Journal Pre-proof

A narrative review of global and national physical activity and sedentary behaviour guidelines development processes - The GUidelines Standards (GUS) project

Karen Milton, Coral L. Hanson, Alice Pearsons, Roger Chou, Emmanuel Stamatakis

PII: S0091-7435(24)00114-2

DOI: https://doi.org/10.1016/j.ypmed.2024.107959

Reference: YPMED 107959

To appear in: *Preventive Medicine*

Received date: 29 February 2024

Revised date: 4 April 2024

Accepted date: 15 April 2024

Please cite this article as: K. Milton, C.L. Hanson, A. Pearsons, et al., A narrative review of global and national physical activity and sedentary behaviour guidelines development processes - The GUidelines Standards (GUS) project, *Preventive Medicine* (2023), https://doi.org/10.1016/j.ypmed.2024.107959

A narrative review of global and national physical activity and sedentary behaviour guidelines development processes - The GUidelines Standards (GUS) project

Karen Milton[1]*, Coral L Hanson[2], Alice Pearsons[2], Roger Chou[3], Emmanuel Stamatakis[4,5]

[1] Norwich Medical School, University of East Anglia, UK

[2] School of Health and Social Care, Edinburgh Napier University, UK

[3] Departments of Medicine, and Medical Informatics and Clinical Epidemiology, Oregon Health and Science University, USA

[4] Mackenzie Wearables Research Hub, Charles Perkins Centre, University of Sydney, Australia

[5] School of Health Sciences, Faculty of Medicine and Health, University of Sydney, Australia

* Corresponding author's details:

Dr Karen Milton
Norwich Medical School
University of East Anglia
Norwich
NR4 7TJ
United Kingdom
Email: k.milton@uea.ac.uk
Tel: +44 1603 593311

Abstract

Background: Clinical and public health guidelines serve to direct clinical practice and policy, based on the best available evidence. The World Health Organization (WHO) and national health bodies of many countries have released physical activity and sedentary behaviour guidelines. Despite significant overlap in the body of evidence reviewed, the guidelines differ across jurisdictions. This study aimed to review the processes used to develop global and national physical activity and sedentary behaviour guidelines and examine the extent to which they conform with a recommended methodological standard for the development of guidelines.

Methods: We extracted data on nine sets of guidelines from seven jurisdictions (WHO, Australia, Canada, Japan, the Netherlands, United Kingdom, and United States). We rated each set of guidelines as high, medium, or low quality on criteria related to the rigour of the development process.

Results: We observed variation in the quality of guidelines development processes across jurisdictions and across different criteria. Guidelines received the strongest overall ratings for criteria on clearly describing the evidence selected and stating an explicit link between the recommendations and the supporting evidence. Guidelines received the weakest overall ratings for criteria related to clearly describing the methods used to formulate the recommendations and reporting external review by experts prior to publication. Evaluated against the selected criteria, the strongest processes were undertaken by the WHO and Canada.

Conclusions: Reaching agreement on acceptable guideline development processes, as well as the inclusion and appraisal procedures of different types of evidence, would help to strengthen and align future guidelines.

Keywords: public health, physical activity, sedentary behaviour, guidelines, development, review

Introduction

Clinical and public health guidelines serve to direct clinical practice and policy, based on the best available evidence. Importantly, the guidelines development process ideally provides opportunities for rigorous scientific consensus on the best available evidence through a systematic method. To aid the rigour and evaluation of such processes, a range of resources are available including the World Health Organization (WHO) handbook for guideline development, which details the step-by-step process to follow (World Health Organization, 2014), and the Appraisal of Guidelines for Research and Evaluation 2 (AGREE 2) tool, which can be used to assess the methodological quality of guidelines (Brouwers et al., 2010).

Clinical guidelines provide recommendations on the optimal treatment, prevention of reoccurrence or deterioration, management, and care for people with specific health conditions. Public health guidelines make recommendations on activities, policies and strategies that can help to prevent disease or improve health. Whereas the evidence that informs clinical guidelines often comes from randomised controlled trials (RCTs), the development of public health guidelines can be more complex because much of the evidence often comes from observational studies.

The World Health Organization (WHO) and national health bodies of many countries have released physical activity guidelines (with many also covering sedentary behaviour). Despite significant overlap in the body of evidence reviewed, the resultant sets of guidelines differ across jurisdictions. For adults, for example, the WHO and the United States guidelines recommend a *minimum range* of 150 – 300 minutes of moderate or 75 – 150 minutes of vigorous aerobic activity per week (Physical Activity Guidelines Advisory Committee, 2018; World Health Organization, 2020), whereas other sets of guidelines, including those from the United Kingdom and Canada, recommend a *minimum threshold* of 150 minutes of moderate or 75 minutes of vigorous intensity activity, without providing a recommended minimum range (UK Chief Medical Officers, 2019; Ross et al., 2020). As another

example of discrepancies, whilst most sets of guidelines have a general recommendation to minimise sedentary time without a specific quantitative threshold, the Canadian guidelines specify a maximum limit of eight hours per day, which includes no more than three hours of recreational screen time (Ross et al., 2020). Inconsistencies across existing physical activity and sedentary behaviour guidelines may be due to differences in 1) the way the guideline development groups and committees were constructed, interacted, and worked; 2) the types of evidence considered; 3) the way the evidence was synthesised and interpreted; and 4) the methods applied to appraise the strength of the evidence.

The GUidelines Standards (GUS) project aims to establish a Delphi-based consensus on the methodological standards applied to the development of future physical activity and sedentary behaviour guidelines. Whilst methodological standards for guideline development exist, they are generic across topics and often omit important elements such as the types of experts that should be involved and the types of evidence that should be reviewed. Furthermore, research on the relationship between physical activity and health is dominated by observational studies, with relatively few robust trials of interventions, posing challenges for the appraisal of evidence. As part of the GUS project, this study aimed to review the processes used to develop global and national physical activity and sedentary behaviour guidelines and examine the extent to which they conform with a recommended methodological standard for the development of guidelines based on the AGREE 2 tool (Brouwers et al., 2010).

Methods

We undertook a review of the development processes of the most recent (as of April 2022) national and global physical activity guidelines that involved reviews of the scientific evidence.

Inclusion and exclusion criteria

Guidelines were eligible for inclusion if they were the most recent national and global physical activity guidelines for children and youth and/or adults, developed through a formal review process, and available in any language. Guidelines were excluded if no formal review was conducted, they were based solely on ratification or endorsement of other existing guidelines, or we were unable to locate the necessary information on process or use of evidence following the pre-defined protocol.

Search strategy and sources (April – September 2022)

We were previously aware of the WHO global guidelines as three members of the authorship team (KM, RC, ES) were on the Guideline Development Group (GDG). We were also aware of a published review of the development of physical activity guidelines in the European region (Tcymbal et al., 2021). To identify the existence of other sets of guidelines, we used the Global Observatory for Physical Activity (Go-PA) Country Cards, published in 2021 (Global Observatory for Physical Activity, 2021). Every country that was identified on the Go-PA Country Cards as having physical activity guidelines was added to our search. We attempted to obtain all available information for each set of guidelines, including the guideline development processes, the evidence inclusion criteria, and how judgements were made about the strength of the evidence. This information may have been contained within the same document as the guidelines or published separately as a technical report or methodological manual. If the methodological detail was not included in the main documents, we considered any additional documents that detailed the methods, including peer-reviewed manuscripts. We followed a pre-defined protocol as follows:

- Google searches (conducted in English) were used to find the scientific reports on the methods and evidence used to inform each set of guidelines. We used the search terms 'physical activity guidelines' and country name.
- Where the relevant information could not be located via web-searches, the Go-PA Country Contacts were emailed and asked to provide the relevant scientific documents (this included

the Country Contacts for the countries identified from the European review by Tcymbal et al., 2021).

- If no reply was received from the Country Contacts within two weeks, a follow-up email was sent. If no reply was received within a further two weeks, countries were excluded.
- Where a country was confirmed to be potentially eligible for inclusion, but no published report was available or reports were available in languages other than English, the Country Contact was asked to complete a PROFORMA (Supplementary file 1) to determine eligibility. Countries that completed the PROFORMA were considered for inclusion; those that did not return the PROFORMA were excluded.

Assessment of quality of guideline development

AGREE 2 provides methodological standards for the development of guidelines and can be used to assess the quality of guidelines (Brouwers et al., 2010). The tool consists of 23 items organised within six domains (scope and purpose, stakeholder involvement, rigour of development, clarity of presentation, acceptability, and editorial independence). The purpose of this study was to examine the methodological rigour of the development of guidelines. As such, we rated each set of guidelines as high, medium, or low quality in relation to each item in domain 3 of AGREE 2 (rigour of development) using pre-defined scoring criteria (Supplementary file 2). We also extracted information related to domains 2 (stakeholder involvement) and 6 (editorial independence) for additional contextual information (Box 1). In addition, we were interested in the types of evidence considered and how the evidence was appraised. These factors are not included in AGREE 2, hence we developed a bespoke set of items to assess these characteristics. We developed a data extraction template *a priori*, including the selected items from AGREE 2 and the bespoke items developed to assess the use of evidence (Supplementary file 2).

Box 1. AGREE 2 domains and items extracted for review

Domain 2: Stakeholder involvement (contextual information only)

Item 4. The guideline development group includes individuals from all the relevant professional groups.

Item 5. The views and preferences of the target population (patients, public, etc.) have been sought.

Item 6. The target users of the guideline are clearly defined.

Domain 3: Rigour of development (scored via pre-defined criteria)

Item 7. Systematic methods were used to search for evidence.

Item 8. The criteria for selecting the evidence are clearly described.

Item 9. The strengths and limitations of the body of evidence are clearly described.

Item 10. The methods for formulating the recommendations are clearly described.

Item 11. The health benefits, side effects, and risks have been considered in formulating the recommendations.

Item 12. There is an explicit link between the recommendations and the supporting evidence.

Item 13. The guideline has been externally reviewed by experts prior to its publication.

Domain 6: Editorial independence (contextual information only)

Item 22. The views of the funding body have not influenced the content of the guideline.

Item 23. Competing interests of guideline development group members have been recorded and addressed.

Data extraction and analysis

The data extraction template was piloted by two members of the research team (KM and RC). This involved copying relevant sections of text from the available documents into the data extraction template. This process highlighted strong consistency between team members in the content and level of detail extracted from the pilot set of guidelines (WHO global guidelines, World Health Organization, 2020). The available data did not allow us to determine how many of each type of

study design were considered, so these items were subsequently removed from the template. Following this adaptation, CLH and AP independently extracted information on the WHO guidelines, given the pilot was undertaken by KM and RC, who were both on the GDG, to avoid potential bias in scoring. Two team members independently extracted data on all other sets of included guidelines (KM and either CLH or AP) except for one country, where information was not available in English; for this country we engaged a native colleague to work with KM on populating the template (see acknowledgements). Where guidelines for children and adults were developed using the same GDG and process, data were extracted once; however, where guidelines for different population groups involved distinct processes, data were extracted for each set of guidelines separately. The extracted data were compared, and the final dataset agreed through consensus.

Equality, diversity, and inclusion statement

The authorship team was gender balanced (three females, two males), included investigators across a wide range of career levels (from PhD student to full professor) and included geographical representation across the United States, the United Kingdom, and Australia.

Patient and public involvement

Patients and the public were not involved in the design or conduct of the study.

Ethical compliance

This study involved an analysis of publicly available documents; thus, ethical approval was not required.

Results

We identified one set of global guidelines and 62 countries with national guidelines on physical activity and/or sedentary behaviour. Whilst each country of the United Kingdom has a separate Go-

PA County Card, they have a combined set of guidelines, meaning that guidelines from 60 jurisdictions were potentially eligible for inclusion (Figure 1). We excluded guidelines from 53 jurisdictions because: they endorsed other guidelines (n=13); no formal review was conducted (n=14); Country Contacts confirmed there were no guidelines (n=2); there was no Country Contact that we could follow-up for information (n=1); or Country Contracts provided no or insufficient information (n=23). Guidelines from seven jurisdictions were included in the final review (Figure 1). In two countries (Australia and Canada) the guidelines for children and adults were developed through distinct processes, thus we extracted data on nine sets of guidelines (Table 1).

INSERT FIGURE 1

Figure 1. Summary of the guideline identification process and reasons for exclusion

Table 1. The nine sets of physical activity guidelines that were included in the review

Country	**Guidelines**
Global	The WHO Guidelines on Physical Activity and Sedentary Behaviour (World Health Organization, 2020)
Australia	Australian 24-Hour Movement Guidelines for Children (5-12 years) and Young People (13-17 years): An Integration of Physical Activity, Sedentary Behaviour, and Sleep (Department of Health and Aged Care, 2019)
	Development of Evidence-based Physical Activity Recommendations for Adults (18-64 years) (Brown et al., 2013)
Canada	Canadian 24-Hour Movement Guidelines for Children and Youth: An Integration of Physical Activity, Sedentary Behaviour, and Sleep (Tremblay et al., 2016)
	Canadian 24-Hour Movement Guidelines for Adults aged 18–64 years and Adults aged 65 years or older: an integration of physical activity, sedentary behaviour, and sleep (Ross et al., 2020)
Japan	The Japanese official physical activity guidelines for health promotion - "ActiveGuide" (Ministry of Health, Labour and Welfare, 2013)
The Netherlands	Physical Activity Guidelines (Health Council of the Netherlands, 2017)
United Kingdom	United Kingdom Chief Medical Officers' Physical Activity Guidelines (UK Chief Medical Officers, 2019)
United States	Physical Activity Guidelines for Americans, 2nd Edition (Physical Activity Guidelines Advisory Committee, 2018)

Contextual information about guideline development process

The contextual information extracted on each set of guidelines is summarised in Table 2. The size of the GDGs ranged from four to 72 members. The structure of these committees varied and was not always described. Four out of nine GDGs included methodologists, i.e. at least one expert trained or specialised in the appraisal of scientific evidence and guidelines development. The GDGs for the WHO and the Australian children guidelines included a Grading of Recommendations, Assessment, Development, and Evaluations (GRADE) methodologist (Guyatt et al., 2008). For both the Canadian children and adult guidelines, guideline development (AGREE 2) consultants were engaged. More than half of guidelines were developed without the involvement of a methodologist.

All guidelines, except for the Netherlands, described the target audience. The most commonly identified target audiences were policymakers and practitioners (such as health professionals); however, researchers and the public were also identified as target audiences for some sets of guidelines. Whilst no target audience was specified for the Netherlands guidelines report, there was a note stating that the Knowledge Centre for Sport Netherlands was working on translating the physical activity guidelines in a practical manner, which included a special focus on informing the professionals who will use the guidelines. Five guideline development processes included consultation with the target audience, for example patients and the public, although the target audience varied across jurisdictions and guidelines. This was usually undertaken via an online survey, although other methods such as focus groups were also used in some cases. In the United States process, the public were invited to attend the committee's first two meetings in person and oral and written comments were invited.

The development process was typically funded by the government or a public health agency. It was not stated how the guideline development process was funded in the Netherlands or the United Kingdom. For the Canadian adult guidelines, representatives of each of the four funding partners sat on the leadership group that oversaw the guideline development process. Six out of nine guideline reports stated that competing interests of the guideline development group members were recorded and addressed.

Table 2. Assessment of stakeholder involvement and editorial independence in relation to the nine sets of guidelines included in the review (domains 2 and 6 of AGREE 2)

	Domain 2					Domain 6	
	Number of experts on the GDG	Structure of the GDG described	Inclusion of a methodologist	Target users of the guidelines specified	Sought the views and preferences of the public	Funder specified	Competing interests assessed
WHO - children and adults	27	✓	✓	✓	✓	✓	✓
Australia - children	35	✓	✓	✓	✓	✓	✓
Australia - adults	4	-	-	✓	-	✓	-
Canada - children	27	✓	✓	✓	✓	✓	✓
Canada - adults	30	✓	✓	✓	✓	✓	✓
Japan - adults	11	-	-	✓	-	✓	-
Netherlands - children and adults	14	-	-	-	-	-	✓
United Kingdom - children and adults	72	✓	-	✓	-	-	✓
United States - children and adults	36*	✓	-	✓	✓	✓	-

* 17 core + 19 consultants; ✓ = evidenced within the guidelines documents; - = no evidence within the guidelines documents

Quality of guideline development process

The quality of each guideline development process according to each item of domain 3 of AGREE 2 is summarised in Table 3. The strongest overall ratings across the included guidelines related to the criteria for selecting the evidence being clearly described (Item 8) and stating an explicit link between the recommendations and the supporting evidence (Item 12), with seven guidelines rated as high and two rated as medium. The weakest overall ratings across the included guidelines related to the methods for formulating the recommendations being clearly described (Item 10), and the guidelines being externally reviewed by experts prior to their publication (Item 13), with only one set of guidelines rated as high, five medium and three low. In terms of the individual sets of guidelines, the strongest reported processes were undertaken by the WHO and Canada (both children and adults), with five high and two medium ratings for each set of guidelines, followed by the United States, with five high, one medium and one low rating. Less robust processes were reported in other jurisdictions, particularly the United Kingdom, with zero ratings of high, six medium and one low.

Table 3. Summary of the quality of the development process for the nine sets of guidelines included in the review (domain 3 of AGREE 2)

	Global - children and adults	Australia - children	Australia - adults	Canada - children	Canada - adults	Japan - adults	Netherlands - children and adults	United Kingdom - children and adults	United States - children and adults
7. Systematic methods were used to search for evidence	●●●	●●	●	●●●	●●●	●●●	●●●	●●	●●●
8. The criteria for selecting the evidence are clearly described	●●●	●●●	●●	●●●	●●●	●●●	●●●	●●	●●●
9. The strengths and limitations of the body of evidence are clearly described	●●●	●●●	●●	●●●	●●●	●	●●●	●●	●●●
10. The methods for formulating the recommendations are clearly described	●●	●●	●●●	●●	●●	●	●	●	●●
11. The health benefits, side effects, and risks have been considered in formulating the recommendations	●●●	●●●	●●	●●●	●●●	●●	●●●	●●	●●●
12. There is an explicit link between the recommendations and the supporting evidence	●●●	●●●	●●●	●●●	●●●	●●●	●●	●●	●●●
13. The guideline has been externally reviewed by experts prior to its publication	●●	●	●●●	●●	●●	●●	●	●●	●

●●● = high; ●● = medium; ● = low

Types of evidence included in the guideline development process

Information on the types and volume of evidence reviewed, as well as the use of any formal appraisal procedures, is summarised in Table 4. Seven out of nine sets of guidelines (all except Japan and the United Kingdom) used an evidence quality and/or certainty grading system. The WHO and the United States assessed both the quality and certainty of the evidence. The WHO process utilised the Assessment of Multiple Systematic Reviews instrument (AMSTAR 2) to rate the credibility of the systematic reviews under consideration (Shea et al., 2017), the Newcastle-Ottawa Scale to assess the quality of the studies (Wells et al., 2013), and GRADE to rate the certainty of the evidence (Guyatt et al., 2008). The United States process used AMSTAR 2 to assess the quality of reviews and risk of bias for original studies (Shea et al., 2017) and the 2018 Physical Activity Guidelines Advisory Committee Grading Criteria to rate the overall quality of the evidence for each health outcome (Physical Activity Guidelines Advisory Committee, 2018). The processes to develop the Australian children, and the Canadian children and adult guidelines utilised GRADE to assess the certainty of the evidence (Guyatt et al., 2008). The Australian adults and the Netherlands guidelines development processes utilised other grading systems.

Most guideline development committees considered both previous systematic reviews and original studies, although two considered original studies only (Canadian children and Japan). For the WHO, Australian children, Canadian adults, and the United States guidelines, it was clearly stated how many systematic reviews and original studies were considered. The Canadian children and Japan guidelines processes considered original studies only, and in both cases the number of studies considered was described. For the Netherlands process, it was clear how many systematic reviews were considered, but the number of original studies considered was unclear. For the Australian adults and United Kingdom guidelines development processes it was unclear how many systematic reviews and original studies were considered.

All guideline development processes considered evidence from RCTs except Japan, and all except Japan and the Netherlands considered evidence from other experimental designs such as non-randomised trials. All guideline development processes considered evidence from observational studies such as cohort studies, although not all clearly specified what types of observational studies were considered. No other study designs, such as modelling studies or qualitative studies, were considered for inclusion in any of the guideline development processes. In most cases (all except Australian adults, Japan, and the United Kingdom) it was clearly described how the rating of evidence was adapted for different study designs.

Table 4. Summary of the evidence inclusion and appraisal process for the nine sets of guidelines included in the review

	Use of an evidence quality grading system	Use of an evidence certainty grading system	What was considered – previous reviews, original studies, or both	Types of evidence considered					It is clear how the rating of evidence was adapted for different study designs	The number of reviews considered is clearly described	The number of studies considered is clearly described
				Randomised controlled trials	Other experimental designs (non-randomised trials)	Observational studies	Modelling studies	Qualitative studies			
WHO - children and adults	✓	✓	Both	✓	✓	✓	-	-	✓	✓	✓
Australia - children	-	✓	Both	✓	✓	✓	-	-	✓	✓	✓
Australia - adults	✓	-	Both	✓	✓	✓	-	-	-	-	-
Canada - children	-	✓	Original studies	✓	✓	✓	-	-	✓	N/A	✓
Canada - adults	-	✓	Both*	✓	✓	✓	-	-	✓	✓	✓
Japan - adults	-	-	Original studies	-	-	✓	-	-	-	N/A	✓
Netherlands - children and adults	✓	-	Both	✓	-	✓	-	-	✓	✓	-
United Kingdom - children and adults	-	-	Both	✓	✓	✓	-	-	-	-	-

United States - children and adults	✓	✓	Both	✓	✓	✓	-	-	✓	✓	✓

✓ = evidenced within the guidelines documents; - = no evidence within the guidelines documents; * Although the process to develop the Canadian adult guidelines included both previously synthesised evidence and original studies, original studies were considered for sleep and combined behaviours only, not physical activity and sedentary behaviour, which relied on previous reviews.

Discussion

This study examined the processes used to develop global and national guidelines on physical activity and sedentary behaviour. Across six sets of guidelines for children and seven sets of guidelines for adults, no two sets of recommendations were identical. Whilst some guidelines were relatively old (released over ten years ago), and thus differences may reflect the evolution of the scientific evidence over time, other sets of guidelines were produced around a similar time and were largely based on the same body of evidence. We used specific domains of the AGREE 2 tool to evaluate differences in the processes undertaken that may explain inconsistencies in the guidelines produced.

We observed variation in the quality of the guideline development processes across jurisdictions and across different elements of AGREE 2. The most robust guidelines development processes were undertaken by the WHO and Canada, followed closely by the United States. Each of these sets of guidelines were graded medium, rather than high, on Item 10 – methods for formulating the guidelines, for not explicitly detailing any areas of disagreement among the committee and how these were resolved. For item 13 – external review by experts, Canada and the WHO were graded medium; in both cases an external review was undertaken but no description was provided on how the guidelines were changed as a result. The United States was graded low on this item; whilst it was reported that each chapter was reviewed by at least two committee members who were not members of the drafting subcommittee, as well as federal staff liaisons, this was not considered to constitute an 'external review by experts'. Overall, Canada, the United States, and the global WHO guidelines were developed using robust methods, and aside from the lack of external review in the United States, were typically downgraded for omitting certain details in the reporting of the guideline development methods. Whilst it is encouraging that robust guideline development processes have been undertaken, this was not the case in all jurisdictions, which is likely a contributing factor to the inconsistent guidelines produced.

In addition to variation in the guideline development processes, differences were observed in the types of evidence considered. Some GDGs considered both systematic reviews and original studies, although some considered original studies only. Perhaps more controversial is the variation in the study designs that were considered eligible for inclusion. Notably, no guideline development process restricted study design eligibility to RCTs; the processes therefore differed to the typical processes to develop clinical guidelines, which often rely exclusively or primarily on RCTs. The only study design considered across all guidelines was observational studies, although the types of observational studies included varied across guidelines, with some including only cohort studies and others also including designs such as cross-sectional and case control studies. The inclusion of, and reliance on, observational studies creates additional challenges for guideline development due to increased potential for confounding and greater susceptibility to bias compared to well-conducted RCTs, which can make it difficult to determine causality, result in less certainty in findings, and lead to erroneous conclusions. For example, confounding by indication in observational studies, which occurs when the probability of studied outcomes is causally related to the exposure being studied, can result in reverse causality. In addition, there was apparent variation in the evidence threshold required to make recommendations; this is evident by the decision in Canada to recommend a maximum threshold for sedentary behaviour, while other jurisdictions considered the evidence insufficient to inform a specific threshold.

The production of inconsistent guidelines is problematic for several reasons. Firstly, it casts doubt over the true nature of the evidence linking physical activity to a broad range of health outcomes. Secondly, there is potential for mixed messaging, leading to confusion among the intended target audiences of the guidelines. Thirdly, public health guidelines are typically used as the threshold against which population prevalence of physical activity is assessed. Inconsistent guidelines therefore have the potential to lead to inconsistent approaches to establishing national prevalence estimates, posing challenges for cross-country comparisons.

Guideline development processes can be a resource intensive undertaking, often involving tens of scientists and administrative personnel over a period of several years. The processes undertaken by the WHO, Canada, and the United States represent significant investment into lengthy and robust scientific processes, involving leading experts and extensive consultation. As demonstrated across the subset of guidelines in this review, this level of rigour is not easily replicable. Adopting rigorous guidelines produced by others provides an efficient and cost-effective approach to the establishment of national guidelines, as well as ensuring consistency in the guidelines across jurisdictions, facilitating national surveillance, the establishment of global prevalence estimates, and cross-country comparisons (World Health Organization, 2020). Pooling of resources across jurisdictions for the development of future physical activity guidelines would minimise duplication of efforts and lead to physical activity and sedentary behaviour guidelines that are consistent and generalisable to most people worldwide.

Strengths of this work include the systematic approach to identifying guidelines, the pre-defined protocol for gathering information on the development processes, and the use of AGREE 2 to rate each guideline development process. However, some limitations should be acknowledged. Potentially eligible countries were excluded because we were unable to locate sufficient information on the guideline development process. A second limitation is that the authorship team relied on the information available in published documents; a set of guidelines may have been downgraded if details related to any element of AGREE 2 were omitted from the available documentation. A key challenge was locating information on guideline development processes. In most cases, multiple documents had to be sought and reviewed to appraise the guideline development process. In addition, it was not possible to determine what factors resulted in discrepancies across guidelines. Wide-spread use of AGREE 2 would improve consistency and transparency in the reporting of future guideline development processes, and making this information more readily accessible, and

available in a single document, would simplify similar review processes in the future. A final limitation is that several authors were members of the WHO GDG; however, we addressed this by ensuring other team members undertook data extraction for the WHO guidelines.

Conclusion

We observed variation in the quality of guideline development processes across jurisdictions and across different elements of AGREE 2. Reaching agreement on acceptable guideline development processes, as well as the inclusion and appraisal of different types of evidence, could help to strengthen and align future physical activity and sedentary behaviour guidelines. Wide-spread adoption of AGREE 2 standards would improve consistency and transparency in the reporting of future guideline development processes.

Competing interests

KM, RC, and ES were on the GDG for the WHO 2020 guidelines on physical activity and sedentary behaviour. KM was also on the GDG for the 2019 United Kingdom Chief Medical Officers guidelines.

Funding

ES is funded by an NHMRC Investigator Grant level 2 (APP1194510).

References

Brouwers M, Kho ME, Browman GP et al. AGREE II. Advancing guideline development, reporting and evaluation in healthcare. *Canadian Medical Association Journal* 2010; 182:E839-842; doi: 10.1503/cmaj.090449.

Brown WJ, Bauman AE, Bull FC et al. *Development of evidence-based physical activity recommendations for adults (18-64 years)*. Canberra: Commonwealth of Australia, 2013.

Department of Health and Aged Care. *Australian 24-Hour Movement Guidelines for Children (5-12 years) and Young People (13-17 years): An Integration of Physical Activity, Sedentary Behaviour, and Sleep.* Canberra: Commonwealth of Australia, 2019.

Global Observatory for Physical Activity. *2nd physical activity almanac*. Pelotas: Federal University of Pelotas, 2021.

Guyatt GH, Oxman AD, Vist GE et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ (Clinical research ed)* 2008;336(7650):924-6.

Health Council of the Netherlands. *Physical activity guidelines 2017.* The Hague: Health Council of the Netherlands, 2017.

Ministry of Health, Labour and Welfare. *ActiveGuide – Japanese official physical activity guidelines for health promotion*. Tokyo: Ministry of Health, Labour and Welfare, 2013.

Physical Activity Guidelines Advisory Committee. *2018 physical activity guidelines advisory committee scientific report*. Washington, DC: US Department of Health and Human Services, 2018.

Ross R, Chaput J-P, Giangregorio LM et al. Canadian 24-hour movement guidelines for adults aged 18–64 years and adults aged 65 years or older: an integration of physical activity, sedentary behaviour, and sleep. *Applied Physiology, Nutrition, and Metabolism* 2020; 45:S57–S102.

Shea BJ, Reeves BC, Wells G et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ* 2017;358:j4008.

Tcymbal A, Gelius P, Abu-Omar K et al. Development of national physical activity recommendations in 18 EU member states: a comparison of methodologies and the use of evidence. *BMJ Open* 2021; 11:e041710. doi:10.1136/bmjopen-2020-041710.

Tremblay MS, Carson V, Chaput J-P et al. Canadian 24-hour movement guidelines for children and youth: An integration of physical activity, sedentary behaviour, and sleep. *Applied Physiology, Nutrition, and Metabolism* 2016; 41:S311–S327.

UK Chief Medical Officers. *UK Chief Medical Officers' Physical Activity Guidelines.* London: Department of Health and Social Care, 2019.

Wells G, Shea B, O-Connell D et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. 2013. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.

World Health Organization. *WHO handbook for guideline development, 2nd edition.* Geneva: World Health Organization, 2014.

World Health Organization. *WHO guidelines on physical activity and sedentary behaviour*. Geneva: World Health Organization, 2020.

Ethical approval

Not applicable

Consent for publication

Not applicable

Availability of data and materials

Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.

Authors' contributions

KM and ES conceived the idea. RC provided significant guidance on the methods. KM, RC, and ES developed the study protocol and data extraction template. KM and RC piloted the data extraction template. KM, CLH and AP extracted data for all sets of guidelines and agreed the final dataset. KM led the draft manuscript. All authors contributed to revising the manuscript and approved the final version.

Acknowledgements

The authorship team would like to acknowledge Dr Andrea Ramirez Varela and Juliana Mejia Grueso, who assisted in collating relevant documents. We would also like to acknowledge Dr Masamitsu Kamada who assisted with data extraction for the Japanese guidelines.

Highlights

- The quality of guideline development processes varied across jurisdictions.
- The most robust processes were undertaken by the WHO and Canada.
- Agreement on acceptable processes would strengthen and align future guidelines.

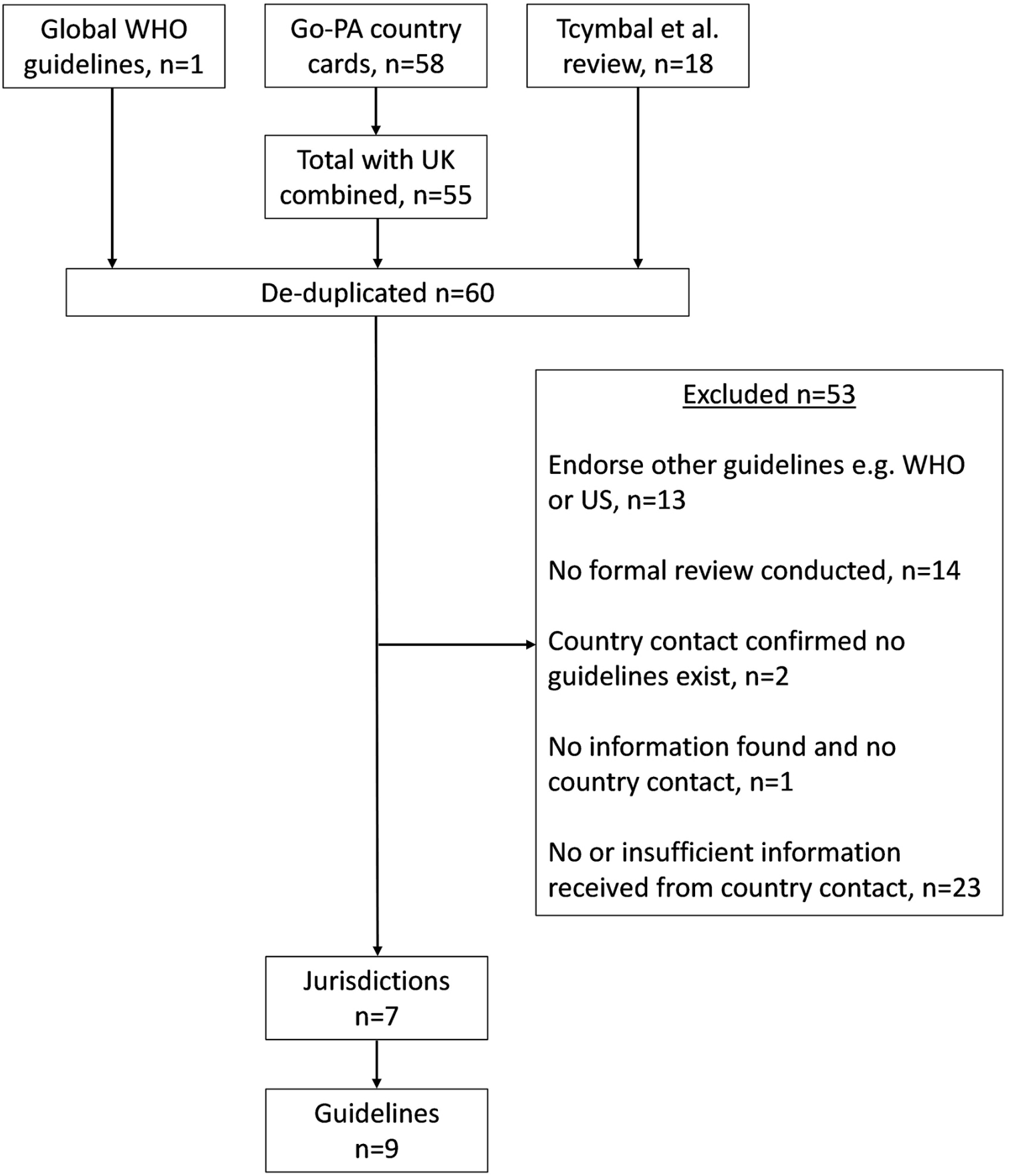
Global WHO guidelines, n=1
Go-PA country cards, n=58
Tcymbal et al. review, n=18
Total with UK combined, n=55
De-duplicated n=60
Excluded n=53
Endorse other guidelines e.g. WHO or US, n=13
No formal review conducted, n=14
Country contact confirmed no guidelines exist, n=2
No information found and no country contact, n=1
No or insufficient information received from country contact, n=23
Jurisdictions n=7
Guidelines n=9

Figure 1