., 2014; Weatherly et al., 2009). The wider literature also indicated that the following contextual factors were important to review when understanding an analyst's approach: intervention setting, country and year of publication. Additional ad hoc analyses were performed where trends became apparent. Lastly, the strength of the narrative synthesis and the conclusions derived from it were considered by reflecting on the quantity of studies and results of the quality assessment.

Results

A total of 15 economic evaluations (17 publications) were included in the review (Figure 1). Searching across Medline, SPORTSDiscus, EconLit, PsychINFO, NHS EED and HTA databases retrieved 7,063 records. Supplementary searching retrieved six additional records including: two records from hand searching on key websites, two from the reference list of a systematic review (Gc et al., 2016), and a further two from searching for the completed studies of two protocols (de Vries et al., 2013; Kolt et al., 2009) in Appendix E [INSERT LINK TO ONLINE FILE A, B, C, D & E]. After removing duplicates 6,129 records remained of which a further 5,907 records were removed as title and abstracts did not meet the eligibility criteria. During the full text screening, 159 citations were examined in further detail, of which 142 studies were excluded. Reasons are outlined in Figure 1.

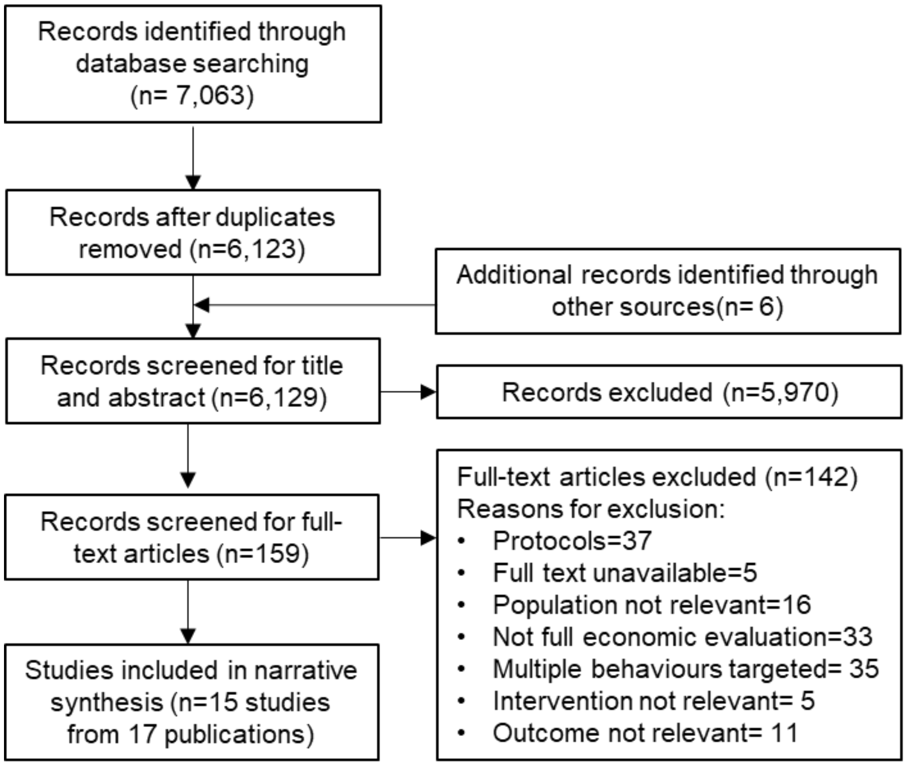

Figure 1. PRISMA flow diagram representing study selection process

Study characteristics

Of the 15 studies, ten were single trial-based economic evaluations and five were model-based; no studies were single trials that had extrapolated or modelled their results. Table 1 provides an overview of study characteristics for the trial- and model-based studies respectively. Studies are arranged by country followed by year of publication. Interventions were set in primary care, community and the home, and setting did not appear to be related to intervention type or country. As shown in Table 1, no studies targeted sedentary behaviour as an independent risk factor from physical activity. The range of interventions was limited to the following types: physical activity programme/ on prescription in primary care (n=9); brief advice in primary care (n=2); home-based informational advice (n=1); physical therapy in a physical therapy setting (n=1); and fall prevention programme in both primary care and the home (n=1). The remaining study compared strategies for recruiting to physical activity interventions in primary care. The overall range of adult-based interventions matches the narrow range identified in a recent review of reviews focussing on the economic results of physical activity interventions (Abu-Omar et al., 2017). Studies came from four high-income countries. More than half (n=8) of the 15 studies came from the UK, with the remaining coming from New Zealand (n=3), the USA (n=2), and the Netherlands (n=2) (Table 1).

Quality assessment

Overall, studies performed well against Drummond's 10-item quality assessment checklist (Drummond et al., 2015) in Appendix D [INSERT LINK TO ONLINE FILE A, B, C, D & E]. Nevertheless, six studies scored 'No' on at least one item: two studies did not state their perspective (item 1); three studies did not include all costs and consequences relevant to their stated perspective (item 4); one study did not discount its costs and consequences (item 7); and one study did not report their price source (item 6). Interpretation on whether item 4 was met by any of the ten trial-based economic evaluations who captured costs and outcomes at two years or less, is up for debate. It could be argued that not all important and relevant costs and consequences can be identified for studies, which do not take a systems approach (e.g. if they do not consider the impact on the wider system in which an intervention is being implemented nor capture the long-term impact) (Rutter et al., 2017; Squires et al., 2016). In order to align with other reviews which have used Drummond's checklist, the quality assessment results for item 4 were based on the checklist's accompanying guidance (Drummond et al., 2015). Costs and consequences identified, measured and valued are discussed in greater depth in the subsequent sections.

Table 1: Overview of economic evaluations

Trial-based economic evaluations								
Study & Year of publication	Stated perspective	Country	Population targeted	Sample size	Intervention	Comparator	Setting	Valuation technique
Iliffe et al. 2014	Health sector	UK	Inactive ≥65 years old who had fallen less than times in the previous 12 months	100	Falls Management Exercise Programme (Weekly group exercise class & 2 home-based exercise sessions)	Usual care (no intervention); Otago Exercise Programme	Primary care & community (as Home-based)	CEA
Edwards et al. 2013; Murphy et al. 2012	Multi-agency public sector	UK	Sedentary, and over 16 years, with risk factors for coronary heart disease, or mild to moderate anxiety, depression or stress.	798	ERS (primary care)	Information leaflet only	Primary care	CUA
Boehler et al. 2011	Health sector	UK	Inactive adults, 16 to 74 years old	46	Opportunistic recruitment strategy for physical activity interventions	Disease register strategy; Hypothetical no intervention strategy	Primary care	CEA

Shaw et al. 2011	Not reported	UK	Inactive, adults (age not defined)	79	Individualised walking programme: a pedometer and a 30-min consultation	Individualised walking programme: a pedometer, but and 5 min brief advice	Primary care	CEA
Larsen et al. 2015	Payer	USA	Inactive Latina women, 18-65 years old	266	Home print-based mail-delivered MVPA intervention linguistically and culturally adapted for Latinas	Wellness contact (information on health topics excluding MVPA)	Home-based	CEA
Young et al. 2012	Societal	USA	Women, following coronary artery bypass surgery	40	Symptom management intervention delivered by telehealth device to improve the physical activity level	Usual care, 2 week follow up call by the primary providers and cardiac specialists	Community	CEA
de Vries et al. 2016	Societal	Netherlands	Sedentary adults (or at risk of losing active lifestyle in near future) with mobility problems, ≥70 years old	130	Patient-centred physical therapy	Usual care for physical therapy, less patient-centred	Physical therapy setting	CUA
Maddison et al. 2015	Not reported	New Zealand	≥18 years old with diagnosis of IHD	171	Exercise prescription and behavioural support	Usual care (participation in usual Cardiac Rehabilitation	Home-based	CEA; CUA

			within previous 3-24 months.		by mobile phone text messages and internet	e.g. education session and psychological support)		
Leung et al. 2012	Public health system and participant	New Zealand	Inactive adults, ≥65 years old	330	Pedometer-based prescription, focus was on step-related goals	Green prescription, focus was on physical activity time-related goals	Community	CEA; CUA
Elley et al. 2011	Societal	New Zealand	Inactive, 40- 74 years old	974	Enhanced green prescription, 10 min of brief advice and a written exercise prescription with telephone support at 9 months and 30min face-to-face support at 6 months	Usual care from GP (not standard green prescription, usual care from GP not defined)	Primary care	CEA

Model-based economic evaluations								
Study & Year of publication	Stated perspective	Country	Population targeted	Model type & size of simulation cohort	Intervention	Comparator	Setting	Valuation technique
Campbell et al. 2015	Health Sector	UK	Sedentary adults, ≥50 years old	Markov model (100,000 simulation cohort)	ERS (primary care)	Usual care (refers to Pavey et al. 2011's definition)	Primary care	CUA
Anokye et al. 2012; Anokye et al. 2014	Health sector; Health sector and participant for CCA	UK	Inactive, ≥33 years old	Markov model (100,000 simulation cohort)	Brief Advice (primary care)	Usual care (no intervention)	Primary care	CUA (and CCA)
Anokye et al., 2011	Health sector	UK	Sedentary adults, 40-60 years old	Decision tree model (1,000 simulation cohort)	ERS (primary care)	Usual care (refers to Pavey et al. 2011's definition)	Primary care	CUA

Pavey et al. 2011	Health sector CUA; Partial-societal for CCA	UK	Sedentary adults, 40-60 years old	Decision tree model (1,000 simulation cohort)	ERS (leisure centre)	Usual care (no active ingredient- PA advice or leaflets)	Leisure-centre	CUA (and CCA)
Over et al. 2012	Health sector	Netherlands	Inactive, 20-65 year olds	Markov model (100,000 simulation cohort)	GP pedometer prescription, counselling combined with pedometer use	Usual care (no intervention)	Primary care	CUA

ERS: Exercise Referral Scheme; GP: General Practitioner; MVPA: Moderate-to-vigorous physical activity; CEA: cost-effectiveness analysis; CUA: Cost-utility analysis

Challenge 1: Attribution of effects

Two thirds (n=10) of the studies in this review, all trial-based, did not compare the costs and consequences of the comparator groups beyond the trial follow up period (Table 3). More specifically, one study compared costs and consequences over a two-year period (Elley et al., 2011), the remaining nine had a time horizon of 12-months or less. For six of these studies, authors referred to their short time horizon as a limitation of their study (Boehler et al., 2011; de Vries et al., 2016; Edwards et al., 2013; Larsen et al., 2015; Leung et al., 2012; Shaw et al., 2011). For instance, it precluded the incorporation of any potential long-term healthcare savings (Larsen et al., 2015). Just one study suggested future modelling exercises could be used to address this challenge (Edwards et al., 2013). Yet, for Shaw et al. (Shaw et al., 2011) a short-time horizon was justified as they reported there was insufficient data to extrapolate their results over the participants' lifetime.

By contrast, all five model-based studies extrapolated a pooled trial-derived effectiveness estimate over the rest of the participants' lifetime; bridging the gap between the short- and long-term evidence (Table 2). Nevertheless, the assumptions underpinning the model-based studies varied considerably. Two studies (Anokye et al., 2011; Pavey et al., 2011) made large assumptions unsupported by evidence about the duration of the effect, assuming that any short-term change in physical activity observed in the trials 6-12 months after the intervention, would be long-lasting. Over et al. (2012) employed a different approach by extrapolating an effect estimate, observed at 18 weeks, over a 40-year time horizon (the life expectancy of the participants). The authors assumed that only 25% of the effect recorded at 18 weeks would remain over the 40-year time horizon; they too reported that their assumptions were unsupported by evidence. These findings demonstrate how studies will vary according to the assumptions made. It is therefore important that end-users of cost-effectiveness results check they agree with the assumptions that underpin the economic evaluation.

Assumptions underlying the two other model-based studies (Anokye et al., 2012; Campbell et al., 2015) were supported by three robust cohort studies. Campbell et al. (2015) replicated

Anokye et al.’s (2012) approach. More specifically, they linked the short-term change in physical activity level observed in trial data, with Hu et al.’s (2007; 2003; 2005) cohort studies that followed a group of active and inactive individuals for a duration of at least 10 years to predict how their activity levels and risk of disease changed over time. Anokye et al. (2012) explain how their identification and use of the cohort studies has strengthened previous modelling attempts in the field of physical activity. Campbell et al.(2015) reported this approach has enabled more conservative assumptions to be made around changing physical activity levels and disease development over time.

Table 2. Time horizon and types of outcomes compared to costs

Trial-based economic evaluations		
Study & Year of publication	Time Horizon (trial follow up)	Types of outcomes compared to costs per valuation technique
Larsen et al. 2015	Trial duration (12 months)	CEA: Cost per minute of increase in physical activity
Iliffe et al. 2014	Trial duration (12 months)	CEA: Cost per participant reaching or exceeding 150 minutes of moderate-to-vigorous physical activity per week
Young et al. 2012	Trial duration (3 months)	CEA: Cost per incremental change in daily estimated energy expenditure; CEA: Cost per the incremental change in minutes spent on moderate-to-vigorous activity
Elley et al. 2011	Trial duration (24 months; 12 months)	CEA: Cost per participant achieving 150 minutes of moderate intensity activity per week
Boehler et al. 2011	Trial duration (3 months)	CEA: Cost per participant achieving 150 minutes of moderate intensity activity per week
Shaw et al. 2011	Trial duration (12 months)	CEA: Cost per additional person achieving the target of a weekly increase of ≥ 15,000 steps.
Maddison et al. 2015	Trial duration (24 weeks / [6 months])	CEA: Cost per MET-hour of walking and leisure activity; CUA: Cost per short-term QALY gain

Leung et al. 2012	Trial duration (12 months)	CEA: Cost per 30 minutes of weekly leisure walking; CUA: Cost per short-term QALY gain
de Vries et al. 2016	Trial duration (6 months)	CUA: Cost per short-term QALY gain
Edwards et al. 2013; Murphy et al. 2012	Trial duration (12 months)	CUA: Cost per short-term QALY gain
Model-based economic evaluations		
Study & Year of publication	Time Horizon (trial follow up)	Types of outcomes compared to costs per valuation technique
Campbell et al. 2015	Lifetime	CUA: Cost per short-term QALY gain (mental health gain); Cost per QALYs associated with coronary heart disease, stroke, type 2 diabetes due to reduced risk for developing these health states
Anokye et al. 2012; Anokye et al. 2014	Lifetime	CUA: Cost per short-term QALY gain (mental health gain); Cost per QALYs associated with coronary heart disease, stroke, type 2 diabetes due to reduced risk for developing these health states CCA: Same outcomes outlined below for Pavey et al.'s (2011) CCA
Anokye et al., 2011	Lifetime	CUA: QALYs associated with coronary heart disease, stroke, type 2 diabetes due to reduced risk for developing these health states

Pavey et al. 2011	Lifetime	CUA: Cost per short-term QALY gain (mental health gain); Cost per QALYs associated with coronary heart disease, stroke, type 2 diabetes due to reduced risk for developing these health states CCA: Mental health (anxiety), Mental health (depression), Metabolic diabetes, Colon cancer, Breast cancer , Lung cancer, Hypertension (cardiovascular), Coronary Heart Disease, Stroke, Musculoskeletal (Osteroporosis), Musculoskeletal (Osteroarthritis), Lower back pain, Rhumatoid arthritis, Falls prevention, Absenteeism at work, Injury (disbenfit), Disability
Over et al. 2012	Lifetime	CUA: QALYs associated with myocardial infarction, stroke, diabetes, colorectal cancer, breast cancer due to reduced risk for developing these health states

RCT: randomised controlled trial; cRCT: cluster randomised controlled trial; CEA: Cost-effectiveness analysis; CUA: Cost-utility analysis; CCA: cost-consequence analysis ; MET: Metabolic Equivalent of Task

Challenge 2: Measuring and valuing outcomes

No studies in this present review conducted a cost-benefit analysis (CBA), despite health economists (Drummond et al., 2015) stating this approach is superior to cost-utility analysis (CUA) (Drummond et al., 2015). Recent UK and US guidelines recommended that studies report a broad range of outcomes alongside their economic analyses, through the use of approaches such as CBA, cost-consequence analysis (CCA) or an impact inventory (NICE, 2014; Sanders et al., 2016). Two studies (Anokye et al., 2012; Pavey et al., 2011) included a CCA conducted alongside a CUA. A broad range of health outcomes were included in their CCA (Table 2) yet the only non-health outcome reported was absenteeism.

Two thirds (n=11) of the studies presented just one type of valuation technique, either a CUA (n=5) or cost-effectiveness analysis (CEA) (n=6) (Table 2). Table 2 demonstrates further how despite having the same aim to increase physical activity levels and same valuation technique, the way results are presented to the end-user are inconsistent. Young et al. (Young et al., 2012) performed two CEAs reporting on the ‘cost per incremental change in daily estimated energy expenditure’ and ‘cost per incremental change in minutes spent on moderate-to-vigorous activity’. Three other studies (Boehler et al., 2011; Elley et al., 2011; Iliffe et al., 2014) performed a different type of CEA reporting on ‘cost per participant achieving 150 minutes of moderate physical activity per week’. The most common way to present the result of the valuation analysis was as ‘cost per short-term quality-adjusted life year (QALY) gain’. Nevertheless, this was reported for just under half (n=7) of the economic evaluations: four trial-based (de Vries et al., 2016; Edwards et al., 2013; Leung et al., 2012; Maddison et al., 2015) and three model-based (Anokye et al., 2012; Campbell et al., 2015; Pavey et al., 2011) studies. All model-based studies conceptualised the long-term gain in QALY in the same way, in terms of the QALYs gained due to not developing coronary heart disease, stroke or type 2 diabetes, or experiencing premature mortality. Over et al.’s (2012) analysis differed slightly, as they also included colorectal and breast cancer.

Rationale for the inclusion and exclusion of trial-derived QALYs varied considerably. Shaw et al. (Shaw et al., 2011) argued against the inclusion of trial-derived QALYs in their analysis, explaining it would be unnecessarily restrictive since evidence already shows that physical activity is associated with a reduction in NCD and premature mortality, which in turn is associated with a much greater gain in QALYs than trial-derived QALYs. Three model-based studies (Anokye et al., 2012; Campbell et al., 2015; Pavey et al., 2011) deemed it appropriate to incorporate both short-and long-term gain in QALYs. They conceptualised the short-term QALY gain as being a one-off gain in mental health, which they assumed would be achieved as a result of becoming physically active for at least 90 minutes per week. They assumed the one-off mental health benefit would last for just one year, which they claimed was a conservative assumption. Campbell et al. (2015) reported that their cost-effectiveness result was highly sensitive to the inclusion and exclusion of the one-off gain in mental health.

Challenge 3: Identifying intersectoral costs and consequences

The most commonly reported perspective was the health sector perspective (n=7) (Table 1 and 2). Six of the eight studies from the UK were from this perspective. In 2014, the UK reference case was updated to recommend the public sector perspective when conducting economic evaluations of public health interventions (NICE, 2014). The multi-agency public sector perspective adopted by Edwards et al. (2013) reflects the start of this paradigm shift. Two more recent UK studies (Campbell et al., 2015; Iliffe et al., 2014) did not adopt a public sector perspective. Despite studies being conducted from the same perspective, the type of costs identified as relevant varied within and across countries and intervention type. This weakness was identified through the quality assessment (Item 4 on Appendix D [INSERT LINK TO ONLINE FILE A, B, C, D & E]), as five studies (Boehler et al., 2011; de Vries et al., 2016; Maddison et al., 2015; Shaw et al., 2011; Young et al., 2012) did not relate their costs to a study perspective. More specifically, two studies did not report their perspective (Maddison et al., 2015; Shaw et al., 2011) and three included a narrower range of costs and consequences than would be expected for their stated perspective (Boehler et al., 2011; de Vries et al., 2016; Young et al., 2012). For example, two studies stated their study was from the societal perspective yet assessed only direct intervention costs and short-term healthcare savings (de Vries et al., 2016; Young et al., 2012), which were the same costs as studies which stated taking a health sector perspective (Table 1 and 2). Weatherly et al. (2009) also found that many studies included only a narrow range of costs within their stated study perspectives.

Figure 2. Cost categories identified across all 15 included studies

Figure 2 shows that seven cost categories were identified across all 15 included studies. Like the findings in this review, Alayli-Goebbels et al. (2014) found the most common type of cost reported was the intervention costs, followed by healthcare costs. Participant out-of-pocket expenses and productivity losses appeared in only a small proportion of studies in this review and Alayli-Goebbels et al.'s (2014) review. Although most studies looked at both the direct and indirect costs of the interventions, only Edwards et al. (2013) looked at the unintended productivity costs to the provider. More specifically, they examined whether the provider where the intervention was set (the leisure centre) experienced a loss in revenue, as a result of providing the intervention.

Challenge 4: Incorporating equity

The two former reviews found that authors did not routinely consider equity in their analysis (Alayli-Goebbels et al., 2014; Weatherly et al., 2009). Table 3 shows that all but one study (Shaw et al., 2011) included in the present review did consider equity. All but one study (Edwards et al., 2013) did this implicitly, conducting subgroup analyses of the cost-effectiveness result (n=6) or targeting the intervention at a population deemed in need of intervention (n=8). Edwards et al. (2013) were the only authors to explicitly discuss equity and to consider socio-economic status in their equity analysis. They did this by asking participants from areas of different levels of deprivation about how much they would be willing to pay to participate in the intervention of interest; thus informing the reader about participants' economic preferences. Notably this was an exploratory analysis and so the results were not incorporated in the CUA.

Table 3. Types of equity considered

	Campbell et al. 2015	Pre-existing condition
	Pavey et al. 2011	Pre-existing condition

Subgroup analyses of cost-effectiveness result	Anokye et al. 2011	Pre-existing condition
	Edwards et al. 2013; Murphy et al. 2012	Medical diagnosis
		Referral reason
		Adherence to scheme
		Gender
		Inequalities
		Age group
	Over et al. 2012	Age group
	Anokye et al.2012 ; Anokye et al. 2014	Age group
Intervention targeted at equity group	de Vries et al.2016	Frail older adults with mobility problems
	Leung et al.2012	Older adults
	Iliffe et al. 2014	Older adults
	Boehler et al. 2011	Older adults
	Maddison et al. 2015	People with ischaemic heart disease
	Elley et al. 2011	Females
	Young et al. 2012	Females
	Larsen et al. 2015	Latinas
Willing to pay question	Edwards et al. 2013; Murphy et al. 2012	Socio-economic status (level of deprivation)

Table 3 details the eight studies which targeted their intervention at a specific population group as well as the six studies that performed subgroup analyses of their cost-effectiveness result. Older adults was the most common equity subgroup targeted for intervention (Boehler et al., 2011; de Vries et al., 2016; Iliffe et al., 2014; Leung et al., 2012), followed by females (Elley et al., 2011; Young et al., 2012). The most common subgroup analyses were on pre-existing condition/ medical diagnosis (Anokye et al., 2011; Campbell et al., 2015; Edwards et al., 2013; Pavey et al., 2011) and age group (Anokye et al., 2012; Edwards et al., 2013; Over et al., 2012). Edwards et al.(2013) carried out seven types of equity analyses, all other authors conducted just one type. Furthermore, no studies attempted alternative equity analyses, such as an equity constraint or equity weighing analysis (Cookson et al., 2017).

New studies

The results of the rapid systematic scoping search are presented in Appendix B [INSERT LINK TO ONLINE FILE A, B, C, D & E]. In brief, four additional studies were identified as meeting the inclusion criteria of this review. Notably, one study (Gao et al., 2018) was an intervention targeting sedentary behaviour as an independent risk factor from physical activity. Furthermore, two studies (Gao et al., 2018; Harris et al., 2018) were both trial-and model-based economic evaluations, as the analysts had extrapolated their within-trial results a lifetime horizon.

Discussion

This review identified 15 economic evaluations of interventions that targeted physically inactive adults, and no economic evaluations of interventions that targeted sedentary adults (where sedentary behaviour was addressed an independent risk factor from physical activity). Like Abu-Omar et al's (2017) review of reviews which focuses on the results of economic evaluations, this present review identified economic evaluations on a limited range of physical activity interventions (Abu-Omar et al., 2017). Studies came from just four high-income countries, with over half (n=8) coming from the UK. This points to an important evidence gap in countries where economic evaluations are deemed appropriate. Examining a country's traditional beliefs around personal responsibility, efficiency and equity can explain why countries such as France and Germany are low users of economic evaluations and can in part explain why no studies in this review originated from these countries (Torbica et al., 2018). Regardless of cultural and institutional differences, globally health economists agree economic evaluations of preventative interventions are expected to have an important impact on future healthcare decision-making (ISPOR, 2018). In order to answer upcoming complex public health challenges, researchers need to go beyond clinical effectiveness methods and use a multidisciplinary suite of methods (Rutter et al., 2017) which includes economic evaluation. A prerequisite for this is an understanding on how key methodological challenges can be addressed.

Challenge 1: Attribution of effects

Modelling exercises

All ten trial-based economic evaluations in this review had a short time horizon; meaning they did not attempt to extrapolate or model the long-term impact of the intervention which could be used to informer longer term investment decision making. Any future reduction in incidence of NCD and premature mortality, attributable to physical activity and sedentary behaviour interventions, is unlikely to manifest until decades after the intervention has taken place. Yet, evaluating these interventions over the wrong timeframe means these interventions may appear ineffective or markedly less effective; they are at risk of not being appropriately prioritised by policymakers (Rutter et al., 2017). Curative interventions that rescue people from very poor health to better health will continue to be favoured, even if they are less cost-effective overall. Alayli-Goebbels et al. (2014) had previously suggested modelling as a way to extend the time horizon of trial-based studies, yet none of the ten trial-based studies in this review performed any modelling exercises. The challenges which can preclude extrapolation include the availability of data, and time and skills of the analyst (Squires et al., 2016).

Cohort studies

Campbell et al. (2015) and Anokye et al. (2012) were the only two studies in this review to identify additional evidence to link up their short- and long-term effect estimate. The three other model-based studies claimed there was insufficient evidence to verify the accuracy of their assumptions (Anokye et al., 2011; Over et al., 2012; Pavey et al., 2011). Notably, the cohort studies which Campbell et al.(2015) and Anokye et al.(2012) draw on were published several years prior to the publication of the three other model-based studies. This suggests that the methodological challenge of 'attribution of effect' may be more dependent upon the analysts' time and skills as opposed to the availability of data.

Challenge 2: Measuring and valuing outcomes

Cost-effectiveness and cost-utility analyses

This review found large inconsistencies in the types of outcomes measured and valued. There is no agreed classification system for physical activity outcomes(Abu-Omar et al.,

2017) since the analysis of raw objective accelerometer data measuring objective physical activity levels is still in its infancy. Presenting a limited range of results can reduce the applicability of the study's findings to other policymakers. Authors' views also differed firstly on whether short-term QALYs should be included in the economic analysis, secondly on whether a short-term QALY gain represented a one-off gain in mental health or general functional health. Presently, within the economic literature the responsiveness of the EQ-5D-3L to detect important differences in the severity of health is being challenged, and had led to the development of the EQ-5D-5L, which measures health on five levels as opposed to just three (Glick et al., 2014). This review has shown that outcomes used in physical activity studies are diverse; therefore, there is a need for analysts to agree on a consistent outcome that best captures the objectives of a physical activity intervention.

Cost-benefit and cost-consequence analyses

No studies in this review performed a CBA and just two presented a CCA alongside their full economic evaluation. There is a lack of CBAs in other public health areas. Hill et al. (2017) and Alayli-Goebbels et al. (2014) identified a small proportion of studies (n=1 and n=8 respectively) who reported conducting a CBA, but due to insufficient reporting gained limited insight into how these were performed such as how outcomes had been monetised (Alayli-Goebbels et al., 2014; Hill et al., 2017). Likewise, four studies claimed to be CBAs in the review by Weatherly et al.(2009), but after further assessment were re-classified as CCAs (n=3) and a CEA (n=1). Although classified as a partial-economic evaluation, CCA is a useful alternative to CBA since all relevant costs and consequences can be presented to the reader in the form of an inventory, rather than simplified into a single outcome measure or index as is the case in CEA and CUA, respectively. If an outcome is deemed relevant to the reader, they can reanalyse the data quantified in the CCA. However, CCA puts more onus on decision makers than CBA or CUA, as it does not roll outcomes into a summary measure that can be compared to a decision rule. An example of a decision rule in the UK is: invest where the incremental cost-effectiveness ratio is less than £30,000 per QALY (NICE, 2014).

Challenge 3: Identifying intersectoral costs and consequences

Inconsistent perspectives

The three most common perspectives stated were the health system, payer and societal perspectives. These match the three most commonly reported perspectives in the broader field of economic evaluation (Husereau et al., 2013). Only Edwards et al. (2013) conducted their analysis from the public sector perspective, a perspective recently recommended in the UK reference case (NICE, 2014). That said, Edwards et al. (2013) did not incorporate participant costs in their CUA, only through an exploratory analysis. Only three studies considered the cost to the participant, which is not surprising since the health sector perspective was the most commonly stated perspective. Participant and voluntary sector costs are deemed important, but previously have not been routinely captured (Weatherly et al., 2009).

It was found that even economic evaluations stated the same perspective did not always include the same costs and consequences. This is likely to be because there is a lack of standard definitions for the various perspective types (Husereau et al., 2013). Even where there are examples of standard definitions, such as those proposed by the Second US Panel on Cost-Effectiveness in Health and Medicine (Sanders et al., 2016), not all economists agree with their definitions, and furthermore the definitions may not be applicable to other countries since there are distinct features of each health system (Torbica et al., 2018). For instance, deciding what costs and consequences to capture within a societal perspective is a normative question, requiring the analyst to make social value judgements (Drummond et al., 2015). This is an important issue, since the exclusion of relevant consequences can lead to an underestimation of cost-effectiveness whilst the exclusion of relevant costs can lead to an overestimation of cost-effectiveness (Hill et al., 2017).

Cost categories identified

The cost categories identified in this review match the five cost categories (healthcare services, intervention costs, patient and family costs, lost productivity costs, future costs) identified as most relevant for inclusion in economic evaluations, by health economists who recently took part in a cross-Europe Delphi study (van Lier et al., 2017). This suggests analysts' choice in costs in this review align with analysts in the more general field of economic evaluation. It should be noted however that there was a difference in one of the categories, as family costs were not identified as a relevant cost category in the studies from this present review. Just two trial-based studies included absenteeism in their study; similarly only two of the model-based studies included it in their CCA. It continues to be debated in the literature as to whether absenteeism is an outcome of cost-offset, and thus whether it should be included in the numerator or denominator part of the incremental cost-effectiveness fraction (Drummond et al., 2015).

Challenge 4: Incorporating equity considerations

Presenting results by subgroups

Equity impact analysis can be as straightforward as presenting cost-effectiveness results by equity subgroups (Alayli-Goebbels et al., 2014; Hill et al., 2017; Weatherly et al., 2009). Six studies in this review presented an equity impact analysis (Anokye et al., 2012; Anokye et al., 2011; Campbell et al., 2015; Edwards et al., 2013; Over et al., 2012; Pavey et al., 2011). The most common subgroup analysed was individuals with pre-existing medical conditions, nevertheless this analysis was performed in just four studies (Anokye et al., 2011; Campbell et al., 2015; Edwards et al., 2013; Pavey et al., 2011). Furthermore, only one study (Edwards et al., 2013) conducted more than one type of equity subgroup analysis. These findings suggest analysts are not performing equity analyses in a comprehensive nor consistent manner. Weatherly et al. (2009) outlined socio-economic status as an important under-researched equity issue in economic evaluations, however only one study in this review researched socio-economic status by asking participants about their willingness to pay for an intervention component (Edwards et al., 2013). Incorporating equity into decisions on physical activity and sedentary behaviour interventions is especially important, since it is amongst the lower socioeconomic groups where physical inactivity is greatest (OECD, 2015).

New studies

Overall, the four studies published since March 2017 did not change the narrative of this review since there remains a dearth of economic evaluations in the field of physical activity and sedentary behaviour. What the studies have demonstrated is that firstly, there is an indication that health economic methods have begun to be applied to targeted sedentary behaviour interventions (Gao et al., 2018). Secondly, that it is feasible and informative to extrapolate beyond the trial (Gao et al., 2018; Harris et al., 2018).

Strengths and limitations

This is the first systematic review conducted since 2009 to review the methods used in economic evaluations of interventions targeted at physically inactive individuals, and the first systematic review to search for economic evaluations targeting sedentary behaviour as an independent risk factor from physical activity. This review included comprehensive literature searching and a rigorous methodology in line with the PRISMA guidelines (Moher et al., 2009). Economic evaluations aim to inform resource allocation decisions (Drummond et al., 2015). Previous reviews have demonstrated that key methodological challenges preclude economic evaluations in the field of public health from achieving this aim (Alayli-Goebbels et al., 2014; Weatherly et al., 2009). By focusing on physical activity and sedentary behaviour, this review has been able to not just provide an overview on whether or not the four key methodological challenges have been addressed in the last decade, but crucially explain in greater depth the methods performed in those few studies where progress has been made.

More specifically, progress has been observed in the 14 studies which have considered equity in their analysis (Table 3) and the small proportion of studies where either: the long-term model presented has been informed by robust epidemiological evidence (Anokye et al., 2012; Campbell et al., 2015); all important and relevant costs and consequences have been outlined to the reader in the form of a cost-consequence analysis (Anokye et al., 2012; Pavey et al., 2011); and/or a multi-sector perspective has been selected (Edwards et al., 2013). An output from the narrative synthesis of this review is a number of recommendations (as outlined in Table 4) explaining how analysts can continue to make progress towards addressing the four methodological challenges. Although, the comprehensive search strategy only goes upto March 2017, a rapid systematic scoping search is presented which highlights four new empirical studies. Two of these studies (Gao et al., 2018; Harris et al., 2018) support the recommendations emerging from this review in terms of linking up the intermediate evidence with longer term policy relevant outcomes.

It was not within the scope of this research to review the methods used in population-level interventions such as national policies or media campaigns. It would therefore be useful for future reviews to explore how economic evaluations are being carried out within this area. In addition, this review focuses on the methods conducted in full economic evaluations and so there is scope to review the methods used in partial evaluations. Nevertheless, full economic evaluations are deemed more informative than partial evaluations, and so it would have been expected that analysts would conduct for instance, a CCA alongside their full economic evaluation, as was done in two studies (Anokye et al., 2014; Pavey et al., 2011) in this review.

Recommendations

Table 4 presents a list of recommendations for researchers and users of economic evaluations from a variety of disciplines (health economics, public health, physical activity etc) to refer to when designing, analysing and appraising economic evaluations of targeted physical activity and sedentary behaviour interventions.

598 **Table 4. Recommendations for future economic evaluations**

Challenge	Recommendation	Explanation
Challenge 1. Attribution of Effects	*Visual representations of disease pathways*	It is necessary for public health researchers to invest time in reviewing the existing evidence base and develop novel modelling skills. Best practice guidelines state well established published models are preferred to those developed specifically for a trial (Ramsey et al., 2015). If skill and time permits, analysts can draw on the structure of the published models (Anokye et al., 2012; Campbell et al., 2015) identified in this review and adapt them according to the local decision-making context. All five models in this review presented a visual depiction of the disease pathway for physical activity. Authors from non-economic disciplines could build on the disease pathways presented in the model-based studies in this review, in order to help policymakers and those designing interventions to consider the long-term costs and consequences of investing or disinvesting in physical activity interventions. The visual could be as simple as a logic model, a visual tool recommended for public health interventions (Moore et al., 2015).
	Long-term objective data derived from cohort studies	Future investment and disinvestment decisions should be informed by economic evaluations which not only assess the short-term impact of interventions, but also impact on the medium- and long-term (Academy of Medical Sciences, 2016). As long-

		term RCTs of physical activity and sedentary behaviour interventions are likely to be impractical or unethical, evidence from non-experimental studies such as cohort studies could be drawn on to evidence the long-term impact of physical activity and sedentary behaviour interventions as done in two studies. In the hierarchy of evidence, cohort studies are recognised as being the next best alternative to RCTs (Murad et al., 2016). The popularity of wireless-enabled wearable activity monitors in high-income countries present researchers with an opportunity to conduct more cohort studies and collect objective data on behaviour change over a longer time period.
Challenge 2. Measuring and valuing outcomes	*Quality of life measurement tools*	Future research should aim to understand whether a short-term gain in QALY represents a one-off benefit in mental health due to becoming physically active. The EQ-5D tool, is the most commonly used tool to measure QALYs but only captures the functional health of an individual. Future studies could use other recently developed quality of life tools such as the ICECAP-A (Al-Janabi et al., 2012; Al-Janabi et al., 2013; Flynn et al., 2015), which has been designed to capture capability in a broader sense, beyond functional health. Another solution is for analysts to agree on a tool which crosswalks between physical activity outcomes and a summary tool like the EQ-5D. There is currently a mapping database of studies that map the EQ-5D tool to other

		outcomes measures (Dakin et al., 2018). No studies on the database have mapped a physical activity specific tool to the EQ-5D; future research should address this gap.
	Cost-consequence analysis	There is a need for further methodological developments in the monetisation of effects in CBAs (Drummond et al., 2015; Sanders et al., 2016). In the meantime, it is deemed more appropriate to conduct a good quality CUA which may be of a narrower perspective, than a poor quality CBA which captures a broader perspective (Hill et al., 2017; Weatherly et al., 2009). In order to report on multiple outcomes which extend beyond health, a CCA or impact inventory conducted alongside a full economic evaluation is recommended (NICE, 2014; Sanders et al., 2016). If the word limit in journals precludes authors from presenting a CCA in the main manuscript, they should present this information in the online supplementary material.
Challenge 3. Identifying intersectoral costs and consequences	*Multi-sector perspective*	Three studies in this review omitted costs which would typically be deemed relevant to their stated perspective, and two studies did not report their perspective. It is imperative for analysts to describe and justify the costs and consequences which they have deemed relevant for their chosen perspectives (Husereau et al., 2013). Inevitably different assumptions on what costs and consequences are included in the analysis leads to different results (Sanders et al., 2016). Furthermore, future studies should aim

		to present at least two types of perspectives and conduct a CCA or impact inventory alongside their CUA or CEA in order to present the various relevant costs and consequences to the various relevant sectors (Alayli-Goebbels et al., 2014; Sanders et al., 2016; Weatherly et al., 2009). A multi-sector perspective where costs and consequences are presented in their disaggregated form (i.e. in a CCA) for each sector is preferred over stating a societal perspective (Drummond et al., 2015; Hill et al., 2017).
	Systems thinking approach	Absenteeism was the only non-health effect identified in the two CCAs in this review (Anokye et al., 2012; Pavey et al., 2011). During the design stage of future economic evaluations analysts could conduct multi-stakeholder and expert consultations to map out which costs and consequences are deemed relevant to physical activity and sedentary behaviour interventions (Squires et al., 2016). A systems thinking approach (Rutter et al., 2017; Squires et al., 2016) is recommended to ensure interventions' indirect and unintended costs and consequences on the whole system are considered, not just those experienced by the health sector or payer. Two recently published frameworks can help analysts apply a systems approach (Cylus et al., 2016; Squires et al., 2016).

Challenge 4. **Incorporating equity**	*Equity impact analysis*	Analysts should present costs and consequences explicitly in their disaggregated form for various equity groups, so policymakers can start to build a better picture on which population groups gain and lose from a specific decision (Hill et al., 2017). From here, analysts can conduct an equity impact analysis. This type of analysis is deemed easier than conducting equity constraint or equity weighting analysis (Hill et al., 2017). The equity effectiveness loop framework (Welch et al., 2008) and PROGRESS-Plus framework (O'Neill et al., 2014) are recommended to help analysts consider, in a structured way, which equity factors may be relevant to their study (Alayli-Goebbels et al., 2014; Welch et al., 2017).
	Participant's preferences	Other types of equity-related analyses not identified in this review, but which future studies could investigate, include the public's perspective on trading off efficiency with equity (in public services) (Weatherly et al., 2009). It is also recommended that future studies, specifically trial-based studies, capture economic information on time, travel and out-of-pocket expenses incurred by the participant. The APEASE criteria by Michie et al. (2014) could also help analysts to consider the acceptability and affordability of an intervention to various stakeholders. Inevitably, these two issues will contribute to the success of interventions aiming to change behaviour (Michie et al., 2014).

Conclusions

A focus on the key methodological challenges in economic evaluations is important, as they can impact on the derived cost-effectiveness result, which ultimately can impact on a policymaker's resource allocation decision. As economic evaluation is a rapidly developing field (Drummond et al., 2015) this systematic review has provided an important update on the most recent methods used in targeted physical activity interventions. The review has also highlighted there is a scarcity of economic evaluations for targeted sedentary behaviour interventions. Importantly, this review makes it explicit to policymakers and researchers from the varied disciplines in which physical activity and sedentary behaviour falls under, that there are still key methodological challenges that need further attention. This review has highlighted that methodological choices vary widely not just between countries but also within them. Ultimately, these analyst-based choices affect the results presented and subsequent resource allocation decisions made. A recent consensus statement has called for collaboration across the disciplines to develop guidance specific to the context of economic evaluations of physical activity interventions (Davis et al., 2014). To date, no guidelines have been developed to address this need. The examples of methodological development identified from the studies in this review and the resulting review recommendations can be used to inform future guidelines and their supplementary materials.

References

Abu-Omar, K., Rütten, A., Burlacu, I., Schätzlein, V., Messing, S., & Suhrcke, M. (2017). The cost-effectiveness of physical activity interventions: A systematic review of reviews. *Preventive Medicine Reports,* 8, 72-78.

Academy of Medical Sciences. (2016). Improving the health of the public by 2040: optimising the research environment for a healthier, fairer future (September, 2016).

Al-Janabi, H., Flynn, T.N., & Coast, J. (2012). Development of a self-report measure of capability wellbeing for adults: the ICECAP-A. *Quality of Life Research,* 21, 167-176.

Al-Janabi, H., Peters, T.J., Brazier, J., Bryan, S., Flynn, T.N., Clemens, S., et al. (2013). An investigation of the construct validity of the ICECAP-A capability measure. *Quality of Life Research,* 22, 1831-1840.

Alayli-Goebbels, A.F.G., Evers, S.M.A.A., Alexeeva, D., Ament, A.J.H.A., de Vries, N.K., Tilly, J.C., et al. (2014). A review of economic evaluations of behavior change interventions: setting an agenda for research methods and practice. *Journal of Public Health,* 36, 336-344.

Anokye, N.K., Lord, J., & Fox-Rushby, J. (2012). National Institute for Health and Clinical Excellence: Public Health Intervention Guidance on Physical Activity – Brief advice for adults in primary care: Economic Analysis. pp. 1-72). London: National Institute for Health and Clinical Excellence,.

Anokye, N.K., Lord, J., & Fox-Rushby, J. (2014). Is brief advice in primary care a cost-effective way to promote physical activity? *British Journal of Sports Medicine,* 48, 202-206.

Anokye, N.K., Trueman, P., Green, C., Pavey, T., Hillsdon, M., & Taylor, R. (2011). The cost-effectiveness of exercise referral schemes. *BMC Public Health,* 11, 954.

Boehler, C.E.H., Milton, K.E., Bull, F.C., & Fox-Rushby, J.A. (2011). The cost of changing physical activity behaviour: evidence from a "physical activity pathway" in the primary care setting. *BMC Public Health,* 11, 370.

Campbell, F., Holmes, M., Everson-Hock, E., Davis, S., Buckley Woods, H., Anokye, N., et al. (2015). A systematic review and economic evaluation of exercise referral schemes in primary care: a short report. *Health Technol Assess,* 19, 1-110.

Cookson, R., Mirelman, A.J., Griffin, S., Asaria, M., Dawkins, B., Norheim, O.F., et al. (2017). Using Cost-Effectiveness Analysis to Address Health Equity Concerns. *Value in Health,* 20, 206-212.

Cylus, J., Papanicolas, I., & Smith, P.C. (2016). A framework for thinking about health system efficiency. *Health System Efficiency*, 1.

Dahlgren, G., & Whitehead, M. (1991). Policies and strategies to promote social equity in health (Mimeo). . Stockholm, Sweden: Insititute for Future Studies,.

Dakin, H., Abel, L., Burns, R., & Yang, Y. (2018). Review and critical appraisal of studies mapping from quality of life or clinical measures to EQ-5D: an online database and application of the MAPS statement. *Health and Quality of Life Outcomes,* 16, 31.

Davis, J.C., Verhagen, E., Bryan, S., Liu-Ambrose, T., Borland, J., Buchner, D., et al. (2014). 2014 Consensus Statement from the first Economics of Physical Inactivity Consensus (EPIC) Conference (Vancouver). *British Journal of Sports Medicine,* 48, 947-951.

de Vries, N.M., Staal, J.B., Teerenstra, S., Adang, E.M.M., Rikkert, M.G.M.O., & Nijhuis-van der Sanden, M.W.G. (2013). Physiotherapy to improve physical activity in community-dwelling older adults with mobility problems (Coach2Move): study protocol for a randomized controlled trial. *Trials,* 14, 434.

de Vries, N.M., Staal, J.B., van der Wees, P.J., Adang, E.M., Akkermans, R., Olde Rikkert, M.G., et al. (2016). Patient-centred physical therapy is (cost-) effective in increasing physical activity and reducing frailty in older adults with mobility problems: a randomized controlled trial with 6 months follow-up. *J Cachexia Sarcopenia Muscle,* 7, 422-435.

Drummond, M.F., Schulpher, M.J., Claxton, K., Stoddart, G.L., & Torrance, G.W. (2015). *Methods for the Economic Evaluation of Health Care Programmes (Fourth Edition)*. Oxford: Oxford University Press.

Edwards, R.T., Linck, P., Hounsome, N., Raisanen, L., Williams, N., Moore, L., et al. (2013). Cost-effectiveness of a national exercise referral programme for primary care patients in Wales: results of a randomised controlled trial. *BMC Public Health,* 13.

Ekelund, U., Steene-Johannessen, J., Brown, W.J., Fagerland, M.W., Owen, N., Powell, K.E., et al. (2016). Does physical activity attenuate, or even eliminate, the detrimental association of sitting time with mortality? A harmonised meta-analysis of data from more than 1 million men and women. *Lancet,* 388, 1302-1310.

Elley, C.R., Garrett, S., Rose, S.B., O'Dea, D., Lawton, B.A., Moyes, S.A., et al. (2011). Cost-effectiveness of exercise on prescription with telephone support among women in general practice over 2 years. *Br J Sports Med,* 45, 1223-1229.

Flynn, T.N., Huynh, E., Peters, T.J., Al-Janabi, H., Clemens, S., Moody, A., et al. (2015). Scoring the ICECAP-A capability instrument. Estimation of a UK general population tariff. *Health Economics,* 24, 258-269.

Gao, L., Flego, A., Dunstan, D.W., Winkler, E.A.H., Healy, G.N., Eakin, E.G., et al. (2018). Economic evaluation of a randomized controlled trial of an intervention to reduce office workers' sitting time: the "Stand Up Victoria" trial. *Scandinavian Journal of Work, Environment & Health*, 503-511.

Gc, V., Wilson, E.C.F., Suhrcke, M., Hardeman, W., & Sutton, S. (2016). Are brief interventions to increase physical activity cost-effective? A systematic review. *British Journal of Sports Medicine,* 50, 408-417.

Glick, H.A., Doshi, J.A., Sonnad, S.S., & Polsky, D. (2014). *Economic Evaluation in Clinical Trials*: Oxford University Press.

Harris, T., Kerry, S.M., Limb, E.S., Furness, C., Wahlich, C., & Victor, C.R. (2018). Physical activity levels in adults and older adults 3-4 years after pedometer-based walking interventions: long-term follow-up of participants from two randomised controlled trials in UK primary care. 15, e1002526.

Hill, S.R., Vale, L., Hunter, D., Henderson, E., & Oluboyede, Y. (2017). Economic evaluations of alcohol prevention interventions: Is the evidence sufficient? A review of methodological challenges. *Health Policy,* 121, 1249-1262.

Hu, G., Jousilahti, P., Borodulin, K., Barengo, N.C., Lakka, T.A., Nissinen, A., et al. (2007). Occupational, commuting and leisure-time physical activity in relation to coronary heart disease among middle-aged Finnish men and women. *Atherosclerosis,* 194, 490-497.

Hu, G., Qiao, Q., Silventoinen, K., Eriksson, J.G., Jousilahti, P., Lindstrom, J., et al. (2003). Occupational, commuting, and leisure-time physical activity in relation to risk for Type 2 diabetes in middle-aged Finnish men and women. *Diabetologia,* 46, 322-329.

Hu, G., Sarti, C., Jousilahti, P., Silventoinen, K., Barengo, N.C., & Tuomilehto, J. (2005). Leisure time, occupational, and commuting physical activity and the risk of stroke. *Stroke,* 36, 1994-1999.

Husereau, D., Drummond, M., Petrou, S., Carswell, C., Moher, D., Greenberg, D., et al. (2013). Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. *BMJ : British Medical Journal,* 346.

Iliffe, S., Kendrick, D., Morris, R., Masud, T., Gage, H., Skelton, D., et al. (2014). Multicentre cluster randomised trial comparing a community group exercise programme and home-based exercise with usual care for people aged 65 years and over in primary care. *Health Technol Assess,* 18, vii-xxvii, 1-105.

ISPOR. (2018). Top 10 HEOR Trends.

Kelly, M.P., McDaid, D., Ludbrook, A., & Powell, J.E. (2005). Economic appraisal of public health interventions.: Health Development Agency,.

Kohl, H.W., Craig, C.L., Lambert, E.V., Inoue, S., Alkandari, J.R., Leetongin, G., et al. (2012). The pandemic of physical inactivity: global action for public health. *The Lancet,* 380, 294-305.

Kolt, G.S., Schofield, G.M., Kerse, N., Garrett, N., Schluter, P.J., Ashton, T., et al. (2009). The Healthy Steps Study: A randomized controlled trial of a pedometer-based Green Prescription for older adults. Trial protocol. *BMC Public Health,* 9, 404.

Larsen, B., Gilmer, T., Pekmezi, D., Napolitano, M.A., & Marcus, B.H. (2015). Cost effectiveness of a mail-delivered individually tailored physical activity intervention for Latinas vs. a mailed contact control. *The International Journal of Behavioral Nutrition and Physical Activity,* 12, 140.

Leung, W., Ashton, T., Kolt, G.S., Schofield, G.M., Garrett, N., Kerse, N., et al. (2012). Cost-effectiveness of pedometer-based versus time-based Green Prescriptions: the Healthy Steps Study. *Aust J Prim Health,* 18, 204-211.

Liberati, A., Altman, D.G., Tetzlaff, J., Mulrow, C., Gøtzsche, P.C., Ioannidis, J.P.A., et al. (2009). The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration. *PLoS Medicine,* 6, e1000100.

Maddison, R., Pfaeffli, L., Whittaker, R., Stewart, R., Kerr, A., Jiang, Y., et al. (2015). A mobile phone intervention increases physical activity in people with cardiovascular disease: Results from the HEART randomized controlled trial. *Eur J Prev Cardiol,* 22, 701-709.

Marmot, M., & Allen, J.J. (2014). Social Determinants of Health Equity. *American Journal of Public Health,* 104, S517-S519.

Michie, S., Atkins, L., & West, R. (2014). *The Behaviour Change Wheel: A Guide to Developing Interventions.* . London: Silverback Publishing.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., & The, P.G. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Medicine,* 6, e1000097.

Moore, G.F., Audrey, S., Barker, M., Bond, L., Bonell, C., Hardeman, W., et al. (2015). Process evaluation of complex interventions: Medical Research Council guidance. *BMJ : British Medical Journal,* 350.

Murad, M.H., Asi, N., Alsawas, M., & Alahdab, F. (2016). New evidence pyramid. *Evidence Based Medicine,* 21, 125-127.

NICE. (2014). National Institute for Health and Care Excellence: Incorporating economic evaluation.

O'Neill, J., Tabish, H., Welch, V., Petticrew, M., Pottie, K., Clarke, M., et al. (2014). Applying an equity lens to interventions: using PROGRESS ensures consideration of socially stratifying factors to illuminate inequities in health. *Journal of Clinical Epidemiology,* 67, 56-64.

OECD. (2015). Organisation for Economc Cooperation and Development; World Health Organization: Promoting Health, Preventing Disease.

Over, E.A.B., Wendel-Vos, G.C.W., van den Berg, M., Reenen, H.H.H.-v., Tariq, L., Hoogenveen, R.T., et al. (2012). Cost-effectiveness of counseling and pedometer use to increase physical activity in the Netherlands: a modeling study. *Cost Effectiveness and Resource Allocation : C/E,* 10, 13-13.

Owen, L., Pennington, B., Fischer, A., & Jeong, K. (2017). The cost-effectiveness of public health interventions examined by NICE from 2011 to 2016. *Journal of Public Health*, 1-10.

Pavey, T.G., Anokye, N., Taylor, A.H., Trueman, P., Moxham, T., Fox, K.R., et al. (2011). The clinical effectiveness and cost-effectiveness of exercise referral schemes: a systematic review and economic evaluation. *Health Technol Assess,* 15, i.

Popay, J., Roberts, H.M., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., et al. (2006). Guidance on the conduct of narrative synthesis in sytematic reviews. Institute for Health Research.

Ramsey, S.D., Willke, R.J., Glick, H., Reed, S.D., Augustovski, F., Jonsson, B., et al. (2015). Cost-Effectiveness Analysis Alongside Clinical Trials II—An ISPOR Good Research Practices Task Force Report. *Value in Health,* 18, 161-172.

Rutter, H., Savona, N., Glonti, K., Bibby, J., Cummins, S., Finegood, D.T., et al. (2017). The need for a complex systems model of evidence for public health. *The Lancet,* 390, 2602-2604.

Sanders, G.D., Neumann, P.J., Basu, A., & et al. (2016). Recommendations for conduct, methodological practices, and reporting of cost-effectiveness analyses: Second panel on cost-effectiveness in health and medicine. *JAMA,* 316, 1093-1103.

Shaw, R., Fenwick, E., Baker, G., McAdam, C., Fitzsimons, C., & Mutrie, N. (2011). 'Pedometers cost buttons': the feasibility of implementing a pedometer based walking programme within the community. *BMC Public Health,* 11, 200.

Squires, H., Chilcott, J., Akehurst, R., Burr, J., & Kelly, M.P. (2016). A Framework for Developing the Structure of Public Health Economic Models. *Value in Health,* 19, 588-601.

Torbica, A., Tarricone, R., & Drummond, M. (2018). Does the approach to economic evaluation in health care depend on culture, values, and institutional context? *The European Journal of Health Economics,* 19, 769-774.

van Lier, L.I., Bosmans, J.E., van Hout, H.P.J., Mokkink, L.B., van den Hout, W.B., de Wit, G.A., et al. (2017). Consensus-based cross-European recommendations for the identification, measurement and valuation of costs in health economic evaluations: a European Delphi study. *The European Journal of Health Economics*.

Wanless, D. (2004). Securing good for the whole population: Final Report. London: HM Treasury.

Weatherly, H.L.A., Cookson, R.A., & Drummond., M.F. (2014). Economic Evaluation of Public Health Interventions: Methodological Challenges. In A.J. Culyer (Ed.), *Encyclopedia of Health Economics* pp. 217-223). USA: Elesvier.

Weatherly, H.L.A., Drummond, M., Claxton, K., Cookson, R., Ferguson, B., Godfrey, C., et al. (2009). Methods for assessing the cost-effectiveness of public health interventions: Key challenges and recommendations. *Health Policy,* 93, 85-92.

Welch, V.A., Norheim, O.F., Jull, J., Cookson, R., Sommerfelt, H., & Tugwell, P. (2017). CONSORT-Equity 2017 extension and elaboration for better reporting of health equity in randomised trials. *BMJ,* 359.

Welch, V.A., Tugwell, P., & Morris, E.B. (2008). The equity-effectiveness loop as a tool for evaluating population health interventions. *Rev Salud Publica (Bogota),* 10 Suppl, 83-96.

WHO. (2018a). World Health Organisation: Noncommunicable diseases.

WHO. (2018b). World Health Organisation: Physical Activity: Key Facts.

Wu, S., Cohen, D., Shi, Y., Pearson, M., & Sturm, R. (2011). Economic analysis of physical activity interventions. *Am J Prev Med,* 40, 149-158.

Young, L., Zimmerman, L., Pozehl, B., Barnason, S., & Wang, H. (2012). Cost-effectiveness of a symptom management intervention: improving physical activity in older women following coronary artery bypass surgery. *Nurs Econ,* 30, 94-103.