



Evaluation of the Government indicator methodology in the Global Matrix on physical activity for children and adolescents

Iryna Demchenko^{a,b,c}, Salomé Aubert^c, Nicholas Kuzik^d, Lowri C. Edwards^e, Malcolm Ward^f, Kelly P. Arbour-Nicitopoulos^g, Javier Brazo-Sayavera^h, Andrea Cortinez-O'Ryan^{i,j}, Gabriela De Roia^k, Silvia A. González^l, Roselle M. Guisihan^m, Justin Y. Jeonⁿ, Yeon Soo Kim^o, Bee Koon Poh^p, Terézia Kovalik Slančová^q, José Francisco López-Gil^{r,s}, Kwok Ng^{t,u,v}, Yves Palad^w, Diego Augusto Santos Silva^x, Simone A. Tomaz^y, Piotr Kazimierz Urbański^z, Mark S. Tremblay^{a,b,c,aa,*}

^a Healthy Active Living and Obesity Research Group, Children's Hospital of Eastern Ontario Research Institute, Ottawa, ON, Canada

^b Department of Health Sciences, Carleton University, Ottawa, ON, Canada

^c Active Healthy Kids Global Alliance, Ottawa, ON, Canada

^d Faculty of Medicine and Dentistry, University of Alberta, Edmonton, AB, Canada

^e Cardiff School of Sport and Health Sciences, Cardiff Metropolitan University, Wales, UK

^f Applied Sports Technology, Exercise and Medicine Research Centre (A-STEM), Swansea University, Wales, UK

^g Faculty of Kinesiology and Physical Education, University of Toronto, Toronto, ON, Canada

^h Department of Sports and Computer Science, Universidad Pablo de Olavide, Seville, Spain

ⁱ Department of Physical Education, Sports and Recreation, Universidad de La Frontera, Temuco, Chile

^j School of Kinesiology, Faculty of Medical Sciences, Universidad de Santiago de Chile, Chile

^k Laboratorio de Estudios en Actividad Física (LEAF), Universidad de Flores, CABA, Argentina

^l School of Medicine, Universidad de los Andes, Bogotá, Colombia

^m Institute of Clinical Epidemiology, National Institutes of Health, University of the Philippines Manila, Manila City, Philippines

ⁿ Department of Sport Industry Studies, Center for Exercise Medicine and Salutogenesis, Yonsei University, Seoul, South Korea

^o Department of Physical Education, Institute of Sports Science, Seoul National University, Seoul, South Korea

^p Faculty of Health Sciences, Universiti Kebangsaan Malaysia, Kuala Lumpur, Malaysia

^q University of Prešov, Faculty of Sports, Prešov, Slovakia

^r School of Medicine, Universidad Espíritu Santo, Samborombón, Ecuador

^s Faculty of Health Sciences, Universidad Autónoma de Chile, Temuco, Chile

^t Physical Activity for Health Research Centre, Health Research Institute & Department of Physical Education and Sport Sciences, University of Limerick, Limerick, Ireland

^u Institute of Sports Science and Innovation, Lithuanian Sports University, Kaunas, Lithuania

^v Department of Teacher Education, Turku Institute of Advanced Studies, University of Turku, Turku, Finland

^w Department of Physical Therapy, College of Allied Medical Professions, University of the Philippines, Manila, Philippines

^x Department of Physical Education, Sports Center, Federal University of Santa Catarina, Florianópolis, SC, Brazil

^y Faculty of Health Sciences and Sport, University of Stirling, Stirling, Scotland, UK

^z Department of Adapted Physical Activity, Poznań University of Physical Education, Poznań, Poland

^{aa} Department of Pediatrics, University of Ottawa, Ottawa, ON, Canada

ARTICLE INFO

Keywords:
International benchmarking
Measurement
Policy
Youth

ABSTRACT

Background/objectives: The Government indicator evaluates national policies and strategies for child and adolescent physical activity, supporting country reporting and international comparisons within the Global Matrix initiative. To grade this indicator, countries/jurisdictions apply either the original method or, in recent editions, an alternative approach. The alternative approach was proposed by the Welsh team, who derived six criteria from the World Health Organization's Health-Enhancing Physical Activity Policy Audit Tool v2 and developed an accompanying scoring rubric. This study aimed to evaluate the Welsh approach in comparison with the original method to guide recommendations for future applications.

* Corresponding author. 401 Smyth Rd, Ottawa, ON, K1H 8L1, Canada.

E-mail address: mtremblay@cheo.on.ca (M.S. Tremblay).

<https://doi.org/10.1016/j.jesf.2026.200520>

Received 3 May 2026; Received in revised form 14 August 2026; Accepted 31 August 2026

Available online 5 September 2026

1728-869X/© 2026 The Society of Chinese Scholars on Exercise Physiology and Fitness. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Methods: The study employed a mixed-methods design, including: 1) inductive content analysis of feedback collected via survey from 17 researchers and policy professionals from various countries, and 2) descriptive and Wilcoxon signed-rank analysis of Government indicator grades from Global Matrices 3.0 and 4.0 in 6 countries that transitioned to the Welsh approach and 28 countries/jurisdictions that retained the original method.

Results: User feedback on the Welsh approach was predominantly positive, noting its structure, transparency, and comprehensiveness as strengths, and indicating greater utility and suitability than the original method. Concerns included limited consideration of policy implementation and feasibility aspects. No statistically significant group-level temporal grade differences were found for both methods.

Conclusion: The Welsh approach adds structure and transparency to the Government indicator assessment. Further refinements to the Welsh approach may improve comparability and policy relevance, while both the Welsh and original methods remain acceptable for interim use within the Global Matrix, with potential applicability beyond it.

1. Introduction

Physical activity (PA) in childhood and adolescence is essential for optimal growth and development.^{1,2} Yet more than two-thirds of children and adolescents worldwide are estimated to have engaged in insufficient PA in the past decades.^{3–7} In addition to limiting health benefits, insufficient PA contributes to broader impacts on health systems and society, prompting public health concerns and global calls for sustained collaborative action.^{8–10} The socio-ecological model of health outlines a multilevel structure of determinants, from intrapersonal to policy level,¹¹ all of which are important for effective PA promotion.^{1,12–14} Among PA surveillance efforts, the Global Matrix (GM) is the only global initiative to consistently capture influences on children and adolescents' PA at different levels of the socio-ecological model.¹⁵ Led by the Active Healthy Kids Global Alliance (AHKGA), the GM reports on the international state of PA among 5–17-year-olds through a series of indicators, representing PA behaviors (e.g., Overall PA, Active Play) along with key sources of influence on PA (e.g., School, Government).¹⁶ Collaborative efforts from a total of 85 countries and jurisdictions¹⁷ led to the successful completion of the GMs 1.0 to 4.0 (2014 – 2022),^{3–6} with the GM 5.0 scheduled for release in 2026.¹⁸ Jurisdictions that participated as separate entities (e.g., Hong Kong, Guernsey) are hereafter referred to as “countries” for simplicity. The detailed methodology of the GM has been previously published.^{5,16} In brief, within GMs, teams of PA experts synthesize the best available population-level evidence in their respective countries, evaluate it against established benchmarks, and grade country performance on each indicator in a harmonized way using an academic letter scale (A+ [excellent] to F [failing]). Grades are summarized in country-specific Report Cards (RCs). Subsequently, national findings are compiled at the global level, enabling cross-country comparisons and helping to inform evidence-based PA promotion efforts worldwide.^{19,20}

The Government indicator is included in the GM framework to represent a crucial source of influence on population PA. It captures government action to support PA among children and adolescents in each respective country by examining the policies, strategies, legislation, and regulations that are in place, as well as evidence of efforts to put these commitments into practice through resource allocation, program implementation, monitoring, and other relevant activities. In the first three GMs, this indicator (defined as “*any governmental body with authority to influence PA opportunities or participation of children and adolescents through policy, legislation, or regulation*”) was evaluated through expert discussion (hereafter referred to as the “original method”).^{3–5} However, partly due to reliance on consensus rather than quantitative metrics, the Government indicator was described by the GM leadership team as challenging to uniformly grade and compare.³ An independent scoping review of global PA policy research has also noted some challenges with its assessment in early GMs, such as a lack of agreed-upon criteria for policy evaluation and clarity of benchmarks.²¹ To improve the grading process, an expert team from Wales initiated an alternative method.²² During their participation in the GM 3.0, Wales established a

national expert, multi-agency, multidisciplinary working group comprising researchers, academics, policy advisors, and specialist practitioners from national organizations to develop the Welsh 2018 RC.²³ A dedicated subgroup was responsible for the Government indicator assessment and, during the analysis of data related to government action, identified six key criteria informed by the “prompts” provided by the World Health Organization's Health-Enhancing PA Policy Audit Tool version 2 (HEPA PAT v2).²⁴ Based on their collective expertise, the subgroup agreed on the relative importance of each criterion and assigned percentage weights, contributing to a total score (100%). The proposed scoring rubric was then presented to the wider expert group for further consideration, leading to agreement on the final rubric, published by Ward et al.²² (hereafter referred to as the “Welsh approach”). The Welsh approach was further discussed and approved by the AHKGA Board as an alternative method that countries could optionally use for their RCs from GM 4.0 onwards.²⁵ The transition to the Welsh approach at the international level occurred when 12 out of 57 countries participating in the GM 4.0 chose to implement it.⁶ Besides the GM 4.0, the Welsh approach was applied by eight countries participating in the 2022 Para RCs initiative – a collaborative 14-country project exploring PA among children and adolescents with disabilities (CAWD), led by the International Federation of Adapted PA and inspired by and modeled after the AHKGA's GM.²⁶

Both the original and Welsh methods involved identifying existing government policies, initiatives, and budget items related to PA in children and adolescents and therefore relied on the same underlying data and broadly similar criteria. The main differences were in how the criteria were defined and operationalized. In the original method, the criteria were broader and based on three benchmarks (Supplementary data, Table S1), whereas in the Welsh approach, they were more specifically defined and organized into six items (Supplementary data, Table S2). The procedure used to assess the data in the original method involved having the expert team review the available qualitative evidence, discuss the findings, and reach consensus on the grade (A+ to F) that best reflected government performance. In the Welsh approach, the criteria were applied more systematically, with each criterion assessed and weighted according to a predefined scoring rubric, and the total score converted into a final grade (A+ to F). Therefore, although both methods assessed the same underlying data, differences in how the criteria were defined, applied, and weighted may contribute to possible variation in the assessment results.

Following GM evaluations, it was recommended to integrate CAWD data into future GMs²⁷ and to evaluate the Government indicator assessment methods as part of potential adjustments to the GM methodological framework.²⁵ Understanding PA experts' perceptions of the Welsh approach and exploring differences in Government indicator grades between the two methods across the two latest GM editions is warranted to guide future national assessments and international comparisons. Insights on its feasibility (i.e., how manageable), suitability (i.e., how appropriate), and utility (i.e., how practical) would support methodology optimization and facilitate the generation of more robust

global evidence for research, policy, and advocacy to advance PA promotion. The assessment of PA policies receives increasing international attention beyond GM, reflecting growing interest in understanding and strengthening policy action.²⁸ The findings of this study will contribute to a better understanding of how methodological choices may shape the assessment of policy environments and highlight the importance of considering both the results and methodological characteristics of different approaches when assessing government actions. Therefore, this study aimed to evaluate the use of the Welsh approach for measuring the Government indicator in RCs and Para RCs and compare it with the original method to guide recommendations for future applications.

2. Methods

2.1. Study design

Designed as a mixed-methods study, this study involved: 1) analysis of insights from the Welsh approach users in the GM 4.0 and 2022 Para RC initiative, collected via a survey, and 2) secondary analysis of quantitative data on the Government indicator from the GM 3.0 and 4.0. Ethical approval was obtained from the Carleton University Research Ethics Board, Ottawa, Canada (Clearance Certificate #123052).

2.2. Survey participants and design

User feedback was collected in April 2025 through an online English-language survey using Google Forms (Google LLC, USA), targeting one representative from each RC and Para RC team that employed the Welsh approach in the GM 4.0 ($n = 12$)⁶ or in the Para RCs initiative ($n = 8$).²⁶ The survey consisted of 62 questions, covering the general information about participants; the Welsh approach feasibility, suitability, utility, challenges, and benefits; user satisfaction, and future recommendations. The survey questions were developed specifically for the purposes of this study by the authorship team (ID, SA, NK, and MST) based on the study objectives and the key aspects of the methodology that the study sought to assess; it was not intended as a formally validated instrument. The questions were subsequently reviewed by the authors (LCE, MW) to ensure that they appropriately cover these aspects. Survey data were de-identified prior to analysis. Further details on participant recruitment and survey content are provided in the Supplementary data, [Appendix A](#).

2.3. Government indicator assessment methods and data sources

The original method determines grades for the Government indicator through a qualitative assessment of evidence across three benchmarks, without assigning numerical scores: 1) government leadership and commitment to providing PA opportunities for all children and adolescents; 2) the allocation of funds and resources for implementing PA promotion strategies and initiatives for all children and adolescents; and 3) demonstrated progress through the key stages of public policy making (Supplementary data, [Table S1](#)). The Welsh approach²² determines grades based on six key criteria, each assigned a predefined weight corresponding to its maximum score, with the weights summing to 100% (100 points). Key criteria include: 1) Number and breadth of relevant policies (maximum score 10 points), 2) Identified supporting actions (maximum score 20 points), 3) Identified accountable organization(s) (maximum score 25 points), 4) Identifiable reporting structures (maximum score 15 points), 5) Identified funding (maximum score 20), and 6) Monitoring and evaluation plan (maximum score 10), detailed in Supplementary data, [Table S2](#). The Government indicator grades from the GM 3.0 and 4.0 were obtained from publicly available sources: the AHKGA website (www.activehealthykids.org) and peer-reviewed academic publications.^{5,6} Only RC grades from countries that participated in both GM 3.0 and 4.0 were extracted for analysis. Grades from the 2022 Para RCs²⁶ were not considered due to the absence

of earlier editions for comparisons.

2.4. Data analysis

Survey responses to closed-ended questions (e.g., multiple-choice, rating scale) were analyzed descriptively (counts, percentages) in Excel (Microsoft Corporation, USA). Group satisfaction and agreement levels were measured using a top-two-box score, comprising the percentage of respondents who selected the two highest response categories (i.e., very satisfied and satisfied, or strongly agree and agree).²⁹ Thresholds were pre-defined by the study team for pragmatic categorization: top-two-box scores of $\geq 80\%$ indicating high, 60–79% moderate, and $< 60\%$ low levels of group satisfaction or agreement. Qualitative data from open-ended questions were analyzed using inductive content analysis³⁰ within predefined survey sections by one reviewer (ID), verified by a second reviewer (SA), with discrepancies resolved by consensus or by involving a third reviewer (MST). Grades and grade changes for the Government indicator across GM 3.0 and 4.0 were summarized using descriptive statistics, calculated after converting categorical variables (letter grades, A+ through F) into corresponding numbers, ranging from 15 (A+) to 2 (F). The Wilcoxon signed-rank test was used to compare grades between GM 3.0 and 4.0, given the paired nature of the observations at the country level, with each included country contributing a grade for each edition. A non-parametric approach was considered appropriate, avoiding reliance on the assumption of normally distributed paired differences. Statistical analyses were performed using SAS 9.4 (SAS Institute Inc., USA). Statistical significance was set at $p < 0.05$. The small sample size was taken into account when interpreting the statistical findings, which were considered alongside the qualitative results.

3. Results

3.1. Survey findings

Seventeen invitees responded to the survey, representing 11 of 12 contacted RC teams (Argentina, Brazil, Chile, Colombia, Malaysia, Philippines, Scotland, Slovakia, South Korea, Uruguay, Wales) and 7 of 8 contacted Para RC teams (Brazil, Canada, Ireland, Philippines, Poland, South Korea, Spain), indicating a 90% team response rate. Each team was represented by one distinct expert, except in one country, where teams from both initiatives were represented by the same individual. A summary profile of survey respondents by professional characteristics is available in Supplementary data, [Table S3](#).

According to the survey responses ($n = 17$), the average size of working groups involved in the assessment of the Government indicator was six within RC (range: 1 to 13) and four within Para RC teams (range: 2 to 13). The research, government, and practice sectors were represented in 93%, 64%, and 57% of the RC working groups, and in 89%, 56%, and 56% of the Para RC working groups, respectively. Among respondents, 59% ($n = 10$) reported a lack of sector representation or expertise, such as government ($n = 5$), community-level policy implementers ($n = 2$), practice sector ($n = 2$), and/or expertise in gathering policy-related data ($n = 1$), which they believed could have facilitated or improved the grading process. The assessment using the Welsh approach required less than one month for 41% of teams ($n = 7$), one to three months for 35% ($n = 6$), over three months for 12% ($n = 2$), and was not reported by 12% ($n = 2$). Among respondents with prior experience using the original method in previous GMs, 70% ($n = 7$) indicated that the Welsh approach was more time-consuming and required either the same amount of human resources (50%, $n = 5$) or more (30%, $n = 3$) than the original method. Group satisfaction with the use of the Welsh approach was 87% ($n = 13/15$) for RCs and 70% ($n = 7/10$) for Para RCs. Respondents' opinions on the feasibility, suitability, and utility of the Welsh approach are presented in [Table 1](#). General satisfaction with the method and its components is summarized in [Table 2](#). A comparison

Table 1

Agreement of survey respondents ($n = 17$) with statements related to perceived feasibility, suitability, and utility of the Welsh approach and its components (key criteria, scoring rubric) for the Government indicator assessment.

Survey items (statements)	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree	Group agreement level
1. General appreciation of the Welsh approach						
easy to understand	6%	71%	18%	6%	0%	77% ($n=13$) Moderate
easy to use	12%	71%	18%	0%	0%	82% ($n=14$) High
feasible (i.e., manageable given the time, resources, circumstances)	18%	47%	18%	18%	0%	65% ($n=11$) Moderate
suitable (i.e., appropriate for Report Cards and GMs)	29%	35%	24%	12%	0%	65% ($n=11$) Moderate
suitable for CAWD subpopulation within Report Cards and GMs	12%	29%	47%	12%	0%	41% ($n=7$) Low
valid (i.e., accurately measures the Government indicator)	24%	47%	18%	12%	0%	71% ($n=12$) Moderate
reliable (i.e., able to produce consistent and repeatable results)	18%	53%	24%	0%	6%	71% ($n=12$) Moderate
useful/beneficial for Report Cards and GMs	35%	35%	18%	12%	0%	71% ($n=12$) Moderate
coherent and comprehensive	24%	53%	12%	6%	6%	77% ($n=13$) Moderate
2. Key criteria						
easy to understand	12%	65%	6%	18%	0%	77% ($n=13$) Moderate
easy to use	6%	65%	12%	18%	0%	71% ($n=12$) Moderate
suitable (i.e., appropriate for Report Cards and GMs)	24%	41%	24%	6%	6%	65% ($n=11$) Moderate
suitable for CAWD subpopulation within Report Cards and GMs	12%	29%	47%	6%	6%	41% ($n=7$) Low
3. Scoring rubric						
easy to understand	12%	77%	6%	6%	0%	88% ($n=15$) High
easy to use	29%	53%	12%	6%	0%	82% ($n=14$) High
suitable (i.e., appropriate for Report Cards and GMs)	35%	29%	29%	6%	0%	65% ($n=11$) Moderate
suitable for CAWD subpopulation within Report Cards and GMs	12%	29%	47%	12%	0%	41% ($n=7$) Low

Notes: High group agreement level ($\geq 80\%$), moderate (60–79%), low ($< 60\%$). Totals may not sum to exactly 100% due to rounding.

Abbreviations: CAWD = children and adolescents with disabilities, GM = Global Matrix.

Table 2

Satisfaction of survey respondents ($n = 17$) with the use of the Welsh approach and its components (key criteria, scoring rubric) for the Government indicator assessment.

Survey items (questions)	Very satisfied	Satisfied	Neutral	Dissatisfied	Very dissatisfied	Group satisfaction level
1. How satisfied are you with the Welsh approach (overall) for the Government indicator assessment in your country's...?						
Global Matrix 4.0 Report Card (if applicable) ^a	13%	73%	7%	7%	0%	87% ($n=13$) High
2022 Para Report Card (if applicable) ^b	10%	60%	20%	10%	0%	70% ($n=7$) Moderate
2. How satisfied are you with the inclusion of each of the following key criteria into the Welsh approach?						
number and breadth of relevant policies	41%	29%	12%	6%	12%	71% ($n=12$) Moderate
identified supporting actions	41%	29%	18%	12%	0%	71% ($n=12$) Moderate
identified accountable organization(s)	47%	29%	24%	0%	0%	77% ($n=13$) Moderate
identified reporting structures	47%	18%	29%	6%	0%	65% ($n=11$) Moderate
identified funding	47%	29%	24%	0%	0%	77% ($n=13$) Moderate
monitoring and evaluation plan	47%	18%	35%	0%	0%	65% ($n=11$) Moderate
3. How satisfied are you with the scoring rubric for each of the following criteria?						
number and breadth of relevant policies	29%	41%	12%	12%	6%	71% ($n=12$) Moderate
identified supporting actions	41%	35%	12%	12%	0%	77% ($n=13$) Moderate
identified accountable organization(s)	29%	41%	12%	18%	0%	71% ($n=12$) Moderate
identified reporting structures	24%	35%	24%	18%	0%	59% ($n=10$) Low
identified funding	29%	41%	18%	12%	0%	71% ($n=12$) Moderate
monitoring and evaluation plan	35%	35%	18%	12%	0%	71% ($n=12$) Moderate

Notes: ^a only survey participants who had experience using the Welsh approach in the Global Matrix 4.0 Report Cards answered this question ($n = 15$);

^b only survey participants who had experience using the Welsh approach in the 2022 Para Report Cards answered this question ($n = 10$); high group satisfaction level ($\geq 80\%$), moderate (60–79%), low ($< 60\%$).

of the Welsh approach with the original method, including perceived need for modifications, is shown in Fig. 1.

Modifications to the Welsh approach were considered necessary by 71% ($n = 12$) of respondents, including 57% ($n = 4/7$) of respondents who served only on the RC team, 100% ($n = 2/2$) of respondents who served only on the Para RC team, and 75% ($n = 6/8$) of respondents who served on both teams. Among survey respondents, 65% ($n = 11$) supported using the Welsh approach as recommended (optional) in the GM 5.0, and 41% ($n = 7$) supported making it mandatory (replacing the

original method) in future GMs (responses were not mutually exclusive). A summary of user-reported advantages, disadvantages, perceived coverage gaps, and challenges in using the Welsh approach, along with strategies teams used to overcome these challenges, is provided in Table 3. Suggestions gathered through open-ended questions for improving the use of the Welsh approach in future RCs and GMs, and the aspects anticipated to benefit from these changes, are summarized in Table 4.

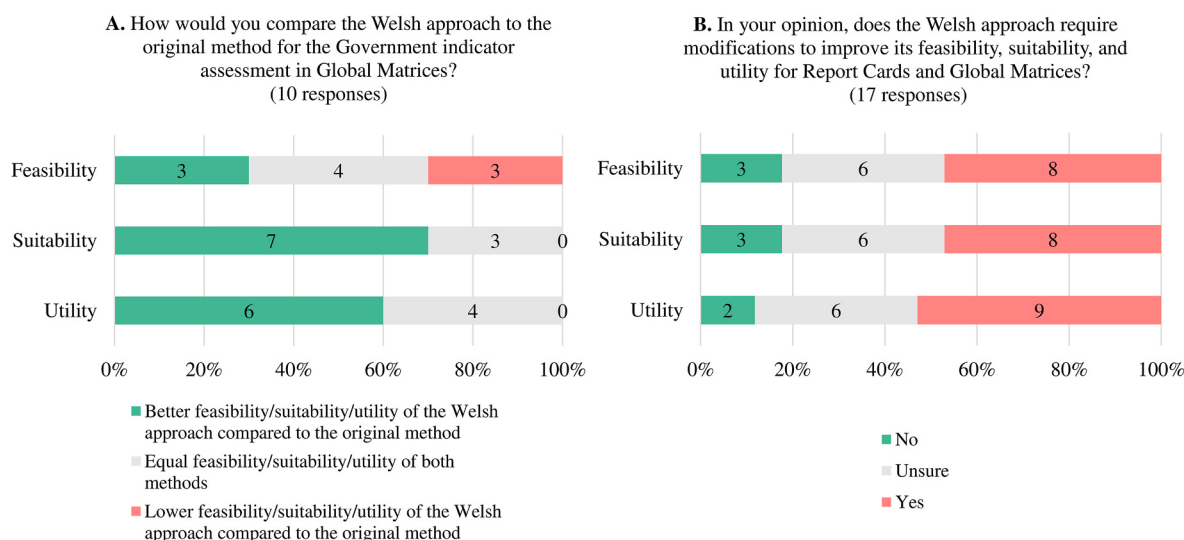


Fig. 1. Comparison of the Welsh approach with the original method used for the Government indicator assessment in Global Matrices (A) and perceived need for modification (B), based on survey responses. *Notes:* Chart A is informed by responses from survey participants with experience using the original method in previous Global Matrices ($n = 10$); Chart B summarizes responses from all survey participants, regardless of their experience with the original method ($n = 17$).

Table 3

Summary of survey findings on advantages, disadvantages, coverage gaps, challenges, and strategies for overcoming challenges related to the use of the Welsh approach, with frequencies of mentions.

Advantages (n)	Disadvantages (n)	Coverage gaps (n)	Challenges (n)	Strategies for overcoming challenges (n)
Structure (6), Transparency (6), Thoroughness/ comprehensiveness/ systematic nature (6), Quantitative component/ scoring (5), Objectivity (4), Clarity (4), Enabling comparisons (4), Scientific rigor (1).	Time investment/workload (6), Complexity (3), Need for specific expertise/ training to use it (3), Criterion weighting (require further revision/ adaptation) (2), Lack of clarity of certain aspects (e.g., data gathering process, definitions) (2), Elements of subjectivity in data interpretation (1), Missing dimension (policy effectiveness) (1), Too broad criteria (1).	Policy implementation (4), Policy impact/ effectiveness (2), Consideration of context specifics (2), Equity (1), Continuity/longevity of policy instruments (1), Consideration of funds origin in scoring (1), Informal/emerging government actions (1), Stakeholder engagement (1), Lack of policy (1).	Ensuring completeness of policy-related data gathering (3), Assessing certain components (e.g., budget, supporting actions) (2), Inexperience/learning the new method (2), Lack of clarity of certain aspects (e.g., definitions) (1), Justifying the grade decline over time despite increased policies (1), Feasibility issues/workload (1), Suitability of the method for the subpopulation of children and adolescents with disabilities (1), Scoring using the predefined weighting of criteria, which may not fully reflect their relative importance and could benefit from further expert review (1).	Collaborations (e.g., discussed with team members, engaged stakeholders from various sectors, consulted with colleagues from other RC teams) (5), Analytical considerations (e.g., context- specific adaptations, structuring) (3), Research (e.g., international examples) (2), Increased time allocation (1), Transparent reporting of the impact of the method change on the grade (1).

Notes: n = number of mentions across responses from all survey participants.

3.2. Government indicator grade differences between the Global Matrix 3.0 and 4.0

Among countries that participated in both GM 3.0 and 4.0, 35/40 had sufficient evidence to consistently assign a letter grade to the Government indicator, including six countries that transitioned to the Welsh approach in the GM 4.0 (group 1) and 28 countries that retained the original method (group 2). Table 5 summarizes the Government indicator grades and grade differences for countries and country groups across both GMs. Wilcoxon signed-rank tests revealed no statistically significant within-group differences between the GM 3.0 and 4.0 grades for either group 1 ($p = 0.25$) or group 2 ($p = 0.62$).

4. Discussion

4.1. User perspectives on the Welsh approach and implications for grading

The Welsh approach received predominantly positive feedback from users, who highlighted its structure, transparency, and

comprehensiveness as key advantages, as well as its greater utility and suitability for GMs and RCs compared with the original method. However, survey respondents represented a set of countries that self-selected to use the Welsh approach in their RCs or Para RCs – a decision that may have been influenced by higher capacity, greater familiarity with structured tools, and better access to data than other teams, potentially contributing to more favorable feedback. Among the aspects explored, feasibility Fig. 1 and suitability for CAWD Table 1 received comparatively less support. Consistent with GM 3.0 and 4.0 evaluations,^{25,31} some teams were small and lacked sector representation, which may have increased workload and reduced the diversity of perspectives, potentially contributing to feasibility concerns. The lower suitability for CAWD was mainly driven by neutral responses, likely due to limited policy documentation available for grading. Furthermore, some participants felt the scoring rubric insufficiently captured policy implementation and effectiveness, meaning the Welsh approach could potentially yield higher grades based primarily on policy existence and scope, even when implementation is limited. Notably, however, policy effectiveness can only be meaningfully assessed after a policy has been

Table 4

A summary of survey respondents' suggestions to improve the application of the Welsh approach in Report Cards and Global Matrices, and the aspects they anticipate will benefit as a result.

Suggestions for improvement (n)	Aspects expected to improve, as indicated by respondents
Revise the score classification thresholds/the weighting of criteria in the scoring rubric (e.g., assign a higher weight to the monitoring and evaluation plan criterion) (8)	Feasibility, Validity, Suitability, Utility
Include parameters for policy implementation and/or policy effectiveness in the assessment (7)	Feasibility, Validity, Suitability, Utility
Provide clearer/more explicit definitions for key terms/criteria (7)	Feasibility, Validity, Reliability, Utility
Provide examples for the scoring process (e.g., how to score partial policy implementation) (7)	Feasibility, Validity, Reliability, Utility
Integrate a discussion/consensus component (e.g., on policy sustainability across political cycles, policy gaps, the influence of political cycles on policy) (4)	Validity, Suitability
Simplify the scoring rubric (e.g., reduce the number of scoring categories, focus on the most critical elements) (3)	Feasibility, Suitability, Utility
Create a user manual on how to apply the Welsh approach (3)	Validity, Reliability, Utility
Allow for contextual flexibility in applying and interpreting key criteria across countries/jurisdictions (e.g., contextual adjustment protocol) (3)	Feasibility, Validity, Suitability
Provide training to Report Card teams on how to apply the Welsh approach (2)	Suitability, Utility
Conduct pilot testing in countries with diverse income levels and governance structures to inform contextual adaptations (2)	Validity, Utility
Make the methodology/scoring rubric adjustable to country/jurisdiction-specific context (2)	Suitability, Utility
Offer guidance on comprehensive data gathering/how to identify whether relevant policy documents exist (2)	Reliability, Utility
Implement inter-rater checks and calibration exercises for scorers (2)	Reliability
Design a scoring system flowchart (1)	Feasibility
Invite government sector representative(s) to the Report Card team (1)	Feasibility
Develop a tool to summarize scores and key notes from policies automatically (1)	Feasibility
Translate into other languages (1)	Feasibility
Incorporate cross-sectoral validation by involving experts from various sectors (e.g., education, health, government) (1)	Validity
Conduct a study comparing grades/scores within the same country using both methods (the original and the Welsh approach) (1)	Validity
Recommend assessment of policy implementation at the local/regional level (beyond a national lens) (1)	Suitability
Include more parameters in the assessment (e.g., target population, policy characteristics) (1)	Utility

Notes: n = number of mentions across responses from all survey participants.

in place for sufficient time, which does not always align with the point at which the assessment is conducted. This aligns with literature demonstrating that many countries have established PA policies, but their implementation and effectiveness are generally rated as low to moderate.^{32,33} The translation of PA policies into effective practice remains slow, uneven, and inequitable due to a number of factors, including political, financial, human, and technical constraints.³⁴ A policy-behavior paradox observed in the GMs^{3,4} (where countries with well-developed supports, including PA policies and strategies, demonstrated lower levels of PA behaviors) further underscores the persistence of gaps between policy and practice and highlights the importance of context-specific considerations for effective implementation.

Regarding the impact of the methodology transition, some notable

changes in the Government indicator grades (between GM 3.0 and 4.0) were observed at the individual country level in South Korea,³⁵ Uruguay,³⁶ Chile,³⁷ and Colombia.³⁸ This could potentially be influenced to some extent by the adoption of the Welsh approach. In the case of South Korea, the Government indicator grade improved from D (GM 3.0) to A (GM 4.0). As suggested by Lee et al., although the quality of policy and program implementation and evaluation remains unclear, this improvement may likely have been driven by the use of the Welsh approach.³⁵ In Uruguay, the Government indicator grade improved from D (GM 3.0) to C (GM 4.0), noting that two new pieces of legislation were approved during this period and contributed to the assessment, but the authors also interpreted the change in grade with caution due to the switch to the Welsh approach.³⁶ In Chile, the grade improved from B– (GM 3.0) to A– (GM 4.0). The high grade in the GM 4.0 was assigned despite substantial fragmentation among those responsible for implementing policies across different government sectors and limited evaluation of implementation quality. Aguilar-Farias et al. therefore noted that the grade may overestimate Chile's performance on the Government indicator, as the authors considered that the Welsh approach has limitations related to capturing these aspects.³⁷ In contrast, in Colombia, the Government indicator grade decreased from B (GM 3.0) to C+ (GM 4.0) following the introduction of the Welsh approach,³⁸ which the team considered helpful in providing a more reliable assessment and mitigating potential sources of bias that may arise from the composition of the RC expert group. Similar positive or negative grade changes occurred in countries continuing with the original method (e.g., India,³⁹ Slovenia,⁴⁰ Lebanon⁴¹), reflecting either enhanced governmental actions or reduced support for PA, for example, during the COVID-19 pandemic. At the group level, grade differences were neither consistent nor substantial enough to reach statistical significance, providing preliminary evidence of comparability between the two methods. Furthermore, the GM Government indicator definition and benchmarks from the original method (with scope-specific adaptations), and the Welsh approach were recently recommended for inclusion in the early childhood 24-h movement behavior RCs through an international Delphi consensus process involving over 50 experts from 15 countries.⁴²

4.2. Recommendations

The findings of this study provide a foundation for the following recommendations:

- 1) It is recommended that an expert group review the scoring rubric for the Welsh approach to support refinement.

Assessing how policies are translated into practice is important for evaluating government support for PA, and there is growing recognition of the need to understand policy implementation processes and evaluation.^{28,43} In the Welsh approach, policy implementation currently begins with the “identified supporting actions” and “identified accountable organizations” criteria, but it could potentially be captured more effectively as a separate criterion or dedicated aspect of the assessment. For example, the RCs' narrative section, which provides a rationale for the grades, could be used to describe specific implementation-related issues that may not be adequately captured by the scoring rubric. Increasing the relative weight of the “monitoring and evaluation plan” criterion could give greater consideration to policy effectiveness in the assessment. Providing a list of definitions for policy-related terms would support both comprehension and consistent application. In the meantime, the current version of the Welsh approach can continue to be used, with any specific implementation concerns captured in the grade rationale and with the glossary of terms from the HEPA PAT v2 document²⁴ consulted as needed.

Table 5

Summary of Government indicator grades and grade differences between GM 3.0 and 4.0 across countries participating in both editions, grouped by indicator assessment method and ordered by the corresponding number difference.

Groupings	Country/ Jurisdiction	GM 3.0		GM 4.0		Number difference (between GM 4.0 and 3.0)
		Grade	Number	Grade	Number	
Countries that used the original method in GM 3.0 and the Welsh approach in GM 4.0 (Group 1 , $n = 6$)	Colombia	B	11	C+	9	-2
	Brazil	D+	6	D+	6	0
	Scotland	C	8	C	8	0
	Chile	B-	10	A-	13	3
	Uruguay	D	5	C	8	3
	South Korea	D	5	A	14	9
	Mean	C-	[7.5] = 7	C+	[9.7] = 9	2
	Range	D to B	5 to 11	D+ to A	6 to 14	
Countries that used the original method in GM 3.0 and 4.0 (Group 2 , $n = 28$)	Slovenia	A	14	D	5	-9
	Botswana	C	8	D-	4	-4
	Lebanon	C+	9	D	5	-4
	Czech Republic	C+	9	D+	6	-3
	Zimbabwe	C-	7	D	5	-2
	Denmark	A-	13	B+	12	-1
	Poland	C+	9	C	8	-1
	Thailand	B+	12	B	11	-1
	Chinese Taipei	B+	12	B+	12	0
	Estonia	B	11	B	11	0
	Finland	A-	13	A-	13	0
	Japan	B	11	B	11	0
	Mexico	C	8	C	8	0
	Portugal	B	11	B	11	0
	South Africa	C	8	C	8	0
	Sweden	B	11	B	11	0
	United Arab Emirates	B+	12	B+	12	0
	Canada	C+	9	B-	10	1
	Hong Kong, China	C	8	C+	9	1
	Lithuania	C	8	C+	9	1
	Australia	D	5	C-	7	2
	New Zealand	B+	12	A	14	2
	China	F	2	D	5	3
	Ethiopia	D	5	C	8	3
	France	C	8	B	11	3
	Jersey	D	5	C	8	3
	Guernsey	D	5	C+	9	4
	India	D	5	C+	9	4
	Mean	C	[8.9] = 8	C+	[9.0] = 9	1
	Range	F to A	2 to 14	D- to A	4 to 14	
Country that used the Welsh approach in GM 3.0 and 4.0 ($n = 1$)	Wales	C+	9	C	8	-1
Countries that assigned an INC grade in at least one of the GM 3.0 or 4.0 ($n = 5$)	England	INC	n/a	INC	n/a	n/a
	Germany	INC	n/a	INC	n/a	n/a
	Nepal	INC	n/a	F	2	n/a
	Spain	INC	n/a	C	8	n/a
	United States	INC	n/a	INC	n/a	n/a

Abbreviations: GM = Global Matrix, INC = incomplete, n = number, n/a = not applicable. **Notes:** The conversion of letter grades to corresponding numbers was based on the following scale: A+ = 15, A = 14, A- = 13, B+ = 12, B = 11, B- = 10, C+ = 9, C = 8, C- = 7, D+ = 6, D = 5, D- = 4, F = 2.

2) RC teams are encouraged to develop effective, contextually relevant processes for assessing the Government indicator in alignment with the GM methodology.

For both methods, effective assessment involves the RC teams allocating sufficient time (e.g., several months) and ensuring a collaborative, multidisciplinary working group with representatives from relevant government sectors and/or other policy actors. Where multi-sectoral or government representation is not feasible, external consultation (e.g., with experts from government-adjacent organizations, other academics) should be considered. First-time users of the Welsh approach are advised to consult experiences from other countries to improve understanding and application of the methodology (e.g., policy analyses for RCs in the Philippines,⁴⁴ Brazil,⁴⁵ Scotland,⁴⁶ and Canada⁴⁷) and to consider strategies for addressing challenges employed by previous users (Table 4).

3) Continued investment in research and knowledge exchange is recommended to strengthen the assessment of government policies and support for PA within the GM and beyond.

Future work could explore the perspectives of RC teams that opted not to use the Welsh approach to better understand perceived barriers and identify opportunities to enhance the method's accessibility and usability across diverse contexts. Given the absence of a gold standard, full validation of the Welsh approach is not feasible; however, comparing grades assigned with the same evidence under both methods in the same GM edition could provide useful insight into potential differences in grading outcomes. The reliability of the Welsh approach may be empirically explored, for example, by having multiple raters score the same dataset and comparing their scores (inter-rater reliability). In the interim, retaining both methods (the original and the Welsh approach) in the GM is considered acceptable, allowing selection of the method best suited to country-specific contexts. However, it should be acknowledged that this, to some extent, reduces harmonization across countries. Webinar(s) or training session(s) led by the AHKGA on

assessing the Government indicator in GMs, including its application to the CAWD subpopulation, could help promote greater harmonization, accuracy, and consistency in assessments.

Importantly, the AHKGA's methodology has broader applicability beyond the GM initiative, as demonstrated by its application in COVID-19-focused RCs in Scotland⁴⁸ and France,⁴⁹ the Para RC initiative for CAWD,²⁶ and the development of a movement behaviors assessment framework for early childhood.⁴² Policy analysis in PA also evolved through other dedicated frameworks and tools (e.g., the system-based Policy Evaluation Network [PEN] framework,⁵⁰ PA policy benchmarking framework "MOVING",⁵¹ the Comprehensive Analysis of Policy on PA [CAPPA] framework,⁵² the PA Environment Policy Index [PA-EPI],⁵³ and GoPA! Policy Inventory, version 3.0³²). Together, these developments can guide efforts to establish more robust and objective methodologies and their adaptation to different populations and contexts.

4.3. Strengths and limitations

Strengths of this study include the integration of quantitative and qualitative components and international representation in the survey, marking it as the first multi-country evaluation of the Welsh approach. Quantitative data on changes in grades were considered alongside qualitative data from expert perceptions, providing complementary perspectives on the use of the method and its potential relationship with grades. The inclusion of representatives from both GM 4.0 and Para RC initiatives broadens the scope and relevance of the findings. Nonetheless, certain limitations should be acknowledged. The quantitative analysis was constrained by a relatively small sample size, and sampling bias could have contributed to an over-representation of favorable perspectives in the qualitative component. Additionally, some participants provided more expressive feedback, which may have led to their perspectives being more prominently represented. Involving two reviewers in content analysis enabled cross-checking of interpretations, helping to mitigate potential researcher bias.

5. Conclusion

The Welsh approach represents progress toward a structured, transparent, and comprehensive assessment of the Government indicator, suitable for the GMs. However, no significant differences in indicator grades were identified within multi-country groups, irrespective of whether the Welsh or original method was applied. Further research is warranted to explore the reliability of the Welsh approach and its applicability across diverse contexts, including beyond GMs. While both the Welsh and original methods remain appropriate for interim use, future GM editions would benefit from refining the Welsh approach, including stronger attention to policy implementation and evaluation, and to harmonization and optimization of assessment processes across countries.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Author contributions

ID, MST, SA, NK: conception and design of the study; ID: analyses and original draft; SA: content analysis; all authors: data interpretation, critical revision for important intellectual content, and final approval of the version to be submitted; MST: directing the project.

Funding

None declared.

Declaration of competing interest

ID was an AHKGA Fellow, and SA, JBS, DASS, and MST were members of the AHKGA Board of Directors during the conduct of this study. LCE and MW were part of the expert team that developed the Welsh approach. All other authors declare no competing interests.

Acknowledgements

The authors gratefully acknowledge the Welsh research team, including Malcolm Ward, Richard Tyler, Lowri C Edwards, Maggie C Miller, Simon Williams, and Gareth Stratton, who developed and published the method evaluated in this study, for their contributions to the field and for informing the present work.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jesf.2026.200520>.

References

- Brusseau T, Fairclough S, Lubans D. *The Routledge Handbook of Youth Physical Activity*. Routledge [VitalSource Bookshelf version]; 2020. vbk://9781000050707.
- Chaput JP, Willumsen J, Bull F, et al. 2020 WHO guidelines on physical activity and sedentary behaviour for children and adolescents aged 5–17 years: summary of the evidence. *Int J Behav Nutr Phys Act*. 2020;17(1):141. <https://doi.org/10.1186/s12966-020-01037-z>.
- Tremblay MS, Gray CE, Akinroye K, et al. Physical activity of children: a Global Matrix of grades comparing 15 countries. *J Phys Activ Health*. 2014;11(s1):S113–S125. <https://doi.org/10.1123/jpah.2014-0177>.
- Tremblay MS, Barnes JD, González SA, et al. Global Matrix 2.0: Report Card grades on the physical activity of children and youth comparing 38 countries. *J Phys Activ Health*. 2016;13(s2):S343–S366. <https://doi.org/10.1123/jpah.2016-0594>.
- Aubert S, Barnes JD, Abdeta C, et al. Global Matrix 3.0 physical activity Report Card grades for children and youth: results and analysis from 49 countries. *J Phys Activ Health*. 2018;15(s2):S251–S273. <https://doi.org/10.1123/jpah.2018-0472>.
- Aubert S, Barnes JD, Demchenko I, et al. Global Matrix 4.0 physical activity Report Card grades for children and adolescents: results and analyses from 57 countries. *J Phys Activ Health*. 2022;19(11):700–728. <https://doi.org/10.1123/jpah.2022-0456>.
- Guthold R, Stevens GA, Riley LM, Bull FC. Global trends in insufficient physical activity among adolescents: a pooled analysis of 298 population-based surveys with 1–6 million participants. *Lancet Child Adolesc Health*. 2020;4(1):23–35. [https://doi.org/10.1016/S2352-4642\(19\)30323-2](https://doi.org/10.1016/S2352-4642(19)30323-2).
- Global Action Plan on Physical Activity 2018–2030: *More Active People for a Healthier World*. Geneva: World Health Organization; 2018. <https://www.who.int/publications/i/item/9789241514187>. Accessed April 2, 2026.
- ISPAH International Society for Physical Activity and Health. The Bangkok declaration on physical activity for global health and sustainable development. *Br J Sports Med*. 2017;51(19):1389–1391. <https://doi.org/10.1136/bjsports-2017-098063>.
- Salvo D, Crochemore-Silva I, Wendt A, et al. Physical activity for public health in the 21st century. *Nat Med*. 2026;32:1479–1489. <https://doi.org/10.1038/s41591-026-04237-5>.
- Sallis JF, Cervero RB, Ascher W, Henderson KA, Kraft MK, Kerr J. An ecological approach to creating active living communities. *Annu Rev Public Health*. 2006;27:297–322. <https://doi.org/10.1146/annurev.publhealth.27.021405.102100>.
- Kohl HW, Craig CL, Lambert EV, et al. The pandemic of physical inactivity: global action for public health. *Lancet*. 2012;380(9838):294–305. [https://doi.org/10.1016/S0140-6736\(12\)60898-8](https://doi.org/10.1016/S0140-6736(12)60898-8).
- Zhao C, Yuan J, Huang WW. Multilevel determinants of physical activity in children and adolescents: a meta-analysis guided by social ecological model. *BMC Sports Sci Med Rehabil*. 2025;17(1):1. <https://doi.org/10.1186/s13102-025-01225-z>.
- Hill J, Shewbridge C. How can OECD countries empower children to be more physically active? *OECD Educ Work Pap*. 2025;339. <https://doi.org/10.1787/f534b2d0-en>.
- Reilly JJ, Aubert S, Brazo-Sayavera J, Liu Y, Cagas JY, Tremblay MS. Surveillance to improve physical activity of children and adolescents. *Bull World Health Organ*. 2022;100(12):815–824. <https://doi.org/10.2471/BLT.22.288569>.
- Morrison SA, Aubert S, Brazo-Sayavera J, Demchenko I, Reilly JJ, Tremblay MS. Celebrating a decade of the Active Healthy Kids Global Alliance and its Global Matrix initiative. *Br J Sports Med*. 2025;59(3):198–199. <https://doi.org/10.1136/bjsports-2024-108984>.
- Participants » Active Healthy Kids Global Alliance. Active Healthy Kids Global Alliance. <https://www.activehealthykids.org/participants/>. Accessed April 2, 2026.
- Global Matrix 5.0 » Active Healthy Kids Global Alliance. Active Healthy Kids Global Alliance. <https://www.activehealthykids.org/5-0/>. Accessed April 2, 2026.

19. Aubert S, Barnes JD, Forse ML, et al. The international impact of the Active Healthy Kids Global Alliance physical activity Report Cards for children and youth. *J Phys Activ Health*. 2019;16(9):679–697. <https://doi.org/10.1123/jpah.2019-0244>.
20. Tremblay MS, Barnes JD, Demchenko I, et al. Active Healthy Kids Global Alliance Global Matrix 4.0—A resource for physical activity researchers. *J Phys Activ Health*. 2022;19(11):693–699. <https://doi.org/10.1123/jpah.2022-0257>.
21. Klepac Pogrmilovic B, O'Sullivan G, Milton K, et al. A global systematic scoping review of studies analysing indicators, development, and content of national-level physical activity and sedentary behaviour policies. *Int J Behav Nutr Phys Act*. 2018; 15(1):1–17. <https://doi.org/10.1186/s12966-018-0742-9>.
22. Ward MR, Tyler R, Edwards LC, Miller MC, Williams S, Stratton G. The AHK-Wales Report Card 2018: Policy measures - is it possible to “score” qualitative data? *Health Promot Int*. 2021;36(4):1151–1159. <https://doi.org/10.1093/heapro/daaa118>.
23. Edwards LC, Tyler R, Blain D, et al. Results from Wales' 2018 Report Card on physical activity for children and youth. *J Phys Activ Health*. 2018;15(s2): S430–S432. <https://doi.org/10.1123/jpah.2018-0544>.
24. Bull F, Milton K, Kahlmeier S. *Health-Enhancing Physical Activity (HEPA) Policy Audit Tool (PAT) Version 2*. Copenhagen: World Health Organization; 2015. <https://iris.who.int/items/76bead5d-986c-4ccd-bf2e-de9276d2dc3d>.
25. Demchenko I, Aubert S, Tremblay MS. Global Matrix 4.0 on physical activity for children and adolescents: project evolution, process evaluation, and future recommendations. *J Phys Activ Health*. 2025;22(4):457–469. <https://doi.org/10.1123/jpah.2024-0513>.
26. Ng K, Sit C, Arbour-Nicitopoulos K, et al. Global Matrix of Para Report Cards on Physical Activity of Children and Adolescents With Disabilities. *Adapt Phys Act Q*. 2023;40(3):409–430. <https://doi.org/10.1123/apaq.2022-0111>.
27. Tremblay MS, Demchenko I, Reilly JJ, Aubert S, Sit C. The Future of Para Report Cards on Physical Activity of Children and Adolescents With Disabilities-A Global Call for Engagement, Data, and Advocacy. *Adapt Phys Act Q*. 2024;41(1):1–8. <https://doi.org/10.1123/apaq.2023-0131>.
28. McLaughlin M, McCue P, Swelam B, Murphy J, Edney S. Physical activity—the past, present and potential future: a state-of-the-art review. *Health Promot Int*. 2025;40(1), daae175. <https://doi.org/10.1093/heapro/daae175>.
29. Albert B, Tullis T. *Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics*. Elsevier Science & Technology; 2013. <https://doi.org/10.1016/C2011-0-00016-9>.
30. Elo S, Kyngäs H. The qualitative content analysis process. *J Adv Nurs*. 2008;62(1): 107–115. <https://doi.org/10.1111/j.1365-2648.2007.04569.x>.
31. Aubert S, Barnes JD, Tremblay MS. Evaluation of the process and outcomes of the Global Matrix 3.0 of physical activity grades for children and youth. *J Exerc Sci Fit*. 2020;18(2):80–88. <https://doi.org/10.1016/j.jesf.2020.01.002>.
32. Klepac Pogrmilovic BK, Ramirez Varela A, Pratt M, et al. National physical activity and sedentary behaviour policies in 76 countries: availability, comprehensiveness, implementation, and effectiveness. *Int J Behav Nutr Phys Act*. 2020;17(1):116. <https://doi.org/10.1186/s12966-020-01022-6>.
33. Ramirez Varela A, Bauman A, Woods CB, et al. Low global physical activity despite two decades of policy progress. *Nat Health*. 2026;1(3):338–354. <https://doi.org/10.1038/s44360-025-00044-3>.
34. *Global Status Report on Physical Activity 2022*. Geneva: World Health Organization; 2022. <https://www.who.int/publications/i/item/9789240059153>. Accessed April 2, 2026.
35. Lee EY, Kim YB, Goo S, et al. Physical activity in the era of climate change and COVID-19 pandemic: results from the South Korea's 2022 Report Card on physical activity for children and adolescents. *J Exerc Sci Fit*. 2023;21(1):26–33. <https://doi.org/10.1016/j.jesf.2022.10.014>.
36. Brazo-Sayavera J, Fernandez-Gimenez S, Pintos-Toledo E, Corvos C, Souza-Marabotto F, Bizzozero-Peroni B. Results from the Uruguay's 2022 Report Card on physical activity for children and adolescents. *J Exerc Sci Fit*. 2023;21(1):104–110. <https://doi.org/10.1016/j.jesf.2022.11.005>.
37. Aguilar-Farias N, Miranda-Marquez S, Toledo-Vargas M, et al. Results from Chile's 2022 Report Card on physical activity for children and adolescents. *J Exerc Sci Fit*. 2024;22(4):390–396. <https://doi.org/10.1016/j.jesf.2024.07.004>.
38. Colombia » Active Healthy Kids Global Alliance. Active Healthy Kids Global Alliance. <https://www.activehealthykids.org/colombia/>. Accessed August 11, 2026.
39. Bhawra J, Khadilkar A, Krishnaveni GV, Kumaran K, Katapally TR. The 2022 India Report Card on physical activity for children and adolescents. *J Exerc Sci Fit*. 2023; 21(1):74–82. <https://doi.org/10.1016/j.jesf.2022.10.013>.
40. Morrison SA, Jurak G, Starc G, et al. Challenges of social change: the 2021 Republic of Slovenia report card on physical activity of children and adolescents. *J Exerc Sci Fit*. 2023;21(4):305–312. <https://doi.org/10.1016/j.jesf.2023.06.003>.
41. Abi Nader P, Hadla R, Majed L, Mattar L, Sayegh S. Results from the Lebanese 2022 report card on physical activity for children and youth. *J Exerc Sci Fit*. 2023;21(1): 14–19. <https://doi.org/10.1016/j.jesf.2022.10.007>.
42. Huang WY, Zhang D, Tremblay MS, et al. Developing international consensus on key indicators for a 24-Hour movement behaviour Report Card in early childhood: a twin-panel Delphi study. *Sports Med Open*. 2025;11(1):92. <https://doi.org/10.1186/s40798-025-00905-x>.
43. Forberger S, Reisch LA, Meshkova B, et al. What we know about the actual implementation process of public physical activity policies: results from a scoping review. *Eur J Publ Health*. 2022;32(Suppl 4):iv59–iv65. <https://doi.org/10.1093/eurpub/ckac089>.
44. Palad YY, Guisihan RM, Aguila MER, Ramos RAA, Cagas JY. An evaluation of policies promoting physical activity among Filipino youth. *Int J Environ Res Publ Health*. 2023;20(4):2865. <https://doi.org/10.3390/ijerph20042865>.
45. Silva DAS, Pelegrini A, Christofaro DGD, et al. Results from Brazil's 2022 Report Card on physical activity for children and adolescents. *Int J Environ Res Publ Health*. 2022;19(16):16. <https://doi.org/10.3390/ijerph191610256>.
46. Tomaz SA, Reilly JJ, Johnstone A, et al. A systematic evaluation of physical activity and diet policies in Scotland: results from the 2021 Active Healthy Kids Report Card. *J Public Health*. 2024;46(2):e240–e247. <https://doi.org/10.1093/pubmed/fdae022>.
47. Kuzik N, Bruijns BA, Cameron C, et al. Examining the grading process for the government indicator of the 2024 ParticipACTION Report Card on physical activity for children and youth. *Health Promot Chronic Dis Prev Can*. 2026;46(9) (in press).
48. Active Healthy Kids Scotland. https://www.activehealthykidsscotland.co.uk/report_card_covid.html. Accessed August 12, 2026.
49. Fillon A, Genin P, Larras B, et al. France's 2020 Report Card on physical activity and sedentary behaviors in children and youth: results and progression. *J Phys Activ Health*. 2021;18(7):811–817. <https://doi.org/10.1123/jpah.2021-0025>.
50. Kamphuis CBM, Forberger S, Lien N, Rehfuess E, Luszczynska A, on behalf of the PEN consortium. Development of a framework to guide research into policies promoting physical activity and healthy diets in the European context: the system-based policy evaluation network (PEN) framework. *Eur J Publ Health*. 2022;32 (Supplement 4):iv3–iv9. <https://doi.org/10.1093/eurpub/ckac068>.
51. Vlad I, Oldridge-Turner K, Klepp KI, et al. The development of the NOURISHING and MOVING benchmarking tools to monitor and evaluate national governments' nutrition and physical activity policies to address obesity in the European region. *Obes Rev*. 2023;24(S1), e13541. <https://doi.org/10.1111/obr.13541>.
52. Klepac Pogrmilovic B, O'Sullivan G, Milton K, et al. The development of the comprehensive analysis of policy on physical activity (CAPPA) framework. *Int J Behav Nutr Phys Act*. 2019;16(1):60. <https://doi.org/10.1186/s12966-019-0822-5>.
53. Woods CB, Kelly L, Volf K, et al. The physical activity environment policy index for monitoring government policies and actions to improve physical activity. *Eur J Publ Health*. 2022;32(Supplement 4):iv50–iv58. <https://doi.org/10.1093/eurpub/ckac062>.