STUDY PROTOCOL

Policy Evaluation Network (PEN): Protocol for systematic literature review examining the evidence for impact of school policies on physical activity [version 1; peer review: 1 approved with reservations]

Kevin Volf¹, Liam Kelly¹, Enrique García Bengoechea¹, Blathin Casey¹, Anna Gobis², Jeroen Lakerveld³, Joanna Zukowska², Peter Gelius⁴, Sven Messing⁴, Sarah Forberger⁵, Catherine Woods¹

¹Physical Activity for Health, Health Research Institute, Department of Physical Education and Sport Sciences, University of Limerick, Limerick, Ireland
²Faculty of Civil and Environmental Engineering, Gdansk University of Technology, Gdansk, Poland
³Amsterdam University Medical Center, Amsterdam, The Netherlands
⁴Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany
⁵Leibniz Institute for Prevention Research and Epidemiology - BIPS, Bremen, Germany

Abstract

Introduction: Over 40 million deaths annually are due to noncommunicable diseases, 15 million of these are premature deaths and physical inactivity attributes an estimated 9% to this figure. Global responses have included the sustainable development goals and the global action plan on physical activity. Both point to policy action in physical activity (PA) to address change, yet the impact of policy is unknown. The protocol described outlines a systematic literature review that will be undertaken by the Policy Evaluation Network to address this knowledge gap.

Protocol: This review of school PA policies is the first of seven planned reviews. The seven best investments for promotion of population PA identified in the Toronto Charter (whole-of-school programmes, transport policy, urban design policy, primary health care policy, public education policy, community programmes and sport programmes) will form the basis of these reviews. Seven individual scientific literature searches across six electronic databases, using key concepts of policy, PA, evaluation and a distinct concept for each area will be conducted. This will be supplemented with a search of the reference list of included articles. Methodological quality will be assessed and overall effectiveness for each included study will be
described according to pre-determined categories. **Conclusions:** The review will provide policy makers with a list of policy statements and corresponding actions which the evidence has determined impact on PA directly or indirectly. By collating the evidence, and demonstrating the depth of the science base which informs these policy recommendations, this review will provide guidance to policymakers to use evidence-based or evidence-informed policies to achieve the 15% relative reduction in physical inactivity as defined by the ‘Global Action Plan on Physical Activity’. **Registration:** PROSPERO CRD42020156630 (10/07/2020).

**Keywords**

physical activity, policy, protocol, systematic review, evaluation

**Corresponding author:** Catherine Woods (catherine.woods@ul.ie)

**Author roles:** Volf K: Conceptualization, Methodology, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; Kelly L: Methodology, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; García Bengoechea E: Conceptualization, Methodology, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing; Casey B: Writing – Review & Editing; Gobis A: Writing – Review & Editing; Lakerveld J: Conceptualization, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; Zuckowska J: Conceptualization, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; Gellius P: Conceptualization, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; Messing S: Conceptualization, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; Forberger S: Conceptualization, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; Woods C: Conceptualization, Methodology, Project Administration, Supervision, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

**Competing interests:** No competing interests were disclosed.

**Grant information:** Health Research Board [JP1-HDHL-PEN-2017-1]. The PEN project is funded by the Joint Programming Initiative (JPI) “A Healthy Diet for a Healthy Life”, a research and innovation initiative of EU member states and associated countries. The funding agencies supporting this work are (in alphabetical order of participating countries): Germany: Federal Ministry of Education and Research (BMBF); Ireland: Health Research Board (HRB); Italy: Ministry of Education, University and Research (MIUR); The Netherlands: The Netherlands Organisation for Health Research and Development (ZonMw); New Zealand: The University of Auckland, School of Population Health; Norway: The Research Council of Norway (RCN); Poland: The National Centre for Research and Development (NCBR). Additionally, the French partners acknowledge the support through the Institute National de la Recherche Agronomique (INRA).

**Copyright:** © 2020 Volf K et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**How to cite this article:** Volf K, Kelly L, García Bengoechea E et al. Policy Evaluation Network (PEN): Protocol for systematic literature review examining the evidence for impact of school policies on physical activity [version 1; peer review: 1 approved with reservations] HRB Open Research 2020, 3:62 https://doi.org/10.12688/hrbopenres.13089.1

**First published:** 04 Sep 2020, 3:62 https://doi.org/10.12688/hrbopenres.13089.1
Introduction

Physical activity (PA) is defined as “any bodily movement produced by skeletal muscles that requires energy expenditure” (Caspersen et al., 1985). The relationship between PA levels and health outcomes is well established (Rutten et al., 2016). Insufficient PA has been identified by the World Health Organisation (WHO) as the fourth leading risk factor for mortality worldwide (WHO, 2009) and in 2012 it was estimated that 9% (range 5.1 – 12.5%) of global premature mortality can be attributed to physical inactivity (Lee et al., 2012). The European region has been strongly affected by the costs of inactivity, absorbing 16.9% of the substantial disability that inactivity causes through its contribution to morbidity from coronary heart disease (CHD), cancer, stroke and diabetes and 21.8% of the healthcare cost (Ding et al., 2016; Lee et al., 2012).

This epidemiological evidence reveals inactivity to be a substantial public health issue and advocacy by public health specialists and the academic community has demanded policy responses to this issue. For the purposes of this document, policy should be understood as “decisions, plans, and actions that are implemented by national or regional governments to achieve specific health promotion goals within a society” (Lakerveld et al., 2020). As indicated by the WHO (WHO Regional Office for Europe, 2010), policy can give support, coherence and visibility at the political level, while making it possible for the organisations involved at national, regional, and local levels – e.g., national government sectors, regional or local authorities, stakeholders, and the private sector – to be logical and consistent in their actions to achieve a shared goal. This definition refers to food and PA environments, systems and behaviours (WHO Regional Office for Europe, 2010). In order to reflect the complexity of the policies that may affect the PA policy environment, a distinction is made between “direct” policies, which refers to policies where improving the PA environment and increasing participation is the primary aim, and “indirect” policies, where the primary aim of the policy is not to increase PA but this may occur as a co-benefit of successful implementation. The International Society for Physical Activity and Health (ISPAH) was established in 2009 (Kohl et al., 2012) and numerous articles and editorials in leading academic and medical journals have pointed out the need to address physical inactivity (Bull & Bauman, 2011; Das & Horton, 2012; Kohl et al., 2012; Woods & Mutrie, 2012).

At its third biennial congress ISPAH promulgated the Toronto Charter calling for political commitment to achieving greater opportunities for PA (Bull et al., 2010). To guide action on this issue the Charter was subsequently accompanied by a document titled Non Communicable Disease Prevention: Investments that Work for Physical Activity (2012). This document declared seven domains which evidence suggested could be effectively targeted to increase PA opportunities. These seven domains provide a structure for systematic literature review search.

Over recent years there has been an acceleration in the production of policy responses to the epidemics of PA and sedentary behaviour (Klepac Pogrmilovic et al., 2018). The Global Observatory for Physical Activity (GoPA) reports that by 2013, 139 countries were members of its PA advocacy alliance and 26.6% of these countries had already published a stand-alone PA plan (Ramirez Varela et al., 2016). Furthermore, in 2013, the WHO published a document which recognised PA as a part of the global effort to combat non-communicable diseases (NCDs) (WHO, 2013).

A significant development occurred in 2017 when, in response to demands for direction on the problem of physical inactivity, the WHO committed to publishing a stand-alone action plan on this issue. This commitment was realised in 2019 when the WHO published the Global Action Plan on Physical Activity (GAPPA), which targeted an even more ambitious PA target than the previous NCD plan (WHO, 2019).

The recent rise in the number of national PA policies allows research into the question of which of these policies are effective in increasing PA. A scoping review published in 2016 provided evidence that research into policy effectiveness lagged behind research that links PA to health and research that links PA interventions to behaviour (Rutten et al., 2016). However, with the increase in the number of PA policies there may have been a concomitant rise in research examining the effectiveness of these policies. Furthermore, to the best available knowledge, no project has linked existing policy statements with research that corroborates or discredits the effectiveness of these statements.

As part of the Joint Programming Initiative on a Healthy Diet for a Healthy Life, researchers from 28 institutes in seven European countries (France, Germany, Ireland, Italy, Norway, Poland, and the Netherlands) and New Zealand combine their expertise to form a Policy Evaluation Network (PEN) (Lakerveld et al., 2020). PEN’s vision is to provide Europe with tools to identify, evaluate and benchmark policies designed to directly or indirectly address unhealthy lifestyle behaviours which contribute to overweight and obesity, while accounting for existing health inequities. Using structured evaluation principles and methods, PEN will examine the content, implementation and impact of lifestyle policies across Europe and will build on existing knowledge. PEN will provide an overview of the ‘best’ public policies most likely to sustainably support more favourable health behaviours. This protocol paper outlines the methodology for the first of several complementary systematic literature reviews designed to determine the impact of policy on physical activity outcomes based on the different policy domains identified in the “seven best investments” document (ISPAH, 2012). This review will provide evidence supporting the development of a tool named the Physical Activity Policy Environment Index (PA EPI), based on similar principles to the food environment policy index (Food EPI) (Swinburn et al., 2013). The PA EPI will provide policy makers with a list of policy statements and corresponding actions which the evidence has determined improve physical activity outcomes across these settings.

Protocol

Original material examining the evidence of what works in terms of direct and indirect policies on PA will be identified in the following ways:
(1) A search, with no date restrictions, of the following electronic databases: four specialized sport science or biomedical databases, MEDLINE (Ebsco), SportDiscus, Cinahl, and Cochrane library, and two general social science databases, Web of Science and Scopus. Search results will be limited to articles that are identified through searching the titles and abstracts.

(2) Manual reference checks of identified original studies.

(3) Publicly available English-language resources and documents of major national and international stakeholders will be searched to identify existing reviews and position papers discussing the evidence of what works in terms of direct and indirect policies on PA, e.g. the WHO's European database on Nutrition, Obesity and Physical Activity (NOPA), Global Action Plan on Physical Activity (GAPPA), the European Physical Activity Strategy (EPAS) (WHO Regional Office for Europe, 2015) and the European Physical Activity Guidelines (EPAG) (European Commission, 2008).

A content analysis was performed on the Toronto Charter supplementary document (ISPAH, 2012). The document declares that the policy interventions should prioritise the following key areas for intervention: whole-of-school programmes, transport policy, urban design policy, primary health care policy, public education policy, community programmes and sport programmes. This document was searched for key words to be included in the search syntax. Researchers consulted with librarians and other research staff for suggestions on search terms.

The search of electronic databases will comprise seven individual searches (corresponding to the seven best investments), each one to be run on each of the databases. The seven searches will be formed by combining the same basic search strategy with seven distinct search concepts. The basic search strategy will consist of three search concepts (Table 1): search concept one (C1), which will combine synonyms for the keyword “policy” with the Boolean Operator “OR”; search concept two (C2), which will do the same with the keyword of “physical activity”; and search concept three (C3), which will do the same for the keyword of “impact”. The three search terms will be combined with the Boolean operator “AND” (Table 1).

Each of the seven searches will further be combined with a search term constructed to replicate one of the seven best investments declared in the document Non Communicable Disease Prevention: Investments that Work for Physical Activity (ISPAH, 2012) (Table 2). It is proposed that individual systematic literature reviews will be performed on each of the seven best investments, with this initial review focusing on the ‘whole-of-school approach’ category, and ‘transport policy’ and ‘public education’ categories to follow. The keywords were selected based on an iterative process which combined preliminary searches, solicitation of professional opinion and content analysis of the seven best investments document.

The following criteria will be applied for searches in databases: language will be limited to English language only.

Eligibility criteria
In order to answer our research question some eligibility criteria were developed to screen out irrelevant articles. Studies will be included based on the following criteria for: 1) type of study, 2) participants/population, 3) exposure/intervention, and 4) outcomes.

General “mandatory” eligibility criteria were formulated as well as specific “preferred” eligibility criteria for each of the seven searches. Publications that do not meet the “mandatory” eligibility criteria will be excluded from review. Publications that do not meet the “preferred” eligibility criteria will be set aside and possibly reassigned to a different search category if they are not duplicates of any publication already included in that search category.

Types of study to be included/excluded. No limitations regarding study type will be placed as long as the study design allows the research questions to be addressed. In addition, reviews using a comprehensive search strategy (including systematic, scoping and realist reviews) and analysing original research on the evidence of what works, and how, in terms of direct and

<table>
<thead>
<tr>
<th>Keyword</th>
<th>Synonyms</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Policy”</td>
<td>(MH “Policy”) OR (MH “Public Policy”) OR (MH “Policy Making”) OR (“policy”) OR (“policies”) OR (“national policy”) OR (“national framework”) OR (“policy framework”) OR (“policy action”) OR (“legislation”) OR (“strategy”) OR (“policy making”)</td>
</tr>
<tr>
<td>“Physical Activity”</td>
<td>(MH “Exercise”) OR (MH “Sedentary Behavior”) OR (“physical activit**) OR (“physical inactivity”) OR (“play”) OR (“physical education”) OR (“sedentar**) OR (“sitting”) OR (“healthy lifestyle”) OR (“health initiative”)</td>
</tr>
<tr>
<td>“Impact”</td>
<td>(“evaluat**”) OR (“impact”) OR (“appraisal”) OR (“effect**) OR (“assessment”)</td>
</tr>
<tr>
<td>“Whole of School Approach”</td>
<td>“Whole-of-school” OR “Whole School” OR “Whole of School” OR WSCC OR “school intervention” OR “school based intervention” OR “school initiative” OR “school based initiative” OR “school program**” OR “School health” OR “Wellness” OR “well-being”</td>
</tr>
</tbody>
</table>

Abbreviations: ‘MH’ = MeSH Heading.
indirect policies on PA; and reviews and policy analysis documents issued by major national and international stakeholder organisations addressing recommendations referring to the same evidence will be eligible for inclusion. Studies will be excluded based on the following criteria; a direct or indirect form of policy intervention is not identifiable, no information is provided regarding the effects of the policy under consideration on the desired outcomes, and published in a language other than English.

**Condition or domain being studied.** Review examining the evidence of what works in terms of direct and indirect policies on PA, and how.

**Participants/population.** Eligibility criteria relating to population characteristics are described in Table 3.

**Exposure(s), intervention(s).** Policies that aim to have a direct or indirect effect on PA behaviour of target groups and populations and on the PA environment that support the behaviour under consideration. “Direct” policy refers to policies where the primary aim is improving the PA environment and increasing PA participation. “Indirect” policy refers to policies where the primary aim is not to increase PA levels but this may occur as a co-benefit of successful implementation.

Grey literature/Other: Similar to the empirical studies, included as a result of a policy intervention.

**Main outcome(s).** All study designs (reviews, empirical evidence and grey literature/other) must include the following outcome(s); a change in PA (or proxy, e.g. fitness), assessed by means of self-report or wearable devices (e.g., accelerometer) and/or; a change in features of the physical and social environment (e.g., facilities, equipment, action plans, programmes) hypothesised to lead to changes in PA outcomes as a result of a policy intervention.

**Study selection and data extraction**

**Download of title and abstract records.** Titles and abstracts identified by the search will be downloaded as “Endnote import” (extension.enw) files. They will be uploaded to Endnote X9, a citation management software, and Rayyan (Ouzzani et al., 2016), online software dedicated to managing reviews. A freely available alternative is Mendeley reference manager. Once records have been uploaded to Rayyan, the software will identify duplicate articles and one of the two identical articles will be removed. The remaining articles will undergo the first round of screening by two researchers (KV and LK) in a shared Rayyan, account.

**Title and abstract review.** Title and abstract review will be performed by at least one reviewer (KV) and checked by another.
reviewer (LK). Checking will involve reviewing title and abstracts decisions to establish whether the second reviewer concurs with the screening decision. Discrepancies will be resolved by discussion to reach consensus, in consultation with a third researcher (EG) when necessary. The screening process will involve comparing the information presented in the title and abstract to the eligibility criteria. Titles and abstracts that appear to conform to the eligibility criteria will be deemed eligible for full text review while those that do not will be discarded from the next stage of the data extraction process.

**Download of full articles.** Full text articles for all title and abstract recorded deemed eligible for full text review will downloaded using the resources provided by their Institution. If a full text record cannot be acquired using these resources, researchers will investigate whether they can be located through use of other libraries to which the research team has access. If a full text article cannot be located through any of these library resources, the authors will be contacted through whichever channels can be identified from the information in the title and abstract.

**Full text review.** Full text review will be performed by at least one reviewer (KV) and checked by another reviewer (LK) using a template. Discrepancies will be resolved by discussion to reach consensus, in consultation with a third researcher (EG) when necessary. The following information will be extracted: first author, year of publication, country, study design, data collection method, sample, recruitment/setting, sample size, and response rate.

**Risk of bias (quality) assessment**

Risk of bias will be assessed by at least one reviewer (EG) and checked by another reviewer (BC). Discrepancies will be resolved by discussion to reach consensus, in consultation with a third researcher when necessary. The results of the quality assessment will be narratively incorporated into the synthesis process. A descriptive summary using the criteria described below will be presented at study level and discussed in the review. Furthermore, the methodological quality will be narratively summarized at review level.

The quality of the included quantitative studies, inclusive of randomised, non-randomised and observational studies (encompassing both longitudinal and cross-sectional studies) will be assessed by means of an adapted ‘Downs and Black’ checklist tool (Downs & Black, 1998). This tool is apt to assess common biases in a range of study types as noted. The checklist will be modified to meet the aims of this review with some items deemed non applicable and subsequently removed.

The AMSTAR measurement tool will be used for the assessment of systematic reviews and comprehensive reviews with a rigorous search strategy, including reviews of reviews. This tool consists of 11 items and has good face and content validity for measuring the methodological quality of systematic reviews (Shea et al., 2007). Not all items are applicable to every type of review being assessed and quality ratings will take account of this circumstance. Similar to Messing et al. (2019), to assess the quality of included studies, we will calculate percentage values for each study. These values will represent the percentages of criteria met per study, based on the criteria specific to each of the assessment tools we use that are applicable in light of the type of study design considered.

**Strategy for data synthesis**

A narrative synthesis will be used to interpret and analyse the data. Outcome data extracted will be tabulated to determine the impact on policy areas and policy actions (policy indicators). Evidence on the effectiveness of policy actions will be described using a method described by Panter and colleagues (2019); policy actions will be coded as “significantly positive” (+), “significantly negative” (-), “no significance test” (?) or “inconclusive” (0). Lastly, for the included mixed methods studies,
reviews and policy analysis documents, the main findings stated in the discussion and conclusions section will be extracted. Main findings of the articles will be copied into a single table along with a reference to the article itself, and details of the overall risk of bias of the study from which the information is extracted. The synthesised data will be presented in a six-column table with the different columns presenting information on the reference, study description, study type, main findings or outcomes, risk of bias and category of evidence, respectively.

Dissemination
Study findings will be presented at professional networking events such as the World Congress on Public Health. The PEN consortium intends to hold a five-day symposium at the biennial conference of the International Society of Physical Activity and Health inviting experts and policymakers to discuss and providing them with the evidence of from the systematic literature reviews.

Study status
The submission of the first of seven intended reviews is being finalised. A further three reviews are being formulated.

Conclusion
The aim of this project is to assist policymakers to achieve the GAPPa target of a 15% relative reduction in the prevalence of insufficient PA (WHO, 2019). The proposed review will attempt to determine what is the impact of policy interventions on PA outcomes in the domain of whole-of-school programmes. By providing this evidence, this review will support the development of the PA EPI. The PA EPI in turn will support policy makers by facilitating the development and benchmarking of policies which will work towards achieving this target. Achieving this target will provide health benefits such as reduced premature mortality as well as substantial co-benefits such as contributing to a sustainable environment and quality education (WHO, 2019).

Data availability
Underlying data
No underlying data are associated with this article.

Reporting guidelines

Data are available under the terms of the Creative Commons Zero “No rights reserved” data waiver (CC0 1.0 Public domain dedication).

Acknowledgements
The authors thank Dr Sarah Taylor and Michael Lawlor for their assistance in developing the search terms used in the study.

References

Panter J, Guell C, Humphreys D, et al.: Can changing the physical...


Thank you for the invitation to review this paper. This protocol paper outlines the details of the first of several planned systematic literature reviews to assess the impact of policy on physical activity outcomes based on the “seven best investments“ document (ISPAH, 2012). This is a very ambitious review series and I commend the authors on undertaking this important work. While the methods to be employed is generally clear, there are several areas that require more detailed explanation/rationale. I believe that doing so will strengthen this paper and ultimately, the review.

Introduction
It would be helpful to include specific aims of the systematic review.

P 4- Eligibility criteria
The authors need to clearly define what they mean by a ‘whole of school approach’ what are the critical components or does a policy just need to be labelled as such? For examples, does a ‘whole of school’ approach need to address the elements identified in ISPAH 2012: i.e. “prioritizing: regular, highly-active, physical education classes; providing suitable physical environments and resources to support structured and unstructured physical activity throughout the day (e.g., play and recreation before, during and after school); supporting walk/cycle-to-school programs and enabling all of these actions through supportive school policy and engaging staff, students, parents and the wider community”.

I was confused by the distinction between the “mandatory” eligibility criteria and specific “preferred” eligibility criteria and the rationale for this. I have not seen this before in a review. Can authors more clearly explain the application of these and the rationale? Perhaps by way of examples?

p. 5 Main Outcomes
The Introduction section implies that the focus of the review will be on the effectiveness of policies

© 2020 Craike M. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
in increasing PA. This seems to be inconsistent with the Main Outcome(s), identified on p. 8, where studies will be eligible if they include “a change in features of the physical and social environment (e.g., facilities, equipment, action plans, programmes) hypothesised to lead to changes in PA outcomes as a result of a policy intervention”. I suggest that the Introduction include a section that alludes to this, and why such outcomes could be considered important (e.g., that focusing on mediators can be used as a proxy for actional PA participation outcomes).

p. 6 Full text review
Given the importance of implementation to the success of policies, I wonder if it would be worthwhile for the authors to attempt to gather data to indicate whether or not the policy was implemented? Or whether, for example, action plans were developed, and resources allocated to implementation. Given the importance of implementation to policy success, I would see this type of information as critical when evaluating the success of such policies.

Also, the inclusion of the validity and reliability of the PA measures should be included in the data extraction.

Each of the policy types is very broad, which will make it difficult to identify policy action that is needed to increase PA. For example, a ‘whole-of-school approach’ can includes several different actions. Have the authors considered providing more detail or categorising the type of whole of school approach to facilitate a clearer and more meaningful assessment about the characteristics of effective whole of school policies?

p. 6, Strategy for data synthesis
I am not clear on what is meant by: “Outcome data extracted will be tabulated to determine the impact on (my emphasis) policy areas and policy actions (policy indicators)”. What is meant here by ‘policy indicators’ for example?

I suggest that the method described by Panter and colleagues (2019) be described in the paper (rather than just referring the reader to the paper). It is very important that the reader understands the criteria that is being used to assess the evidence of effectiveness.

I am unclear about how the findings from systematic reviews could be combined with primary research studies. This needs to be further explained.

Will there be any attempt to distinguish effectiveness based on the following: school level (primary or secondary school level), gender, level of socio-economic disadvantage?

Is the rationale for, and objectives of, the study clearly described?
Partly

Is the study design appropriate for the research question?
Partly

Are sufficient details of the methods provided to allow replication by others?
Partly

Are the datasets clearly presented in a useable and accessible format?
Not applicable

**Competing Interests:** I have previously worked with one of the co-authors of this paper. This has not affected my ability to review impartially.

**Reviewer Expertise:** Physical activity behaviour change; systematic reviews.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

---

**Author Response 18 Nov 2020**

**Kevin Volf, University of Limerick, Ireland**

Dear Dr Craike,

On behalf of the review team, I would like to thank you for taking the time for your thorough review of our paper. The team would especially like to thank you for the timely manner in which your review was conducted. We have received your feedback and have made changes to the manuscript accordingly, and believe this has strengthened the quality of the paper.

Please find below a point-by-point response to the issues raised in your review, with changes from the previous version highlighted in red, for your consideration.

Yours sincerely,
Kevin Volf, B.Sc., MM.Sc. PhD Candidate

**Comment 1:**

**Introduction**

It would be helpful to include specific aims of the systematic review.

We agree with the comment and the aims have added to the final paragraph of the Introduction section. To clarify the focus of this paper is to provide information on the protocol being used for seven reviews, this was unclear in the original manuscript. The text has been edited to make this point clearer throughout. Please see amendments below, text in bold is ‘new’ and has been added for clarity:

[Introduction page 6, paragraph 2]

At its third biennial congress ISPAH promulgated the Toronto Charter calling for political commitment to achieving greater opportunities for PA (Bull et al., 2010). To guide action on this issue the Charter was subsequently accompanied by a document titled Non-Communicable Disease Prevention: Investments that Work for Physical Activity (2012). This document declared seven domains which evidence suggested could be effectively targeted to increase PA opportunities. **These were whole-of-school programmes, transport policies and systems, urban design regulations and infrastructure, primary health**
care, public education, community-wide programmes and sport systems and programmes that promote ‘sport for all’. These seven domains provide a policy setting structure for systematic literature review search.

[Introduction page 7 paragraph 1]
This protocol paper outlines the methodology for seven complementary systematic literature reviews as part of PEN. Each review is designed to determine the impact of policy, either directly or indirectly, on physical activity outcomes across the different policy domains identified in the “Seven Best Investments” (ISPAH, 2011). These policy domains are schools, transport, urban design, healthcare, public education, the community and sport. These reviews will provide evidence supporting the development of a tool named the Physical Activity Environment Policy Index (PA EPI), based on similar principles to the Food Environment Policy Index (Food EPI) (Swinburn et al., 2013). The PA EPI will provide policy makers with a list of policy statements and corresponding actions which the evidence has determined improve PA outcomes across domains. The aim of each PEN review is to evaluate the status of the evidence base for the impact of policy on PA outcomes across the different policy domains identified in the “seven best investments”.

Comment 2:
P 4- Eligibility criteria
The authors need to clearly define what they mean by a ‘whole of school approach’ what are the critical components or does a policy just need to be labelled as such? For examples, does a ‘whole of school’ approach need to address the elements identified in ISPAH 2012: i.e. “prioritizing: regular, highly-active, physical education classes; providing suitable physical environments and resources to support structured and unstructured physical activity throughout the day (e.g., play and recreation before, during and after school); supporting walk/cycle-to-school programs and enabling all of these actions through supportive school policy and engaging staff, students, parents and the wider community”.

The reference to ‘whole of school approach’ is directly related to the ‘Seven Investments’ paper. The use of ‘Whole of school’ within the search strategy as part of the schools domain, is not for the purpose of evaluating the impact of ‘whole of school programmes’, rather to review policies pertinent to the school setting, to determine if any evidence of policy impact exists. The text has been modified as follows, text in bold is ‘new’ and has been added for clarity;

[Introduction page 7 paragraph 1]
This protocol paper outlines the methodology for seven complementary systematic literature reviews as part of PEN. Each review is designed to determine the impact of policy, either directly or indirectly, on physical activity outcomes across the different policy domains identified in the “Seven Best Investments” (ISPAH, 2011). These policy domains are whole-of-school programmes, transport policies and systems, urban design regulations and infrastructure, primary health care, public education, community-wide programmes involving multiple settings and sport systems and programmes.
that promote ‘sport for all’

It is proposed that individual systematic literature reviews will be performed on each of the seven best investment domains, with this initial review focusing on schools and subsequent reviews focusing on ‘transport’, ‘public education’ and ‘sport’ domains.”

Comment 3:
I was confused by the distinction between the “mandatory” eligibility criteria and “preferred” eligibility criteria and the rationale for this. I have not seen this before in a review. Can authors more clearly explain the application of these and the rationale? Perhaps by way of examples?

In essence, the “mandatory” and “preferred” nomenclature refer to “elements that are common to all reviews” and “elements that only relate to a particular review” respectively. While this was explained in the search details, to make this clearer we have edited the mandatory and preferred labels to “general” (applies to all mooted reviews) and “specific” (only applies to a particular review) as follows, text in bold is ‘new’ and has been added for clarity:

[Methods, p7 paragraph 5]:

“The search of electronic databases will comprise seven individual searches (corresponding to the seven best investments), each one to be run on each of the databases. The seven searches will be formed by combining the same basic search strategy (i.e., general eligibility criteria) with seven distinct search concepts (i.e., specific eligibility criteria for each domain). The basic search strategy will consist of three search concepts (Table 1): search concept one (C1), which will combine synonyms for the keyword “policy” with the Boolean Operator “OR”; search concept two (C2), which will do the same with the keyword “physical activity”; and search concept three (C3), which will do the same for the keyword “impact”. The three search terms will be combined with the Boolean operator “AND” (Table 1).

Each of the seven searches will further be combined with a specific search concept constructed to reflect only one of the seven best investments declared in the document Non-Communicable Disease Prevention: Investments that Work for Physical Activity (ISPAH, 2011) (Table 2). It is proposed that individual systematic literature reviews will be performed for each of the seven best investment domains, with an initial review focusing on schools and subsequent reviews focusing on ‘transport’, ‘public education’ and ‘sport’ domains.

Comment 4:
p. 5 Main Outcomes
The Introduction section implies that the focus of the review will be on the effectiveness of policies in increasing PA. This seems to be inconsistent with the Main Outcome(s), identified on p. 8, where studies will be eligible if they include “a change in features of the physical and social environment (e.g., facilities, equipment, action plans, programmes) hypothesised to lead to changes in PA outcomes as a result of a policy intervention”. I suggest that the
Introduction include a section that alludes to this, and why such outcomes could be considered important (e.g., that focusing on mediators can be used as a proxy for actional PA participation outcomes).

We agree with the inconsistency and the text has been reworded in the following location, text in bold is ‘new’ and has been added for clarity:

[Introduction page 7 paragraph 1]

This protocol paper outlines the methodology for seven complementary systematic literature reviews as part of PEN. Each review is designed to determine the impact of policy, either directly or indirectly, on physical activity outcomes across the different policy domains identified in the “Seven Best Investments” (ISPAH, 2011).

Comment 5a:

p. 6 Full text review
Given the importance of implementation to the success of policies, I wonder if it would be worthwhile for the authors to attempt to gather data to indicate whether or not the policy was implemented? Or whether, for example, action plans were developed, and resources allocated to implementation. Given the importance of implementation to policy success, I would see this type of information as critical when evaluating the success of such policies. […]

While we agree that level of implementation of policies is important, and believe that ‘impact on PA’ will be reflected in the degree of implementation presented, it is beyond the scope of this study to carry out a detailed examination of type of implementation beyond what is reported as evidence of impact. This very important topic may indeed by the subject of a follow up review.

Comment 5b:

[...] Also, the inclusion of the validity and reliability of the PA measures should be included in the data extraction.

Detailed information on any outcome measures used to assess PA will be provided in the data extraction tables and presented in the corresponding reviews.

Comment 6:

Each of the policy types is very broad, which will make it difficult to identify policy action that is needed to increase PA. For example, a ‘whole-of-school approach’ can includes several different actions. Have the authors considered providing more detail or categorising the type of whole of school approach to facilitate a clearer and more meaningful assessment about the characteristics of effective whole of school policies?

This links back to Comment 2, the focus of the systematic review is not to evaluate ‘whole of school approach’ but rather the available evidence on impact of direct or indirect policies within specific settings, e.g., the school setting, on PA outcomes. Therefore, the reviewer is
quite correct, we envisage that policy action areas will be broad, e.g., physical education or extra-curricular sport, but the specific policy actions, e.g., require a minimum PE duration or require a PE curriculum may be more defined. Until the reviews are carried out, the level of evidence is unknown.

Comment 7:
p. 6. Strategy for data synthesis
I am not clear on what is meant by: “Outcome data extracted will be tabulated to determine the impact on (my emphasis) policy areas and policy actions (policy indicators)”. What is meant here by ‘policy indicators’ for example?

We agree with the reviewer that this is not clear. In order to improve clarity this sentence has been removed and the third paragraph on page 12 has been restructured as follows, text in bold is ‘new’ and has been added for clarity:

[Strategy for data synthesis, page 12 paragraph 4]

“A narrative synthesis will be used to interpret and analyse the data. The results of the data synthesis will be presented in a table. Within this table, a list of short descriptive statements will be compiled based on the policy actions identified in the scientific literature. Evidence on the effectiveness of these policy actions will then be described using a method used by Panter and colleagues (2019) […]”

Comment 8:
I suggest that the method described by Panter and colleagues (2019) be described in the paper (rather than just referring the reader to the paper). It is very important that the reader understands the criteria that is being used to assess the evidence of effectiveness.

Detail has been added to the second half of the first paragraph on page 12 to describe the coding system inspired by Panter et al. It now reads as follows, text in bold is ‘new’ and has been added for clarity:

[Strategy for data synthesis, page 12 paragraph 4]

[...] Specifically, when a study presents quantitative evidence about the effectiveness of one policy action a symbol will be assigned next to that particular policy action in the table. There are four categories of symbols reflecting the four possible outcomes: “significantly positive evidence” (+), “significantly negative evidence” (-), “no significance test” (?) or “inconclusive” (0).

Comment 9:
I am unclear about how the findings from systematic reviews could be combined with primary research studies. This needs to be further explained.

This may indeed be challenging, but until we examine the main results and conclusions of any systematic reviews that have reference to policy impact in the targeted domains, we
cannot be sure of whether their evidence is strong enough and clearly defined in order to 
assess impact. While we agree it may be difficult, in the search strategy we will include 
these documents in the initial search and depending on what we find include or exclude in 
the final data synthesis.

Comment 10: 
Will there be any attempt to distinguish effectiveness based on the following: school level 
(primary or secondary school level), gender, level of socio-economic disadvantage?

Until data extraction is complete, it is difficult to identify any specific stratification variables 
that would be pertinent to the presentation of the results. However, the data extraction 
tables will be devised to take into account any relevant stratification variables for each 
review as pertinent to the topic and domain.

As the reviewer's comment applies to the systematic review corresponding to the school 
setting only evidence of effectiveness will be presented in a manner that makes it clear 
which policy actions were shown to be effective at different school levels, across gender and 
SES where appropriate. The following sentence has been added to the strategy for data 
synthesis section to clarify this, text in bold is 'new' and has been added for clarity.

[Strategy for data synthesis, page 12 paragraph 3]

“In addition, data extraction tables will be designed to distinguish any demographic, 
environmental or other variables pertinent to synthesising the data. For example, in 
the schools review, data extraction columns will be included to reflect evidence of 
effectiveness stratified by gender, school level (primary, secondary, combined) or 
socio economic status where appropriate.

Competing Interests: No competing interests were disclosed.