PART E. SYSTEMATIC REVIEW LITERATURE SEARCH METHODOLOGY

Table of Contents

OVERVIEW

Under the direction of the Office of Disease Prevention and Health Promotion (ODPHP), the National Institutes of Health (NIH), the Centers for Disease Control and Prevention (CDC), and the President's Council on Fitness, Sports and Nutrition (PCFSN), ICF (a contractor), herein referred to as the literature review team, was responsible for supporting the 2018 Physical Activity Guidelines Advisory Committee in reviewing the scientific literature used to support the development of its report.

The literature review team used a methodology informed by best practices for systematic reviews (SRs) developed by the United States Department of Agriculture's (USDA) Nutrition Evidence Library (NEL),[1] the Agency for Healthcare Research and Quality (AHRQ),[2] the Cochrane Collaboration,[3] and the Health and Medicine Division of the National Academies of Sciences, Engineering, and Medicine SR standards[4] to review, evaluate, and synthesize published, peer-reviewed physical activity research. The literature review team's rigorous, protocol-driven methodology was designed to maximize transparency, minimize bias, and ensure the SRs conducted by the Committee were relevant, timely, and of high quality. Using this evidence-based approach enabled compliance with the Data Quality Act,[5] which states that federal agencies must ensure the quality, objectivity, utility, and integrity of the information used to form federal guidance. Strict quality control processes were implemented throughout the Committee's process to ensure transparency, integrity, reproducibility, and research excellence in design, implementation, and synthesis of the SRs.

The 2018 Scientific Report process was led by the Committee, with support from a federal leadership team. All work completed by the literature review team was under the direction and review of the Committee members. The literature review team[i] comprised several groups:

- A training and quality control team that developed an abstraction tool and accompanying abstraction guide, developed and implemented training and quality control protocols, and ensured overall quality and integrity of the Committee's SRs,
- SR liaisons, who managed the literature review team's workflow for their designated Subcommittee(s) and/or Work Group,
- Librarians, who reviewed search strategies, confirmed search results, and retrieved full text articles,
- A triage team that participated in a 5-hour triage training before conducting title and abstract triage of original articles, existing reports, SRs, meta-analyses (MAs), and pooled analyses identified through the literature searches, and
- Abstractors, who participated in a three-phase, five-week virtual training before abstracting data from original articles, existing reports, SRs, MAs, and pooled analyses. They also assessed

[i] All literature review team staff were required to disclose potential conflicts of interest or professional bias before working on this team. No conflicts of interest or bias were identified.

bias of original articles and assessed the quality of existing reports, SRs, MAs, and pooled analyses.

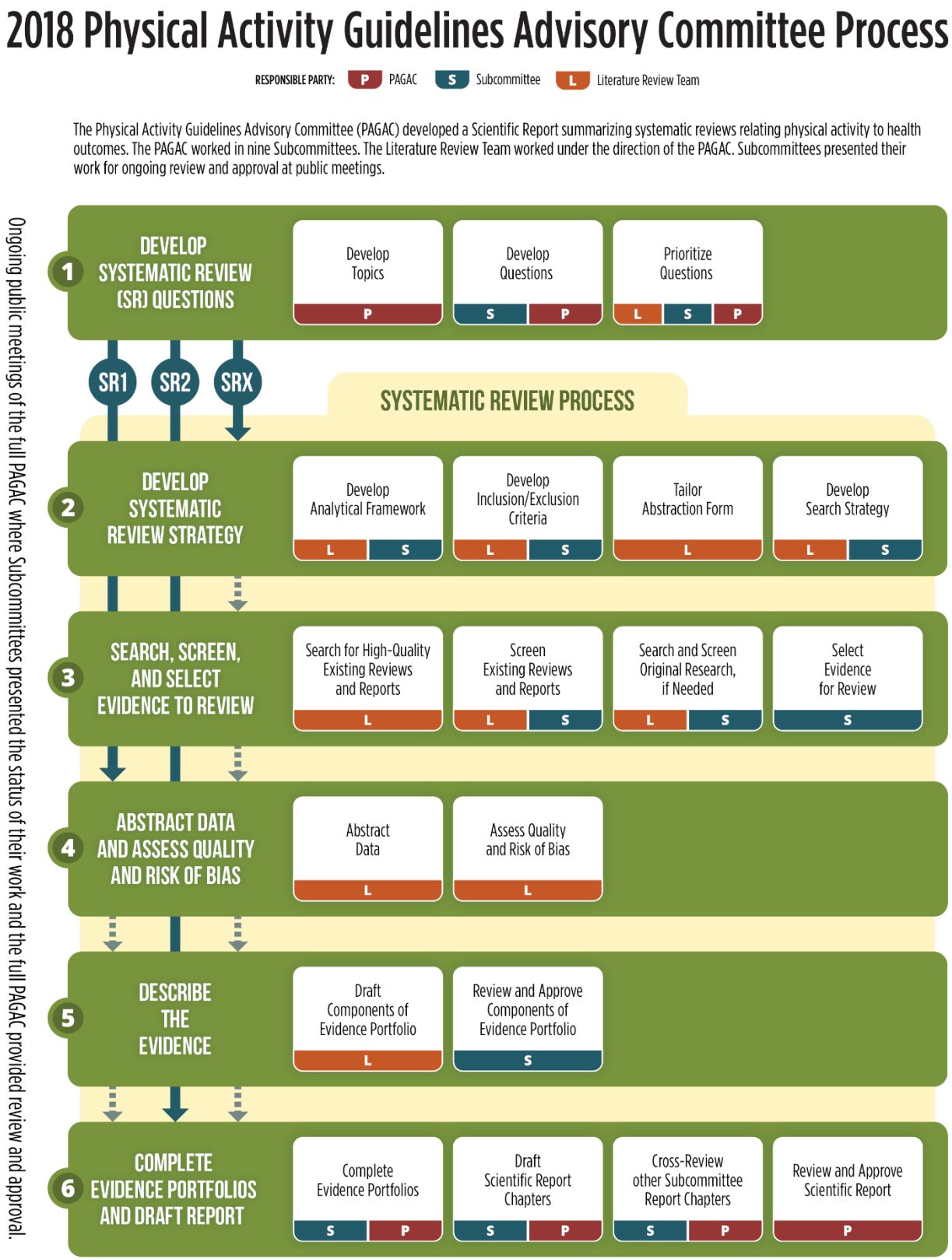
A six-step process was used to develop the Scientific Report:

- Step 1: Develop systematic review questions
- Step 2: Develop systematic review strategy
- Step 3: Search, screen, and select evidence to review for each question
- Step 4: Abstract data and assess the quality and risk of bias of the research
- Step 5: Describe the evidence
- Step 6: Complete evidence portfolios and draft Scientific Report

Figure E-1 provides a visual representation of this process. The model displays the six overarching steps and the associated tasks within each step. It also shows that at any given time, multiple SRs were being executed. For each SR, Steps 2 through 6 were completed sequentially. Throughout the life of the Committee Subcommittees presented the status of their work at in-person public meetings for review and approval by the full Committee. The responsible parties for each task (full Committee, Subcommittee, and/or literature review team[ii]) are included in the model.

[ii] Because federal staff served in an support role, specific tasks were not assigned to them.

Figure E-1. 2018 Physical Activity Guidelines Advisory Committee Process Model



SYSTEMATIC REVIEW PROCESS

Step 1: Develop Systematic Review Questions

In 2014, a federal planning group led by ODPHP, NIH, CDC, and PCFSN organized a potential scope and state of the science meeting with experts from around the country to gather information on whether sufficient new evidence was available to update the *2008 Physical Activity Guidelines for Americans*.[6] Based on the *Physical Activity Guidelines Advisory Committee Report, 2008,*[7] and a summary of areas of rapidly developing science, the group identified a number of key areas with new research available: youth younger than age 6 years, older adults, brain health across the life span, dose-response, and sedentary behavior. In early 2016, the literature review team conducted a scoping exercise to determine the amount of literature published on topics included in the 2008 Scientific Report[7] since the completion of that report. The Committee used the list of key topics from 2014, the summary of the scoping exercise, and their expertise to determine the final list of topics to examine.

At their first public meeting, the Committee decided on topics and formed Subcommittees. The Subcommittee members then developed and refined clearly focused SR questions and subquestions within each topic, which were used to systematically search the existing literature. The development of the SR questions took place during Subcommittee calls.

Prioritize SR Questions

After formulating a list of SR questions, Subcommittee members ranked the questions based on the following:

- Potential for greatest public health impact
- Potential to inform public health policy and/or programs
- Existence of mature scientific evidence
- Potential generalizability

The SR questions and their prioritization were reviewed and revised by the full Committee during their second public meeting. Any refinements to the questions or questions developed after the second public meeting were presented to leaders of all the Subcommittees for their review and approval.

Step 2: Develop Systematic Review Strategy

Develop Analytical Frameworks

The Subcommittees developed an analytical framework for each of their SR questions. Analytical frameworks are a visual representation of the search that provided the foundation for each search strategy. The frameworks were used throughout the process to clearly define key variables and terms, help determine the inclusion and exclusion criteria, inform the development of the literature search strategy, and control the scope. For each question, Subcommittee members were asked to develop the components of the analytical framework using the PICO (Population, Intervention or Exposure, Comparison, and Outcomes) method. The analytical frameworks specified the criteria for the types of population (participants), types of interventions (and comparisons), and the types of outcomes of interest. The frameworks were discussed and refined during Subcommittee calls. In some cases, these discussions resulted in refinements to the SR questions. The development of the analytical framework was often done in conjunction with the next step in the process (developing the inclusion and exclusion criteria).

Develop Inclusion and Exclusion Criteria

SR liaisons developed a template to draft inclusion and exclusion criteria for each search to determine whether studies were eligible to be included in the SR and ensure that the evidence being considered in the SRs was relevant to the U.S. population. The template was shared with Subcommittee members for review, feedback, and approval. To promote consistency, all the SRs included four basic criteria, with additional criteria used as appropriate. The four constant criteria were:

- Publication language: Studies had to be published with full text in English.
- Publication status: Studies had to be published in peer-reviewed journals or a high-quality report identified by the Committee.
- Research type: Studies had to be existing SRs, MAs, pooled analyses, reports, or original research, determined to have appropriate suitability and quality by the Committee.
    - Existing reviews, including SRs, MAs, and pooled analyses, were considered if they met the inclusion criteria for the SR question; no priority was given to the selection of any specific type of review.
- Study subjects: Studies had to include human subjects

As appropriate, Subcommittee members considered additional criteria to identify the optimal evidence to answer each of their SR questions. These criteria related to the following:

- Age of study subjects
- Health status of study subjects
- Comparison groups included or excluded
- Date of publication
- Study design
- Intervention/exposure
- Outcome

Develop Search Strategies

A search strategy was created to identify peer-review literature for each SR conducted. Each search strategy included the following items:

- Search terms
- Boolean logic used to combine search terms
- Databases searched
- Limits: Search date range, languages searched, types of articles included (e.g., peer-reviewed articles, database-specific filters)

The search strategy also recorded the date(s) the searches were conducted and the number of articles identified with each search.

Three databases (PubMed®, CINAHL, and Cochrane) were used for each SR. These databases were identified because they represented comprehensive repositories of citations, abstracts, and full articles in fields relevant to the Committee's SRs.

The SR liaisons and librarians (from both ICF and the National Institutes of Health Library), and Subcommittee members worked together in an iterative process to develop each strategy. A list of core physical activity search terms was developed and shared with the Committee. Each Subcommittee could add or remove physical activity terms, as appropriate for each of their SR questions. Core terms were: "Aerobic activities," "Aerobic activity," "Cardiovascular activities," "Cardiovascular activity," "Endurance activities," "Endurance activity," "Exercise," "Physical activities," "Physical activity," "Physical conditioning," "Resistance training," "Sedentary lifestyle," "Strength training," "Walking," and

"Sedentary." Population- and/or health outcome-specific search terms were developed for each SR question. As appropriate, population- or health outcome-specific search terms (e.g., cancer, all-cause mortality) were shared among the SR liaisons for consistency across Subcommittees.

Once the search terms were approved by the Subcommittee members, the SR liaisons conducted a draft search to get an estimate of how many results (articles) were identified using the search strategy. If the number of results seemed unreasonable or inaccurate to the Subcommittee members based on their expertise, the SR liaisons worked with Subcommittee members to refine the search strategy to ensure that it adequately captured articles that addressed the SR question. If the Subcommittee members considered the number of results to be reasonable and accurate, the SR liaisons shared the list of articles identified through the search for Subcommittee review, feedback, and approval.

The analytical framework, inclusion and exclusion criteria, and search strategy for each SR question can be found in the question-specific evidence portfolios, and can be accessed at www.health.gov/paguidelines.

Step 3: Search, Screen, and Select Evidence to Review

Searching, screening, and selecting scientific literature was an iterative process that sought to objectively identify the most complete and relevant body of evidence to answer each SR question. Working from the analytical frameworks, search strategies, and inclusion and exclusion criteria, the SR liaisons searched, screened, and selected the scientific literature in a systematic way to provide transparent evidence for each Subcommittee's deliberations.

Identify Sources of Evidence to Answer SR Questions

Each SR question was answered using:

- Existing reviews and/or reports,
- Original research (de novo SR),or
- A combination of both existing reviews and/or reports and original research.

For each SR, existing reviews and reports were searched and screened first. These documents are valuable sources of summarized evidence that were used to prevent duplication of effort and promote efficient time and resource management. The decision to use existing reviews and/or reports, original research, or a combination of both existing reviews and/or reports and original research was made by

Subcommittee members for each SR after their review of the initial search results or the title, abstract, or full-text triage results.

Search for High-Quality Existing Reviews

Existing reviews were identified by using the search strategy, which specifically was restricted to identify only publications that were SRs, MAs, and pooled analyses. Two librarians independently reviewed the search strategies carried out by the SR liaisons to ensure quality and comprehensiveness, providing recommendations as needed. The librarians also duplicated each search to identify any errors in searching procedures and reviewed documentation of each search strategy.

After completing each search, duplicates were removed, resulting in a set of articles for triage. The list of articles identified for triage was shared with Subcommittee members, who provided review, feedback, and approval.

Search for High-Quality Existing Reports

The SR liaisons conducted a search of nine resources and websites[iii] using the search terms "physical activity," "exercise," and "sedentary" to identify and gather high-quality existing reports with potential relevance to SR questions that were not identified through the search for high-quality existing reviews. The search resulted in 1,277 titles that were reviewed for relevance independently by two SR liaisons, resulting in a pool of 195 potentially relevant reports. When discrepancies were identified, a third SR liaison reviewed the titles to help reach consensus.

The SR liaisons reviewed the list of report titles and descriptions and shared with their Subcommittee(s) any they thought might be relevant. If the Subcommittee members agreed that an existing report was relevant to a SR question, the report moved to triage.

Search for Original Research Articles

If the Subcommittee determined that a complete (de novo) SR or partial (supplemental de novo) SR was necessary because, for example, of a lack of relevant existing reviews, SR liaisons developed a strategy for a complete or partial search that was specifically tailored to the Subcommittee's needs for

[iii] Resources and websites searched to identify high-quality reports included: AHRQ Evidence Reports: http://www.guideline.gov/resources/ahrq-evidence-reports.aspx; Campbell Collaboration Library of Systematic Reviews: http://www.campbellcollaboration.org/lib/; Cochrane Library: Accessed through NIH Library; Grey Literature Report: http://www.greylit.org/; Health and Medicine Division: http://www.nationalacademies.org/hmd/Reports.aspx; National Guideline Clearinghouse: http://www.guideline.gov; NICE: http://www.evidence.nhs.uk/; Rand Corporation: Accessed through NIH Library; and World Health Organization: http://www.who.int/gho/publications/en/.

Subcommittee review. SR liaisons then implemented the approved search strategy. Librarians reviewed the search strategies to ensure quality and comprehensive nature, and the searches were duplicated to identify any errors in searching procedures.

After completing the search, duplicates were removed, resulting in the set of articles for triage. The list of articles identified for triage was shared with Subcommittee members, who provided review, feedback, and approval.

Triage Articles

Once the literature search was complete, all article titles and abstracts were independently screened, or triaged by two members of the triage team, by one triage team member and one Subcommittee member, or by two or more Subcommittee members. When discrepancies were identified, an additional screener reviewed the titles or abstracts to help reach consensus.

- Title and abstract triage: Two screeners independently reviewed each article's title, then reviewed each remaining article's abstract, to determine whether it met the criteria for inclusion in the review. The list of articles identified and the triage results were shared with Subcommittee members. Subcommittee members were asked to provide review, feedback, and approval. The triage process was conducted and recorded in the online database developed for the Committee, which recorded all triage and abstraction data.
- Full-text triage: Full text was retrieved for the remaining articles after title and abstract triage and shared with Subcommittee members. Subcommittee members conducted triage on the full-text articles and excluded articles that did not meet the inclusion criteria. In addition, during the abstraction process, abstractors identified any concerns about inclusion, which the SR liaison brought to Subcommittee members for review and final decision. Any changes to the initial triage determinations based on full-text review were updated in the online database. SR liaisons shared the final list of included and excluded articles with the associated rationale for exclusion with Subcommittee members for their review.

Conduct Supplemental Searching Activities

Subcommittee members and federal support staff were encouraged to share additional articles that may have contributed to the evidence after the search strategy was executed. Subcommittee members and staff identified these articles through their expertise and familiarity with the literature or through hand searching of included article reference lists.

- If an article was identified that met the inclusion criteria (i.e., was published during the time frame searched and used existing search terms or reasonable variations of the included search terms) but had not been captured by the search strategy, it went through article triage.
- If an article was identified that had not been captured by the search strategy and did not meet the time frame requirement, the search could be "re-opened" to allow the article and other relevant articles published since the search was conducted into the potential body of evidence for consideration. Before re-opening the search, Subcommittee members had to confirm that the article would meet the inclusion criteria, provide evidence that it would alter the conclusion statement and/or the evidence grade, and request approval from the leaders of all the Subcommittees.

Determine Sources of Evidence

After reviewing the full text of all the included existing reviews and reports, the Subcommittee members decided whether these sources of evidence could be used to answer the SR question in full, in part, or not at all.

- If the existing reviews and reports selected could be used to answer the SR question in full, the literature review team proceeded to Step 4: Abstract Data and Assess Quality and Risk of Bias.
- If the existing reviews and reports selected could be used to answer the SR question in part (i.e., in combination with a de novo SR), the literature review team proceeded to Step 4 for the selected existing reviews and reports. Concurrently, the Subcommittee members discussed which components of the SR question were not addressed by the selected existing reviews and/or reports. SR liaisons developed and implemented a search strategy to answer the remaining components of the question, as described in the Search for Original Research Articles section. The revised search strategy was shared with the Subcommittee members for feedback and approval before implementation.
- If none of the existing reviews and reports could be used to answer the SR question (or if no existing reviews and/or reports were identified by the search strategy), the SR liaison implemented a search strategy to search for original research articles.

Step 4: Abstract Data and Assess Quality and Risk of Bias

An objective data abstraction approach was used to present and summarize the characteristics of studies that addressed a SR question. The goals of data abstraction were to accurately identify and concisely describe the key elements of each study, while capturing consistent information from each article across the whole body of evidence. Abstractors were hired, trained, and certified to perform all abstracting duties, and strict quality control procedures were used throughout the abstraction process.

Conduct Abstraction Training and Quality Control

Abstractor candidates participated in a three-phase, five-week virtual training that culminated in a certification process. All abstractors were certified before abstracting articles for the Committee. The training was supported by an abstractor training manual that contained detailed instructions, definitions, reporting instructions, response options, and examples (including screen shots of the online database), as well as annotated versions of the articles used in the training. In addition to initial training sessions, the training and quality control team provided group retraining and recalibration and one-on-one consultation and training to abstractors. On an ongoing basis, the training and quality control team provided feedback and developed guidance documents (e.g., FAQs) based on frequently asked questions and common errors.

Two abstractors (referred to as a "pair") independently conducted all data abstraction tasks. Abstractors were assigned batches of articles to review in the online database. After both abstractors completed the batch, the pair reviewed their entries, discussed discrepancies, and reached agreement:

- When abstractors were able to settle discrepancies, the online database was updated to reflect the decision.
- When needed, the abstractors contacted a training and quality control team member to discuss their disagreements or gain clarification. A training and quality control team member conducted an independent review of the specific data elements where discrepancies existed and provided guidance. After a decision was reached by abstractors, the online database was updated to reflect the decision.

Concurrent with abstraction, the training and quality control team independently abstracted data for 12.5 percent (at a minimum) of existing reviews, reports, and original research and then compared their entries with those of the abstractor pair to identify discrepancies. A higher percentage of articles were reviewed by the training and quality control team when abstractors moved from abstracting SRs, MAs,

pooled analyses, or reports to abstracting original articles and when new research questions required changes in the abstraction form.

Abstract Data

Data were entered into an online database using standard abstraction items, one for existing reviews and reports and another for original research (Standard Abstraction Items – SR, MA, Pooled Analyses, and Reports and Standard Abstraction Items—Original Research). The forms were modeled after similar forms used for the 2008 Advisory Committee and the Guide to Community Preventive Services SRs, and were tailored for each SR based on input from Subcommittee members. The pair of abstractors independently read and reviewed each article, abstracted key information, and entered it into the online database, which was prepopulated with basic information about the article (e.g., citation, abstract). After all quality control processes were conducted, complete abstraction data were used to populate individual article evidence summary tables.

Assess Quality for Existing SRs, MAs, and Pooled Analyses[iv]

In addition to abstracting key information from SRs, MAs, and pooled analyses, the pair of abstractors independently assessed each existing review's quality. Quality for each SR, MA, or pooled analysis was assessed using $AMSTAR_{ExBP}$.[8] $AMSTAR_{ExBP}$, a modified version of "A Measurement Tool to Assess Systematic Reviews" (AMSTAR),[9] was used to assess the methodological quality of SRs and MAs. $AMSTAR_{ExBP}$ is an adaptation of AMSTAR that focuses on MAs that examine the effects of exercise training on blood pressure. The training and quality control team made additional revisions to adapt $AMSTAR_{ExBP}$ for the Committee (SR, MA, and Pooled Analysis Quality Assessment Using Tailored AMSTARExBP Instrument). The adaptation made by the training and quality control team for the Committee was based on a methodology improvement publication for AMSTAR.[10] The main revisions clarified reporting instructions for scoring quality items in different types of reviews and were not intended to modify the tool itself. The results of the SR, MA, and pooled analysis quality assessment were used to develop quality assessment charts and were shared with Subcommittee members for review.

[iv] If authors of a publication conducted an SR followed by an MA, the study was classified as an MA. If authors referred to a study as a pooled analysis, the publication was classified as pooled analysis, independently of being accompanied by a SR or not. Publications that consisted only of SRs, for which the authors did not also conduct a meta-analysis, were classified as an SR. Subcommittee members classified existing reviews as SRs, MAs, or pooled analyses consistent with abstractions and the evidence portfolio.

Assess Quality for Existing Reports

In addition to abstracting key information from existing reports, pairs of abstractors also independently assessed each report's quality. The literature review team developed, with feedback from the USDA NEL, a set of questions that assessed the integrity and appropriateness of the methodology, recommendations, and references in existing reports (Existing Reports Quality Assessment Instrument). The results of each reports' quality assessment were used to develop quality assessment charts and were shared with Subcommittee members for review.

Assess the Risk of Bias for Original Research

In addition to abstracting key information from each original research article, pairs of abstractors assessed each study's risk of bias. Risk of bias, or internal validity, was assessed for each original study using an adapted version of the USDA NEL Bias Assessment Tool (BAT).[11] The NEL BAT uses a domain-based evaluation to help determine whether any systematic error exists that could either over- or underestimate the study results. Selection, performance, detection, and attrition bias are addressed in the NEL BAT.

The NEL BAT is tailored by study design, with different sets of questions applying to randomized controlled trials (RCTs) (14 questions), non-randomized controlled trials (14 questions), and observational studies (12 questions). To adapt the NEL BAT for the Committee, the training and quality control team made minor revisions to expand the reporting instructions to facilitate decision making and provide examples relevant to the Committee's topics, questions, and study designs (Original Research Bias Assessment using Adapted Nutrition Evidence Library Bias Assessment Tool Instrument). The results of studies' risk of bias assessments were used to develop the risk of bias summary charts and were shared with Subcommittee members for review.

Step 5: Describe the Evidence

To facilitate the Committee's review and analysis of the evidence, the literature review team prepared evidence portfolios for each SR question. For transparency, the evidence portfolios documented the full process followed for each of the SRs, including the sources of evidence, conclusions, evidence grades, description of evidence, populations analyzed, individual evidence summary tables, risk of bias and quality assessment charts, search strategy, literature tree, references, and rationale for exclusion of articles excluded at abstract or full-text triage. After the SR liaison compiled the evidence portfolios, all

evidence portfolios and reference lists were edited and reviewed for consistency. SR liaisons submitted evidence portfolios to the corresponding Subcommittee for review, feedback, and approval.

This step was often done concurrently with Step 6: Complete Evidence Portfolio and Draft Advisory Committee Scientific Report.

The evidence portfolio for each SR question can be accessed at www.health.gov/paguidelines.

Step 6: Complete Evidence Portfolios and Draft Scientific Report

Develop Conclusion Statements

Subcommittee members reviewed and deliberated on the body of evidence (i.e., included existing reviews, original research articles included in existing reviews, and/or included original research) to develop conclusion statements that answered each of their SR questions and any subquestions. Conclusion statements were tightly associated with the evidence, focused on general agreement among the studies around the independent variable(s) and outcome(s), and acknowledged areas of disagreement or limitations, where they existed. The conclusion statement(s) reflected only the evidence reviewed and not information Subcommittee members might have known from another source.

Grade the Evidence

Along with the SR evidence portfolios, the Committee members were given a rubric, the 2018 Physical Activity Guidelines Advisory Committee Grading Criteria (Table E-1), to guide their assessment and grading of the strength of the evidence supporting each conclusion statement. The rubric was adapted from the USDA NEL Conclusion Statement Evaluation Criteria rubric[12] and revised slightly by Committee members to reflect the specific characteristics of physical activity literature. Grading the strength of the evidence was based on applicability of the populations, exposures, and outcomes studied; generalizability to the population of interest; risk of bias and study limitations; quantity and consistency of findings across studies; and magnitude and precision of effect.

Subcommittees presented their conclusion statements and strength of evidence grades to the full Committee during public meetings for deliberation and approval. When necessary, Subcommittee members revised the conclusion statements and grades. Any changes to conclusion statements and strength of evidence grades had to be re-presented to the full Committee during public meetings.

Develop Narrative Summary and Research Recommendations

After the Subcommittee members developed a conclusion statement and grade for a SR question and any SR subquestions, they developed a narrative summary of their analysis and research recommendations related to the question. The summary included a review and synthesis of the evidence, rationale for evidence grades, and limitations. The research recommendations listed key areas where additional research could enhance the evidence base by addressing gaps identified in the existing research, advancing the field of physical activity research, and informing future editions of the Physical Activity Guidelines.

Draft the PAGAC Scientific Report

Subcommittee members drafted a summary for each SR question using the body of evidence. The SR question summaries were compiled into the Committee's Scientific Report.

PAGAC EVIDENCE ASSESSMENT TOOLS

Standard Abstraction Items – SR/MA/Pooled Analyses/Reports[v]

Summary of Individual SR/MA/Pooled Analysis/Report

- Type of Review/Source
    - Systematic Review/Meta-Analysis/Pooled-Analysis
        - Total Number of Studies
    - Report
        - Report Organization/Sponsor
        - Report Type
- Purpose of the Review/Report
- Author Stated Funding Source
- Exposure Definition
    - Measures Steps?
    - Measures Bouts?
    - High Intensity Interval Training (HIIT)?
- Timeframe[vi]

[v] All items ending with a question mark have yes/no responses.

[vi] Records the years covered in the search of the SR, MA, or report. If authors searched from the earliest date available in a database (e.g., from the database's inception) it was abstracted as "inception to end date of search."

- Description of Outcomes
    - Measures Change in Fitness?
- Report's Conclusions

Study Population[vii]

- Sex
- Race/Ethnicity
- Age
- Socioeconomic Status
- Population Density
- Weight Status
- Disability Status
- Pregnancy Status
- Cancer
- Chronic Condition
- Other

Standard Abstraction Items—Original Research[viii]

Study Overview

- Purpose
- Study Design
- Do the authors refer to supplementary material or previous publications for detailed methods?
- Country
- Author Stated Funding Source
- Author Stated Sample Power
- Sample Size - Initial
- Final Sample Size
- Attrition (%)
- Was the study an intervention?
- Type of Intervention
    - Provision of Information/Education
    - Behavioral
    - Environmental
    - Policy/Legislation/Regulation
    - Laboratory-based
    - Technology
    - Other

[vii] All populations analyzed and presented in the data related to the outcome of interest are recorded.
[viii] All items ending with a question mark have yes/no responses.

- Physical Activity Exposure Assessment
    - Self-reported
    - Device-measured
    - Direct Observation
    - Other
    - Measures Steps?
    - Measures Bouts?
- Outcomes and Measurement
    - Measures Change in Fitness?
    - Addresses Adverse Events?

Study Population[ix]

- Sex
- Race/Ethnicity
- Age
- Socioeconomic Status
- Population Density/Urbanicity
- Weight Status
- Disability Status
- Pregnancy Status
- Cancer
- Chronic Condition
- Other

Intervention Components

- Length of Overall Physical Activity Intervention
- Frequency of Physical Activity
- Intensity of Physical Activity
- Duration of Physical Activity
- Physical Activity Type
    - Cardiorespiratory
    - Strength
    - Balance
    - Flexibility
    - Active Play, Free Play, or Outdoor Play
    - Other
- High Intensity Interval Training (HIIT)?
- Was Intention to Treat Analysis Conducted?

[ix] All populations analyzed/presented in the data related to the outcome of interest are recorded.

SR, MA, and Pooled Analyses Quality Assessment Using Tailored AMSTAR$_{ExBP}$ Instrument

- Were the review questions and inclusion and exclusion criteria clearly delineated prior to executing the search strategy?
- Were the population variables defined and considered in the methods?
- Was a comprehensive literature search performed?
- Was there duplicate study selection and data extraction?
- Was the search strategy clearly described?
- Was relevant grey literature included in the review?
- Was a list of studies (included and excluded) provided?
- Were the characteristics of the included studies provided?
- Was Frequency, Intensity, Time, and Type (FITT) defined for each study and examined in relation to the outcome effect sizes?
- Was the scientific quality (risk of bias) of the included studies assessed and documented?
- Did results depend on study quality, either overall, or in interaction with moderators?
- Was the scientific quality of the included studies used appropriately in formulating conclusions?
- Were the data appropriately synthesized in a qualitative manner and if applicable, was heterogeneity assessed?
- Was the effect size index chosen justified, statistically?
- Was individual-level meta-analysis used?
- Were practical recommendations clearly addressed?
- Was the likelihood of publication bias assessed?
- Was the conflict of interest disclosed?

Existing Reports Quality Assessment Instrument

- Were the review questions and inclusion and exclusion criteria clearly delineated prior to executing the search strategy?
- Did the inclusion criteria permit grey literature?
- Was a comprehensive literature search performed?
- Was the scientific quality of the included source assessed and documented?
- Are limitations reported and discussed?
- Are the conclusions substantiated by and logically connected to the evidence and findings presented?
- Was there a clear list of practical recommendations provided for future research or work on the topic?
- Are the recommendations relevant to the purpose of the report and supported by the evidence, findings, and conclusions?
- Were the potential conflicts of interest among report funders, authors, expert, or stakeholders assessed and explained?
- Was a reference list or a bibliography for the cited literature provided?

Original Research Bias Assessment using Adapted Nutrition Evidence Library Bias Assessment Tool Instrument[x]

- Were the inclusion and exclusion criteria similar across study groups?
- Was the strategy for recruiting or allocating participants similar across study groups?
- Was the allocation sequence randomly generated?
- Was the group allocation concealed (so that assignments could not be predicted)?
- Was distribution of health status, demographics, and other critical confounding factors similar across study groups at baseline? If not, does the analysis control for baseline differences between groups?
- Did the investigators account for important variations in the execution of the study from the proposed protocol or research plan?
- Was adherence to the study protocols similar across study groups?
- Did the investigators account for the impact of unintended or unplanned concurrent interventions or exposures that were differentially experienced by study groups and might bias results?
- Were participants blinded to their intervention or exposure status?
- Were investigators blinded to the intervention or exposure status of participants?
- Were outcome assessors blinded to the intervention or exposure status of participants?
- Were valid and reliable measures used consistently across all study groups to assess inclusion and exclusion criteria, interventions and exposures, outcomes, participant health benefits and harms, and confounding?
- Was the length of follow-up similar across study groups?
- In cases of high or differential loss to follow-up, was the impact assessed (e.g., through sensitivity analysis or other adjustment method)?
- Were other sources of bias taken into account in the design and/or analysis of the study (e.g., through matching, stratification, interaction terms, multivariate analysis, or other statistical adjustment such as instrumental variables)?
- Were the statistical methods used to assess the primary outcomes adequate?

[x] Item relevance depended on the study design reported.

Table E-1. 2018 Physical Activity Guidelines Advisory Committee Grading Criteria

Criteria	Strong	Moderate	Limited	Not Assignable
Applicability	Study populations, exposures, and outcomes are directly related to the question	Some of the study populations, exposures, or outcomes, are directly related to the question	Most of study populations, exposures, and outcomes relate to the question indirectly	All of the study populations, exposures, and outcomes relate to the question indirectly
Generalizability (to the U.S. population of interest)	Studied population, exposure, and outcomes are free from serious doubts about generalizability	Minor doubts about generalizability	Serious doubts about generalizability due to narrow or different study population, exposure, or outcomes studied	Highly unlikely that the studied population, exposure, and/or outcomes are generalizable to the U.S. population
Risk of bias or study limitations (as determined by NEL BAT and/or $AMSTAR_{ExBP}$)	Studies are of strong design; free from methodological concerns, bias, and execution problems	Studies are of strong design with minor methodological concerns OR studies of weaker study design	Studies of weak design OR inconclusive findings due to design flaws, bias, or execution problems	Serious design flaws, bias, or execution problems across the body of evidence
Quantity and Consistency (of the results across the available studies)	Many studies have been published and the results are highly consistent in direction and approximate size of effect	A moderate number of studies have been published with some inconsistency in direction or size of effect	Few studies have been published with some inconsistency in direction or size of effect	Findings are too disparate to synthesize OR single small study unconfirmed by other studies
Magnitude and precision of effect	The magnitude and precision of the estimated effect provide considerable confidence in the accuracy of the findings	The magnitude and precision of the estimated effect provide confidence in the accuracy of the findings	The magnitude and precision of the estimated effect provide some but not a lot of confidence in the accuracy of the findings	Magnitude and precision of effect cannot be determined

REFERENCES

1. U.S Department of Agriculture (USDA). Nutrition evidence library—about. USDA website. https://www.cnpp.usda.gov/nutrition-evidence-library-about. Accessed January 16, 2018.

2. Agency for Healthcare Research and Quality. *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*. Rockville, MD: Agency for Healthcare Research and Quality; 2014. AHRQ Publication No. 10(14)-EHC063-EF. https://effectivehealthcare.ahrq.gov/topics/cer-methods-guide/overview. Accessed January 16, 2018.

3. Higgins JP, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.1.0. The Cochrane Collaboration; 2011. http://handbook-5-1.cochrane.org. Updated March 2011. Accessed January 16, 2018.

4. Institute of Medicine. Finding what works in health care: standards for systematic reviews. March 2011. http://www.nationalacademies.org/hmd/~/media/Files/Report%20Files/2011/Finding-What-Works-in-Health-Care-Standards-for-Systematic-Reviews/Standards%20for%20Systematic%20Review%202010%20Insert.pdf. Accessed January 16, 2018.

5. Federal Trade Commission. Data Quality Act. Section 515 of the Treasury and General Government Appropriations Act for FY 2001, Pub. L. No. 106–554.

6. U.S Department of Health and Human Services. *2008 Physical Activity Guidelines for Americans*. Washington, DC: U.S Department of Health and Human Services; 2008. https://health.gov/paguidelines/guidelines. Published 2008. Accessed September 22, 2017.

7. Physical Activity Guidelines Advisory Committee. *Physical Activity Guidelines Advisory Committee Report, 2008*. Washington, DC: U.S Department of Health and Human Services; 2008. https://health.gov/paguidelines/guidelines/report.aspx. Published 2008. Accessed January 4, 2018.

8. Johnson BT, MacDonald HV, Bruneau ML Jr, et al. Methodological quality of meta-analyses on the blood pressure response to exercise: a review. *J Hypertens*. 2014;32(4):706-723. doi:10.1097/HJH.0000000000000097.

9. Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol*. Feb 2007;7:10. doi:10.1186/1471-2288-7-10.

10. Burda BU, Holmer HK, Norris SL. Limitations of A Measurement Tool to Assess Systematic Reviews (AMSTAR) and suggestions for improvement. *Syst Rev*. April 2016;5:58. doi:10.1186/s13643-016-0237-1.

11. Office of Disease Prevention and Health Promotion. Scientific Report of the 2015 Dietary Guidelines Advisory Committee. Washington, DC: U.S Department of Health and Human Services; 2015. https://health.gov/dietaryguidelines/2015-scientific-report/05-methodology.asp. Accessed January 4, 2018.

12. Office of Disease Prevention and Health Promotion. Scientific Report of the 2015 Dietary Guidelines Advisory Committee. Table C.2, NEL Grading Rubric. Washington, DC: U.S Department of Health and Human Services; 2015. https://health.gov/dietaryguidelines/2015-scientific-report/05-methodology.asp#table-anchor-c.2. Accessed January 10, 2018.