




Original Research

Evidence on Scaling in Health and Social Care: An Umbrella Review

ROBERTA DE CARVALHO CORÔA,^{*,†,‡}
AMÉDÉ GOGOVOR,^{*,†,‡} ALI BEN CHARIF,[§]
ASMA BEN HASSINE,^{*,||}
HERVÉ TCHALA VIGNON ZOMAHOUN,^{‡,#}
ROBERT K. D. MCLEAN,^{**,††} ANDREW MILAT,^{‡‡,§§}
KARINE V. PLOURDE,^{*} NATHALIE RHEAULT,^{*,†}
LUKE WOLFENDEN,^{|||,##,***}
and FRANCE LÉGARÉ ^{*,†,‡}

^{*}VITAM - Centre de recherche en santé durable, Centre intégré universitaire de santé et services sociaux de la Capitale-Nationale; [†]Unité de soutien au système de santé apprenant Québec; [‡]Faculty of Medicine, Université Laval; [§]CubecXpert; ^{||}Faculty of Nursing, Université Laval; [#]School of Physical and Occupational Therapy, Faculty of Medicine and Health Sciences, McGill University; ^{**}International Development Research Centre; ^{††}Faculty of Medicine and Health Sciences, Stellenbosch University; ^{‡‡}School of Public Health, University of Sydney; ^{§§}Centre for Epidemiology and Evidence, NSW Ministry of Health; ^{|||}School of Medicine and Public Health, University of Newcastle; ^{##}National Centre of Implementation Science, University of Newcastle; ^{***}Hunter New England Population Health

Policy Points:

- More rigorous methodologies and systematic approaches should be encouraged in the science of scaling. This will help researchers better determine the effectiveness of scaling, guide stakeholders in the scaling process, and ultimately increase the impacts of health innovations.
- The practice and the science of scaling need to expand worldwide to address complex health conditions such as noncommunicable and chronic diseases.

The Milbank Quarterly, Vol. 101, No. 3, 2023 (pp. 881-921)

© 2023 The Authors. *The Milbank Quarterly* published by Wiley Periodicals LLC on behalf of The Milbank Memorial Fund.

This is an open access article under the terms of the Creative Commons Attribution-Non-Commercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

- Although most of the scaling experiences described in the literature are occurring in the Global South, most of the authors publishing on it are based in the Global North. As the science of scaling spreads across the world with the aim of reducing health inequities, it is also essential to address the power imbalance in how we do scaling research globally.

Context: Scaling of effective innovations in health and social care is essential to increase their impact. We aimed to synthesize the evidence base on scaling and identify current knowledge gaps.

Methods: We conducted an umbrella review according to the Joanna Briggs Institute Reviewers' Manual. We included any type of review that 1) focused on scaling, 2) covered health or social care, and 3) presented a methods section. We searched MEDLINE (Ovid), Embase, PsycINFO (Ovid), CINAHL (EBSCO), Web of Science, The Cochrane Library, Sociological Abstracts (ProQuest), Academic Search Premier (EBSCO), and ProQuest Dissertations & Theses Global from their inception to August 6, 2020. We searched the gray literature using, e.g., Google and WHO-ExpandNet. We assessed methodological quality with AMSTAR2. Paired reviewers independently selected and extracted eligible reviews and assessed study quality. A narrative synthesis was performed.

Findings: Of 24,269 records, 137 unique reviews were included. The quality of the 58 systematic reviews was critically low ($n = 42$). The most frequent review type was systematic review ($n = 58$). Most reported on scaling in low- and middle-income countries ($n = 59$), whereas most first authors were from high-income countries ($n = 114$). Most reviews concerned infectious diseases ($n = 36$) or maternal–child health ($n = 28$). They mainly focused on interventions ($n = 37$), barriers and facilitators ($n = 29$), frameworks ($n = 24$), scalability ($n = 24$), and costs ($n = 14$). The WHO/ExpandNet scaling definition was the definition most frequently used ($n = 26$). Domains most reported as influencing scaling success were building scaling infrastructure (e.g., creating new service sites) and human resources (e.g., training community health care providers).

Conclusions: The evidence base on scaling is evolving rapidly as reflected by publication trends, the range of focus areas, and diversity of scaling definitions. Our study highlights knowledge gaps around methodology and research infrastructures to facilitate equitable North–South research relationships. Common

efforts are needed to ensure scaling expands the impacts of health and social innovations to broader populations.

Keywords: scaling science, scalability, diffusion of innovation, capacity building, delivery of health care, health plan implementation, evidence-based practice, review literature as topic.

THE SCALING OF HEALTH AND SOCIAL CARE INNOVATIONS IS A topic of growing importance in the contemporary world because of its potential to reduce waste and inequities in health and social settings and to expand the benefits of effective innovations.^{1–6} The emerging science of scaling seeks to provide systematic, critical, principles-based knowledge to inform and improve the scaling of health and social care.^{7,8} Since 2010, a framework proposed by ExpandNet, a World Health Organization (WHO) initiative, has been widely consulted to guide scaling of evidence-based health care innovations. It characterizes “scaling” practices as those whose goal is to increase the reach and adoption of innovations.² In the literature, “scaling” is often interpreted as identical to “spread.” However, they present nuances of method and purpose. Although “spread” usually means replicating an innovation somewhere else,⁹ the word “scaling” can also refer to the various methods used and steps required to increase the impact of an innovation,⁷ i.e., the processes used to expand its reach, improve its quality, or increase access for vulnerable groups. In addition, “vertical” scaling describes the process of using macro structures (such as policy, regulation, or financial tools) to institutionalize an innovation throughout a whole system, whereas “horizontal” scaling refers to the rolling out of an innovation from one specific setting (e.g., a pilot project) to others.^{2,5,10} Furthermore, the term “scaling” is often suffixed with “up,” “out,” or “deep.” Briefly, “scaling up” means increasing throughput to more individuals (e.g., offering the same course to more health professionals), “scaling out” refers to expanding the innovation by adapting it to new systems or populations (e.g., offering the same course on an online platform),¹¹ and “scaling deep” corresponds to improving the outcomes of the innovation by improving the structures and processes used (e.g., updating courses from an equality, diversity, and inclusion perspective).⁷ Inspired by these definitions,^{2,5,7} for the purposes of this paper we define “scaling of health and social care innovations” as a systematic evidence-informed process whose aim is to increase the positive impacts of an innovation that has proven

effective for improving quality in care and the welfare of individuals and populations. Scaling can focus on any health or social outcome, e.g., the functionality of a health system (access to essential services, quality of care, the demand for essential services, or its resilience), medication access for tuberculosis infections, COVID-19 screening, or treatment of depression. It can also focus on food security, empowerment of vulnerable communities, poverty, education, or human development.^{12–14} We define “innovations” as activities that are new as well as those that have already been successfully implemented in specific settings before being scaled to further contexts.¹⁵ In general, they are perceived as new by end users.¹⁶

There is considerable interest in ensuring that new knowledge born out of research is put to good use to improve population health.^{17–19} The fields of knowledge translation (KT) and implementation science have improved the prospects of putting research into practice. Within these more established fields, scaling science is considered complementary.^{7,8,20} Most KT and implementation models such as the Knowledge-to-Action model²¹ do not explicitly include scaling as an important and distinctive phase of the KT process.⁸ We believe greater attention to the science of scaling can improve the impact of implementation and KT efforts.

The COVID-19 pandemic increased the urgency of scaling health and social care innovations in high-income countries (HICs) as well as in lower- and middle-income countries (LMICs).^{19,22–24} For example, the pandemic hastened the scaling of a UNICEF oxygen therapy project to over 90 countries,²³ and in Australia, findings from a study of COVID-19 transmission in schools and childcare settings in the early stages of the pandemic informed widely applied policies to reduce transmission in these settings.²⁴ In addition, many policymakers believe that scaling effective health and social care innovations will reduce per capita cost of health care through economies of scale.^{25,26} Moreover, policymakers who are concerned about the rapidly increasing inequities within and among countries consider scaling as an effective means to address political and community health inequities.^{27–30} These pressing motivations to scale effective innovations are now giving rise to a new interest in the underlying mechanisms at play in scaling, and how to harness them. These mechanisms influencing an action or a system, through which an intervention works or fails to work, can be practical (e.g., human resources), social (e.g., cultural mores), or psychological (e.g., reactions or

reasonings), all of which can vary depending on the area targeted for scaling.^{31,32}

There is as of yet less evidence about scaling mechanisms or processes in health and social care than in business and agriculture. Indeed, there is evidence of a persistent failure to scale effective health innovations across health care systems around the world.^{7,33,34} This may, in part, be attributable to a lack of scientific knowledge about scaling as well as a failure to share practices across countries.^{7,34,35} The use of reviews of research evidence, and in particular, systematic reviews, is recommended to guide health policy, to support efforts to scale health and social innovations, and to identify gaps in the review evidence base. Previous studies have suggested that although systematic reviews on scaling in health and social care are increasing, they are often highly focused on a specific topic or clinical domain (e.g., HIV testing uptake) and provide incomplete and haphazard coverage of the evidence needs of stakeholders.^{34,36–40}

In this umbrella review, we therefore aimed to synthesize the evidence base for scaling in health and social care and to identify current knowledge gaps. The results will help to build and advance the emerging science of scaling.

Methods

Study Design

We conducted an umbrella review whose purpose was to synthesize the findings of available reviews. We adopted the Joanna Briggs Institute (JBI) methodology for umbrella reviews,⁴¹ and we used the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines to structure this review (Additional File 1).⁴² The study protocol was registered with the International Prospective Register of Systematic Reviews (CRD42020183774) on November 11, 2020. The present paper is the first iteration of a living umbrella review for which a protocol has been published.⁴³

Eligibility Criteria

We collected all types of evidence matching the “PICO” criteria (Participants, Intervention, Comparator, Outcome) and “PICo” (Population, phenomena of Interest and Context) to capture the evidence from

quantitative and qualitative reviews pertaining to scaling. For example, in addition to systematic reviews, we included critical reviews, scoping reviews, and narrative reviews.

Participants or Population. We considered all reviews that included primary studies focusing on individuals, health systems, services, or organizations that had been exposed to the scaling of a health or social care innovation.

There were no restrictions based on sociodemographic factors (e.g., age, ethnicity, socioeconomic status) or health conditions (e.g., comorbidities).

Interventions/Phenomena of Interest. We considered all types of interventions. This umbrella review not only explored the scaling of health and social care innovations but also any aspect of or any topic relating to such scaling, including (but not limited to) concepts, models, analytical models, tools, cost or impact assessments, and user engagement data. We considered “scaling” as inclusive of all variants of spread, scale or scaling up, scaling out or scaling deep, horizontal, and vertical. By adopting a broad view, we aimed to cast a wide net and thereby increase the learning potential of the study.⁴³

Comparator. There were no restrictions.

Outcomes. We considered all outcomes reported in the included reviews, including health outcomes. We sought any outcomes such as (but not limited to) patients/caregivers’ and/or health care providers’ and/or policymakers’ perceptions and experiences of barriers, facilitators, acceptability and feasibility of scaling innovations, impact (e.g., adaptability, efficacy, effectiveness), coverage (e.g., proportion of the target population reached by the scaling), adoption, fidelity, penetration, maintenance of the innovation, health outcomes (e.g., impact on morbidity, mortality), patient-reported outcomes (e.g., quality of life, satisfaction), and health care resources (e.g., cost effectiveness of the scaled innovation, cost of staff resources).

We excluded the review if it reported only on the impact of an effective health innovation but was not about scaling.

Types of Study. We considered all types of review (quantitative, qualitative, and mixed methods) that addressed, synthesized, or summarized preexisting evidence in the field of scaling. We define a review as a synthesis of evidence that includes a clear research question (i.e., mentions at least the population and the intervention) and describes the methods used to identify and select the primary research studies (i.e., mentions at

least one data source and keywords).^{44,45} Reviews could include studies with any research design. We included all reviews that had made explicit their methodological approach, whether they had been conducted according to their chosen design or not. We excluded primary research studies, conference abstracts, comments, opinions, letters, and editorials.

Context/Setting. We included reviews in any type of health or social care setting located in any geographical setting. We define “health and social care” as follows: Health care consists of services provided in institutional or community settings, any form of access to a health-related service (such as dental, podiatry or optical services), and access to health care practitioners (such as nurses, physiotherapists, or general practitioners).⁴⁶ Social care consists of interventions that support frail or vulnerable individuals by meeting needs or enabling them to meet needs that arise as a result of physical, mental, or emotional impairment.⁴⁷

Data Sources and Search Strategy. A systematic literature search was performed to identify published studies in the following electronic bibliographic databases: MEDLINE via Ovid, Embase, PsychINFO via Ovid, CINAHL via EBSCOhost, Web of Science, The Cochrane Library, Sociological Abstracts via ProQuest, Academic Search Premier via EBSCO, and ProQuest Dissertations & Theses Global.

Our information specialist (NR) developed an Ovid-MEDLINE and Web of Science strategy with input from the project team and a second information specialist. An iterative revision process was conducted by the members of the research team. We used the following main concepts: “scaling,” “reviews,” and “health and social care.” We improved a search strategy for scaling developed earlier by members of our team.^{33,34} We adapted the Canadian Agency for Drugs and Technologies in Health (CADTH) search filter for Systematic Reviews/Meta-Analysis/Health Technology Assessment for this project.⁴⁸ A second information specialist reviewed the search strategy using the Peer Review of Electronic Search Strategies tool.⁴⁹ The final version was approved by the team members and then translated into the other databases. Details on the search strategies are presented in Additional File 2. No language restriction was applied. We searched from inception onwards until August 6, 2020. Articles not available electronically were ordered via interlibrary loan.

Additionally, to identify gray literature, we searched Google and the websites of the following organizations: WHO (WHO-ExpandNet), the UK’s National Institute for Care and Excellence, NSW

Government (Health Australia), the Canadian Foundation for Healthcare Improvement (CFHI), the International Development Research Centre, the CADTH's Grey Matters checklist, the Agency for Healthcare Research and Quality, the Global Community of Practice on Scaling Development Outcomes, the Institute for Healthcare Improvement, the National Implementation Research Network, the State Implementation and Scaling-up of Evidence-based Practices Center, the Health Foundation, the Australian Prevention Partnership Center, and Scaling Pathways. For each term, we reviewed the first 100 reports, except in ExpandNet and the CFHI, for which we searched all the reports in their websites (Additional File 3). The search took place on July 20, 2021.

Selection Process. The search results were imported and stored in an EndNote X9 library for reference management and duplicate removal.^{50,51} The resulting records were exported to the Internet-based system Covidence for the selection process.⁵² Pilot screening of the titles and abstracts of 2.5% randomly sampled reviews was completed to ensure common understanding of the selection criteria (1: Is it a review? 2: Does its aim concern scaling or inform scaling? 3: Is it about health or social care?). Paired independent reviewers screened titles and abstracts against the eligibility criteria. When abstracts did not appear to meet the criteria for exclusion, were ambiguous, or were missing, the studies were retained and reviewed in full. The reviewers' full understanding of the selection criteria was validated again before beginning the full-text selection. The full-text eligibility assessment was done by paired reviewers independently (ABH, KVP, and RCC). In case of an "unclear response" regarding the eligibility of studies, authors resolved disagreements through discussion and, if necessary, consulted a third senior author (FL or AG). Any reasons for exclusion were recorded in Excel.

Data Extraction. We developed a data extraction form based on the JBI form for review of systematic reviews. A pilot data extraction was conducted on five of the selected reviews to ensure correct comprehension of the grid. Paired independent reviewers conducted the data extraction (ABH and RCC). Any disagreements between the reviewers were resolved through discussion; if consensus was not reached, they consulted a third senior reviewer (KVP). We did not extract data from the primary studies included in the reviews. We summarized the review findings but not the results of primary studies included in the reviews. If the article required translation into English to enable data collection,

we used DeepL Translator software.⁵³ Extracted information included the following:

- Review characteristics: journal name; first author; language; year of publication; type of review as reported by authors (e.g., systematic, scoping, rapid, realist, literature review); source of funding; country of the corresponding author; aim of the study; country or region of interest (e.g., South Africa, Canada, sub-Saharan Africa, LMICs); PICO or PICo; inclusion and exclusion criteria; diseases, conditions, or types of care of interest (e.g., HIV, malaria, mother and child health, cancer, diabetes); scaling terms used in the search (e.g., scal*, spread, large-scale, diffusion, scalability); number and type of databases sourced and searched (e.g., MEDLINE, CINAHL, Google, WHO website); date range of database searching; whether quality of the primary studies was appraised; appraisal instruments used to assess their quality (e.g., Grading of Recommendations Assessment, Development and Evaluation; Critical Appraisal Skills Programme; AMSTAR2, A Measurement Tool to Assess systematic Reviews 2; Mixed Methods Appraisal Tool); results summary in the abstract; review results (e.g., number of studies included, study design of primary studies included, number and characteristics of participants such as sex as reported, age); heterogeneity if applicable.
- Focus area of the reviews: The aim of the review and the results, if reported (yes/no), were organized based on previous reviews into the following focus areas: interventions,³⁴ barriers or facilitators,³⁹ frameworks or concepts,⁵⁴ scalability,⁵⁵ measurement,⁵⁶ infrastructure⁵⁷, economic evaluation or costs,⁵⁸ and other areas (see Additional File 4 for complete list and definitions).
- Scaling characteristics, in all reviews: How authors defined “scaling” (the source and the citation); mention of “sustain” or “sustainability” (yes/no); tool development (yes/no; if yes, tool name); mention frameworks related to scaling (yes/no), whether used in the methods or developed (yes/no; if yes, framework name).

For reviews focusing on the three most frequent areas, we extracted further data:

- Focus area: barriers and facilitators. Reviews that reported results on barriers and facilitators related to the following scaling domains: C1) health care infrastructure (e.g., providing medical equipment or changing procedures within a health system), C2) policy and regulation (e.g., revising policy to allow widespread community-based case management of a disease), C3) financing (e.g., changing payment mechanisms, funding), C4) human resources (e.g., training and deployment of health care providers, changing roles of administrators), and C5) patient or public engagement (e.g., involving patients/public in recruitment or promotion). Scaling domains C1 to C5 were based on Ben Charif's five components of scaling strategies.³⁴
- Focus area: interventions. Reviews that reported intervention results (yes/no) also related to scaling domains C1 to C5, i.e., health care infrastructure, policy and regulation, financing, human resources, and patient or public engagement,³⁴ as well as relating to scaling acceptability, adoption, appropriateness, coverage, cost, feasibility, fidelity, and sustainability.⁵⁹
- Focus area: scalability. Reviews that reported results on scalability also related to scaling domains C1 to C5, i.e., health care infrastructure, policy and regulation, financing, human resources, and patient or public engagement.³⁴

Assessing Quality of Reviews. Two independent reviewers used AMSTAR2⁶⁰ to assess the methodological quality of all included reviews (ABH and RCC). Disagreements between the reviewers over particular studies were resolved by discussion, with involvement of a third author when necessary (KVP). The AMSTAR2 is a validated, widely used instrument for assessing the quality of systematic reviews of both randomized and nonrandomized trials of health interventions. It consists of a set of 16 questions about characteristics of systematic reviews, e.g., components of PICO, information on studies that were excluded, and consideration of risk of bias and how it was handled in the interpretation of results. The evaluator answers “yes” or “no” for the presence of each characteristic. A “partial yes” is allowed in some instances (Additional File 5). We adapted criteria 1 (about the components of PICO) and 8 (about describing the included studies) by adding the options “yes*” and “partial yes*” when all the elements were present except for the comparator. AMSTAR2 was not designed to generate an overall score, but

an overall ranking can be performed with the seven most critical criteria (in bold in Additional File 5). We did not include scores for criteria 11 and 15 because they referred to meta-analyses, which were performed in only two reviews. We used AMSTAR2 with all review types because we consider that methodological items 1 to 10 should be respected even in nonsystematic studies (e.g., realist reviews, narrative reviews). However, no overall score was generated for nonsystematic reviews. For systematic reviews, a score of “high” was attributed when none or only one noncritical criterion was unmet, “moderate” when more than one noncritical criteria was unmet, “low” when there was one critical flaw with or without unmet noncritical criteria, and “critically low” when there was more than one critical flaw. Reviews were not excluded based on their quality assessment.

Data Synthesis

We used the 2020 PRISMA flow diagram to describe the process of study selection.⁴² We used tables to describe detailed characteristics of included studies (e.g., first author’s name, year of publication, study aim, method details), analysis details (e.g., results summary, results details, whether authors reported heterogeneity). The data were charted using frequencies for year of publication, country of corresponding author, sources of funding, language, review type, and scaling results. In our tables we synthesized the following data as “yes,” “no,” or “not reported”: authors stated a PICO, the search used the term “scaling,” timeframe for search was indicated, appraisal included study quality and whether they used quality or risk appraisal instruments, and all scaling characteristics. For the funding sources, we subclassified governmental funding into two groups. “Governmental research councils” included all organizations specialized in funding research, e.g., national research councils such as the Canadian Institutes of Health Research and the US National Institutes of Health, whereas “Governmental/intergovernmental organizations” included government bodies or ministries whose role included funding research among other activities, and government or international development agencies (e.g., USAID, WHO). We classified the targeted countries within the reviews as well as the countries of corresponding authors (using authors’ stated affiliation) as HICs or LMICs based on the Organisation for Economic

Co-operation and Development income levels.⁶¹ For the diseases or conditions, we classified under “HIV/AIDS” all reviews related to HIV or AIDS, whereas reviews related to malaria, tuberculosis, or influenza were classified as “other infectious diseases.” We classified all reviews related to maternal health, newborn or child health, pregnancy, breastfeeding, and family planning as “maternal-child health.” Regarding the sources of documents searched, we recorded the number of databases searched and whether gray literature was searched or not.

We classified reviews into mutually exclusive focus areas based on their aims, i.e., reviews of barriers and facilitators, reviews of economic evaluation or scaling cost, reviews of frameworks, reviews of infrastructure, reviews of scaling interventions, reviews of measures, reviews of scalability, and emerging focus areas. The two emerging focus areas were bibliometric analysis and ethical, legal, and societal issues (ELSI). If the aims related to multiple focus areas, the final classification was chosen based on the focus area most discussed in the review.

No meta-analysis was performed, given that pooling the results of qualitative and quantitative reviews can introduce significant overlap and bias. Because our purpose was to synthesize evidence about scaling in health and social care in the literature, we present the whole body of knowledge and include the results of all the reviews, regardless of any overlap across primary studies.⁶²

The deviation from the protocol⁴³ was that we did not search clinical trial registries or the reference list of each included review, and we did not contact experts in the field to complete data collection as additional data sources. Given the fact that we excluded most of the gray literature for not qualifying as “reviews” according to our eligibility criteria, we judged it would have been only a marginal gain to contact experts for additional data sources.

Results

Review Search and Selection

Our electronic search identified 21,575 potentially relevant records. Of these, 120 reports met the review criteria, including 114 unique reviews and 6 associated reports. The gray literature search identified 2,694 reports. Thirty-two were not available for full-text examination; 23 unique

Systematic Reviews ^a	<i>n</i> = 58	%
High	2	3.4%
Moderate	0	0.0%
Low	14	24.1%
Critically low	42	72.4%

^aThis group includes the following types of review: systematic review, systematic review + meta-analysis, hermeneutic systematic review + case study, and qualitative systematic review.

reviews (plus 3 associated reports) were included (Figure 1). A reference list of included reviews is shown in Additional File 6. A list of excluded reports with the reason for exclusion is shown in Additional File 7.

Methodological Quality Assessment

The AMSTAR2 overall score was “critically low” for most of the systematic reviews (72%, *n* = 42/58) (Table 1, Additional File 5). Two scored “high” (one focusing on interventions, the other on costs) but did not respect the noncritical criterion (10) of reporting on sources of funding of included studies. Fourteen studies received the score “low,” 13 for not providing a list of excluded studies or reasons for their exclusion (critical criterion 7) and 1 for not using a satisfactory technique for assessing the risk of bias (critical criterion 9). Considering all types of reviews (*n* = 137), the criteria most respected were explaining inclusion criteria, describing included studies in adequate detail, and reporting conflict of interest. The least respected criteria were explicitly stating that the review methods were established before conducting the review and reporting funding of included studies. In the systematic reviews, the least respected criteria were assessing the risk of bias and taking it into account when interpreting results.

Review Characteristics

An overview of the 137 included reviews can be found in Table 2. Details of each review are presented in Additional Files 6 and 7. Except for one

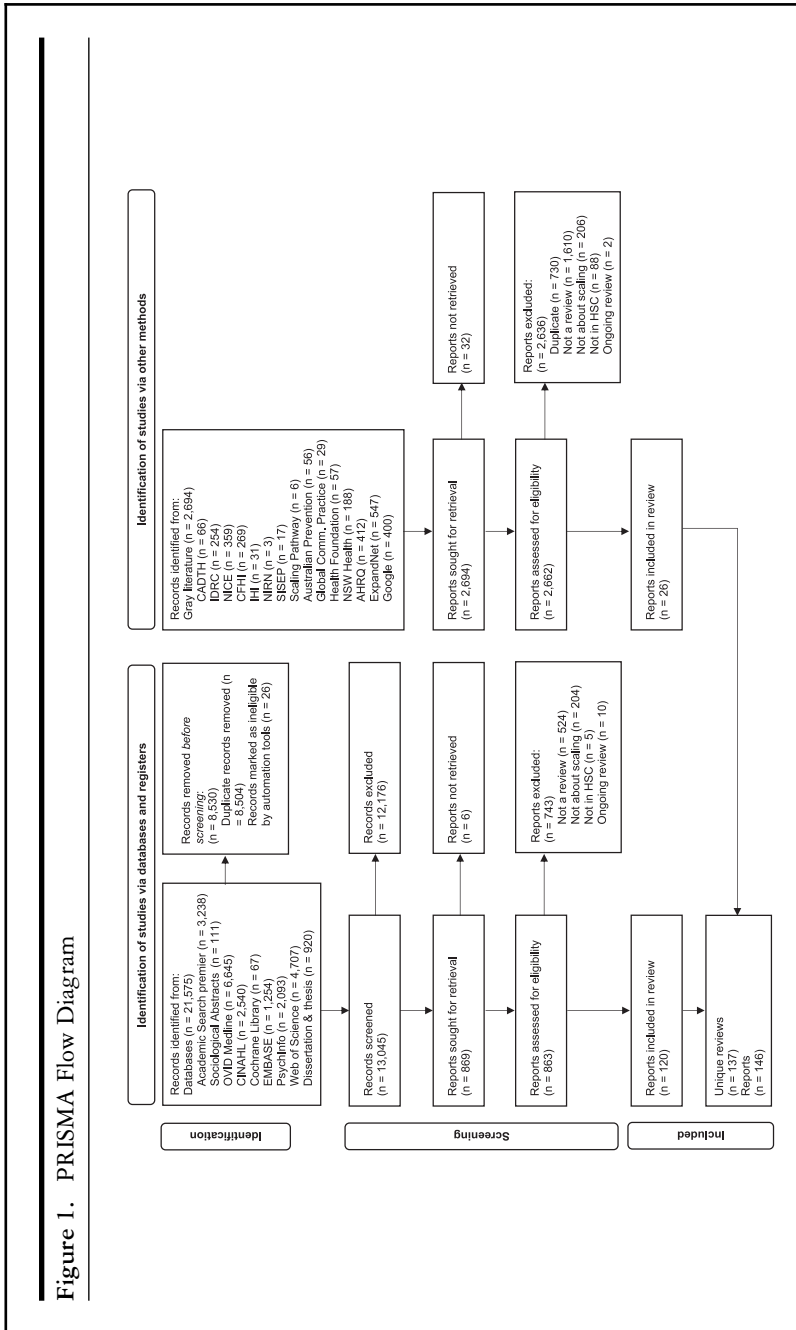


Table 2. Review Characteristics

	<i>n</i> = 137	%
Language of the publication		
English	136	99.3%
German	1	0.7%
Year of publication		
2005-2009		
2005	2	1.5%
2006	3	2.2%
2007	2	1.5%
2008	3	2.2%
2009	1	0.7%
2010-2014		
2010	9	6.6%
2011	4	2.9%
2012	8	5.8%
2013	10	7.3%
2014	11	8.0%
2015-2019		
2015	10	7.3%
2016	11	8.0%
2017	5	3.6%
2018	17	12.4%
2019	22	16.1%
2020-2021 ^a		
2020 ^a	14	10.2%
2021 ^a	5	3.6%
Review type ^b		
Group of systematic reviews		
Systematic review	54	39.4%
Systematic review + meta-analysis	2	1.5%
Qualitative systematic review	1	0.7%
Hermeneutic systematic review + case study	1	0.7%
Group of nonsystematic reviews		
Literature review	29	21.2%
Scoping review	9	6.6%
Rapid review	6	4.4%
Narrative review	5	3.6%
Comprehensive review	4	2.9%
Integrative review	2	1.5%
Qualitative review	2	1.5%

Continued

Table 2. (Continued)

	<i>n</i> = 137	%
Realist review	2	1.5%
Structured review	2	1.5%
Critical interpretive synthesis	1	0.7%
Critical review	1	0.7%
Qualitative interpretive review	1	0.7%
Synthetic review	1	0.7%
Review + interviews or Delphi	14	10.2%
Sources of funding of the review ^c		
Governmental/intergovernmental organizations	31	22.6%
Governmental research councils	28	20.4%
Charitable trusts	21	15.3%
Commercial organizations	1	0.7%
Mixed sources	12	8.8%
No source of funding reported	44	32.1%
Country of the corresponding author		
HICs ^d	114	83.2%
United States	43	31.4%
United Kingdom	25	18.2%
Canada	14	10.2%
Australia	13	9.5%
The Netherlands	6	4.4%
Belgium	3	2.2%
Switzerland	3	2.2%
Denmark	1	0.7%
Germany	1	0.7%
Ireland	1	0.7%
Italy	1	0.7%
Scotland	1	0.7%
Sweden	1	0.7%
Switzerland and United States	1	0.7%
LMICs ^d	22	16.1%
South Africa	9	6.6%
India	4	2.9%
Bangladesh	1	0.7%
Brazil	1	0.7%
Congo	1	0.7%
Kenya	1	0.7%

Continued

Table 2. (Continued)

	<i>n</i> = 137	%
Nigeria	1	0.7%
Pakistan	1	0.7%
Senegal	1	0.7%
Tanzania	1	0.7%
Zambia	1	0.7%
LMIC + HIC collaborations ^d	1	0.7%
China and Australia	1	0.7%
Income levels of targeted countries ^d		
LMIC	59	43.1%
HIC	11	8.0%
LMIC + HIC	2	1.5%
Not specified	65	47.4%
The authors stated the elements of PICO ^c		
Yes	16	11.7%
No	121	88.3%
Diseases, conditions, or care types of interest		
Health care		
Maternal–child health	28	20.4%
HIV/AIDS	25	18.2%
Other infectious diseases	11	8.0%
Mental health	4	2.9%
Cancer	2	1.5%
Chronic diseases	2	1.5%
Eye care	1	0.7%
Palliative care	1	0.7%
Social care and prevention		
Physical activity	4	2.9%
Sexual education	4	2.9%
Violence	2	1.5%
Nutrition	6	4.4%
Reviews not focused on a particular disease		
Health innovation regardless of disease	29	21.2%
Improvement the health system ^f	18	13.1%
Search strategy used scaling-related terms ^f		
Yes	50	36.5%
No	64	46.7%
Not reported	23	16.8%
Timeframe for the search ^g		
Yes	105	76.6%

Continued

Table 2. (Continued)

	<i>n</i> = 137	%
Not reported	32	23.4%
Number of databases searched ^h		
> 1	94	68.6%
One database	28	20.4%
Not specified	5	3.6%
Not reported	10	7.3%
Gray literature		
Yes	103	75.2%
Not reported	34	24.8%
Quality of the studies was appraised		
Yes	40	29.2%
Not reported	97	70.8%

Abbreviations: HICs, high-income countries; LMICs, low- and middle-income countries; PICO, Participants, Intervention, Comparator, Outcome; WHO, World Health Organization.

^aYear incomplete, searched in databases August 6, 2020 and in gray literature July 20, 2021.

^b"Review type": as described by the authors of their review; we did not verify whether the reviews met usual review characteristics of each review typology. "Literature review": any review, review of published and gray literature, or evidence review.

^c"Governmental/intergovernmental organization": government (e.g., Dutch government, health ministry), government or international development agencies, European organizations, USAID, WHO, World Bank, Global Fund. "Governmental research councils": university grants, student grants, national research councils, Canadian Institutes of Health Research, National Institutes of Health, National Institute for Health and Care Research, National Health and Medical Research Council. "Charitable trust": charitable foundations such as Grand Challenges Canada and Save the Children. "Mixed sources": studies with more than one source of funding. "No source of funding": studies did not report on funding sources or clearly indicated no grants or funding were received for the review.

^dIncome country groups: based on the Organisation for Economic Co-operation and Development Development Assistance Committee official list for development assistance⁵⁶ and on authors' reporting of the countries they included or would target.

^ePICO models: We also included PICo: Population, phenomena of Interest, and Context.

^fAuthors stated, either in the Methods or in an appendix, that they had used the following search terms, with or without the addition of "up," "out," or "down": scale, scaling, spread, rollout, diffusion, large-scale, system-wide dissemination, and scalability. "Not reported": authors did not report search strategy or none of these terms were used.

^g"Not reported": author did not state the date range of their search.

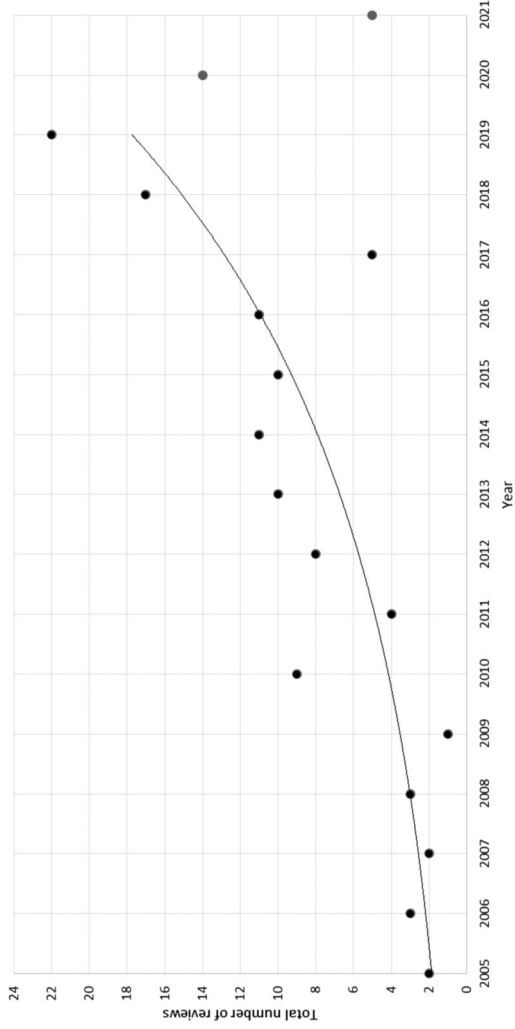
^hThe authors stated the search strategies used which databases or journal; "Not specified": authors stated they would search peer-reviewed journals or databases without mentioning their names; "Not reported": used when any information about the databases or journal searched was found in the review.

review, published in German, all reviews were published in English between 2005 and 2021, with a constant increase in publication frequency per year (Figure 2). The most frequent review type was systematic review ($n = 58$). The funding source was most commonly a governmental/intergovernmental organization ($n = 31$), followed by governmental research councils ($n = 28$). However, the majority of the reviews did not report a source of funding ($n = 44$). Corresponding authors were mostly from the United States ($n = 43$), the United Kingdom ($n = 25$), and Canada ($n = 14$). The majority of the reviews did not specify a country or a region of interest ($n = 65$) in their search methods; when one was specified ($n = 72$), they were mostly LMICs ($n = 59$). The corresponding authors of 71.2% of these 59 reviews focusing on LMICs were situated in HICs. The majority of authors did not state a clear PICO model ($n = 121$) but at least indicated inclusion criteria ($n = 130$) or exclusion criteria ($n = 80$). Most of the reviews were interested in scaling related to infectious diseases such as HIV/AIDS ($n = 25$); malaria, tuberculosis, or influenza ($n = 11$); or maternal–child health ($n = 28$). A significant number of reviews did not focus on a particular disease or condition ($n = 47$). The majority did not use scaling terms in their strategy search ($n = 64$). Most authors searched more than one database ($n = 94$) and the gray literature ($n = 103$). Most authors stated the timeframe of their search ($n = 105$). Few reviews assessed the quality of the studies they reviewed ($n = 40$). Most reviews reported data on study design ($n = 90$), but few reported on the number of study participants ($n = 29$), their age ($n = 18$), or their sex ($n = 30$).

Scaling Characteristics

The 137 reviews focused on seven areas: scaling interventions ($n = 37$), barriers and facilitators ($n = 29$), frameworks ($n = 24$), scalability ($n = 24$), cost ($n = 14$), measures ($n = 5$), and infrastructure ($n = 2$). Two additional focus areas added to our a priori list were bibliometric analysis ($n = 1$), and ELSI ($n = 1$) (Figure 3, Additional File 6). Only 36.5% ($n = 50/137$) of the authors defined scaling (Table 3). The WHO definition was the most frequently used ($n = 26$) (Additional File 8). Overall, most of the authors used the term “sustain” or “sustainability” ($n = 104$) (Table 3). In four reviews, a tool related to scaling was developed (a formula to calculate incremental coverage,⁶³ the Becoming Breastfeeding

Figure 2. Distribution of Reviews per Year



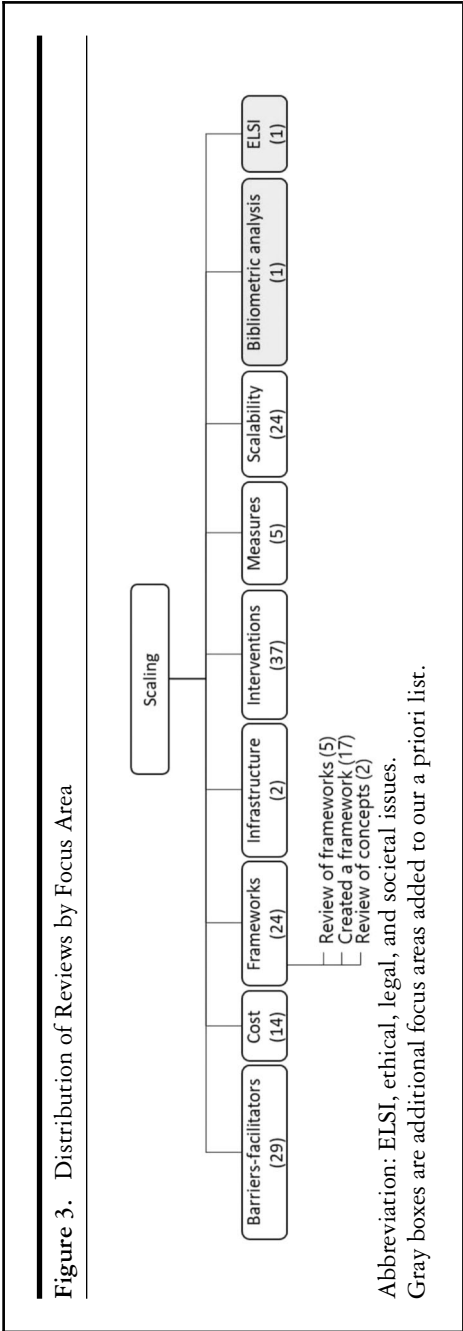


Table 3. Results Related to Scaling		
	<i>n</i> = 137	%
"Scaling" was defined		
Yes	50	36.5%
No	87	63.5%
Mention of sustainability ^a		
Yes	104	75.9%
No	33	24.1%
Developed a tool		
Yes	4	2.9%
No	133	97.1%

^aThe word "sustain" or "sustainability" was used.

Friendly Index,⁶⁴ the Intervention Scalability Assessment Tool,⁶⁵ and Scale-up Readiness Benchmarks⁶⁶) (Additional File 8). Of the 137 reviews, 73 cited at least one framework related to scaling, and 34 reviews developed their own framework or model (Table 4). A scaling framework was used in the methods of 30 reviews, with the ExpandNet framework the most frequent ($n = 7$).

We explored some focus areas in more detail (Additional File 8). Under the focus area "frameworks" ($n = 24$), 5 reviews aimed to review existing scaling frameworks, 17 created a new framework, and 2 reviewed concepts of scaling (Figure 3). Under the focus area "barriers and facilitators" ($n = 29$), 27 reviews reported barriers or facilitators associated with health care infrastructure (C1) or human resources (C4). The least frequent barrier or facilitator reported was in the domain of policy and regulation (C2) ($n = 23$) (Table 5). Under the focus area "interventions" ($n = 37$), all reviews reported on the scaling domain of health care infrastructure (C1), and the least commonly reported was in the domain of policy and regulation (C2) ($n = 22$). Most reviews focusing on "interventions" reported on two outcomes: coverage ($n = 16$) and adoption ($n = 10$) (Table 6). Of the 24 reviews focusing on "scalability," 23 reviews reported domains related to health care infrastructure, and the least frequently reported were domains related to policy and regulation ($n = 14$) (Table 5).

Table 4. Frameworks Related to Scaling

	Mentioned		Used in Methods		Developed	
	<i>n</i> = 137	%	<i>n</i> = 137	%	<i>n</i> = 137	%
Yes						
In the focus area—Barriers/facilitators	73	53.3%	30	21.9%	34	24.8%
In the focus area—Costs	14	10.2%	10	7.3%	2	1.5%
In the focus area—Frameworks	3	2.2%	0	0.7%	2	1.5%
In the focus area—Infrastructure	24	17.5%	5	3.6%	20	14.6%
In the focus area—Interventions	2	1.5%	0	0%	1	0.7%
In the focus area—Measures	14	10.2%	7	5.1%	6	4.4%
In the focus area—Scalability	4	2.9%	1	0.7%	0	0%
No	12	8.8%	7	5.1%	3	2.2%
	64	46.7%	107	78.1%	103	75.2%

Table 5. Scaling Domains Reported

	C1: Health Care Infrastructure ^a	C2: Policy and Regulation ^a	C3: Financing ^a	C4: Human Resources ^a	C5: Patient/ Public Engagement ^a
In the focus area—Barriers and facilitators (<i>n</i> = 29)					
Yes	27 (93.1%)	23 (79.3%)	26 (89.7%)	27 (93.1%)	24 (82.8%)
No	2 (6.9%)	6 (20.7%)	3 (10.3%)	2 (6.9%)	5 (17.2%)
In the focus area—Interventions (<i>n</i> = 37)					
Yes	37 (100.0%)	22 (59.5%)	23 (62.2%)	36 (97.3%)	27 (73.0%)
No	0 (0.0%)	15 (40.5%)	14 (37.8%)	1 (2.7%)	10 (27.0%)
In the focus area—Scalability (<i>n</i> = 24)					
Yes	23 (95.8%)	14 (58.3%)	15 (62.5%)	20 (83.3%)	17 (70.8%)
No	1 (4.2%)	10 (41.7%)	9 (37.5%)	4 (16.7%)	7 (29.2%)

^a As defined in Ben Charif and colleagues.²⁹

Table 6. Scaling Outcomes Reported

	Coverage	Cost	Feasibility	Acceptability	Fidelity	Adoption	Sustainability	Appropriateness
Yes	16 43.2%	7 18.9%	6 16.2%	7 18.9%	1 2.7%	10 27.0%	3 8.1%	2 5.4%
No	21 56.8%	30 81.1%	31 83.8%	30 81.1%	36 97.3%	27 73.0%	34 91.9%	35 94.6%

Discussion

We used an umbrella review to synthesize the evidence base for scaling in health and social care and found a total of 137 eligible reviews, among which 23 were from the gray literature. The vast majority of included systematic reviews scored critically low in terms of quality assessment, and many of the other review types did not respect the assessment criteria we considered important for nonsystematic reviews. We observed an increasing annual publication rate. The most frequent types of review were systematic reviews, and only two were meta-analyses. Nearly one-third of the reviews did not identify their source of funding. Among the 93 reviews that did report their funding, most were funded by governmental/intergovernmental sources. Although most first authors were from or based in HICs at the time of publication, most reviews were focused on scaling in LMICs. The most frequently covered clinical topic was infectious diseases, followed by maternal–child health, two important health issues in LMICs. Reviews focused on nine areas, most frequently on interventions, barriers and facilitators, frameworks, and scalability, whereas fewer reviews focused on measures, costs, or infrastructures. The most frequently used definition and framework were, respectively, the WHO definition of scale and its ExpandNet framework. Finally, within the three most frequent focus areas (barriers and facilitators, interventions, and scalability), the scaling domains of health care infrastructure and human resources were the most frequently reported on, whereas those related to policy and regulation were the least reported. Among reviews focusing on interventions, the outcomes most often reported were coverage, adoption, cost, and acceptability. These results lead us to make the following observations.

First, our study highlights the difficulty of using reviews (including this umbrella review) to draw conclusions about the impacts of scaling. The very low quality of the reviews and the heterogeneity of the scaled innovations may partly explain this. Even for the authors who claimed they were performing systematic reviews, the AMSTAR2 evaluation showed they did not respect the main requirements for this type of review. Combined with the small number of reviews that evaluated the quality of their included original studies, this failure to respect design modalities increased the risk of bias in our interpretation of their results. Defining scaling indicators or the use of standard implementation indicators would help identify common elements in scaling

innovations and make meta-analysis possible. This gap in methodology can be solved by being more systematic in reporting research results and by following reporting guidelines.²⁰ Also, it indicates the need to develop and advance a more systematic and rigorous science of scaling, i.e., study it, publish on it, and publish better.

Second, increasing publication rates since 2005 can be explained by the relative novelty of the science and practice of scaling in the health and social care sector, especially in HICs.¹ Even before the COVID-19 pandemic, most health care systems across the world were struggling with increasing health-related costs,^{67–69} higher consumer-driven demands on health care systems,^{70–72} and a reduced health care workforce.^{73,74} More and more health care systems are now looking for new ways to address these problems. One method, well known in LMICs, sometimes by another name (e.g., replication or expansion), is the scaling of effective evidence-based innovations. In HICs, research funding has historically focused more on pilot projects and less on scaling, but this is changing as governments seek to reduce waste through economies of scale.⁷⁵ Consequently, there has been an increase in government-funded research into the science and practice of scaling in HICs^{76–79} and hence more publications.

Third, as previous reviews have observed,^{33,34,55,80} the vast majority of reviews that reported their region of interest ($n = 72$) focused on scaling initiatives in LMICs ($n = 59$). Although the cause of this finding was not clear from our results (we propose several hypotheses below), it suggests significant potential for South to North capacity strengthening, learning, and education. Why might the majority of scaling reviews focus on the Global South? LMICs need to deliver care at a national level under severe financial constraints and, as reflected in our findings, often use scaling to address their most urgent health challenges, especially infectious disease epidemics (e.g., HIV/AIDS, malaria, tuberculosis, or influenza) as well as health issues related to social inequalities (e.g., mother and child mortality). In addition, international agencies contribute expertise and funding to scaling in LMICs to tackle humanitarian crises and improve health indicators, partly with the goal of reducing global health threats.⁷ Infectious disease programs in LMICs, such as providing access to antiretrovirals for HIV/AIDS or vaccinations for COVID-19, may be faster to scale than chronic disease programs, and their effects may be more immediate and obvious than the effects of scaling programs for chronic diseases; as a result,

they are published more frequently. Although scaling experience and expertise from LMICs are invaluable for HICs, the scaling challenges are different. However, LMICs are now also facing increasing challenges related to chronic illness. Scaling research that addresses a broader range of communicable and noncommunicable diseases is warranted.

Fourth, whereas almost half the reviews in our umbrella review focused on scaling in LMICs, two-thirds of the first authors of these reviews were from HICs or based in HICs. Innovative LMIC scaling practices can become a source of learning for scaling in HIC contexts.⁸¹ Authors located in HICs are clearly benefiting from the knowledge of LMIC experts to design, implement, and guide these scaling programs, as well as to gain funding and publish their findings.^{82–84} These results highlight the need to fully acknowledge the important LMIC contributions to the field of scaling and to support the expansion of scaling research infrastructure in LMICs. Scaling strategies should be considered global knowledge, and an equitable relationship in South–North exchanges may help advance the evidence base as well as give due credit to the rich knowledge and experience of LMICs in scaling. In this regard, a group of researchers and innovators from across the Global South has recently issued a call to action to enhance visibility of Global South contributions to global communities working on scaling.⁸⁵

Fifth, reviews covered all our a priori focus areas. Because scaling is a relatively new science, we expected to mostly find reviews about scaling interventions and reviews that analyzed scaling barriers, facilitators, and success factors. Costs and measures were the focus of fewer scaling reviews. These are significant knowledge gaps because both are elements valued by planners and policymakers when deciding whether to start scaling an innovation or not. Also, as previously observed,^{5,34,86} there are few practical tools available for scaling preparation and execution. Gray, for example, reported that scaling programs used implementation indicators to measure progress or coverage, rather than scaling indicators.⁸⁷ Another gap concerned equity. Only one review focused exclusively on ELSIs, showing that values-related considerations in decision making about scaling, such as equity of care, community needs, and population participation, have been slow to make an appearance in the field.^{7,8,30,88} Moral as well as technical justifications must be provided before, during, and after scaling is undertaken.^{5,7,15} Updated conceptual frameworks for scaling should begin with moral justification as an essential pillar of scaling.⁷ For example, in vaccine distribution, scaling considerations should include planning licensing

schemes that ensure fairness and equity of access; otherwise, scaling can simply reinforce existing inequities. The evaluation of equity impacts must be dynamic and ongoing so that innovation characteristics can be adapted if necessary and unpredicted obstacles addressed in all scaling phases.⁸⁹ In addition, health care systems are increasingly adopting the Quintuple Aim for Health Care Improvement, which maintains that all improvement and innovation efforts should focus on the individuals and communities who need them the most⁹⁰ and moreover should involve them in the entire scaling process.⁹¹

The novelty of the science also explains the lack of clarity about the definition of scaling and the frequent confusion between scaling and implementation or sustainability. Our finding that the WHO definition and its ExpandNet scaling framework were used to guide many reviews confirms that its influence is considerable in scaling innovations. However, a variety of definitions continue to be used in the broader field of scaling science. We suggest mixed interpretation of this finding. On the one hand, efforts to attain consensus on key terminology may help to clarify, unite, and advance the science of scaling. On the other hand, a diversity of scaling definitions may encourage diverse participation, cross-fertilization, and qualitative richness for the field.

Finally, among reviews focusing on barriers and facilitators, scaling interventions, or scalability (the three most frequent focus areas), the factors or strategies most frequently reported on were in the domains of health and social care infrastructure and human resources. This body of evidence on establishing the appropriate infrastructures and best practices in capacity building could be indispensable to scaling researchers and practitioners seeking strategies for ensuring success. Examples of success factors included developing a scaling team within the organization to coordinate scaling with all the stakeholders,^{2,10,92,93} putting in place beforehand all the infrastructure needed to realize the scaling,^{33,39} monitoring the scaling process with indicators related to program objectives,^{64,94,95} and training health workers to improve their ability to use the innovation.^{57,96} Since our previous review,^{33,34} we observed an increase in the number of reviews that mentioned engaging the community and public and private partners as a scaling strategy or facilitator. Policy and regulation was the strategy the least reported, indicating a knowledge gap on this subject and highlighting the need for further reflection on the relation between scaling and macro systems. This gap possibly exists because it is more difficult for researchers to act on this strategy. It is important for scaling practitioners and researchers

to know how to negotiate with macro policymakers, and they need evidence on how to face legal and/or governmental constraints in adapting their scaling interventions.⁸⁹ However, although the issue of change management and governance is becoming more prominent, this gap could also exist because in fact the main bottlenecks in scaling are not at the policy or regulatory level but must be faced locally at the level of building the necessary infrastructure and addressing critical human resource issues. Indeed, current difficulties faced in HICs and LMICs have to do with a growing lack of human resources in relation to the increasing needs for health and social care.⁹⁷

Our review has a few limitations. We may not have identified all eligible reviews because indexing in the science of scaling is an ongoing process. We did not seek the impact on our results of the overlap of primary studies included in the reviews because their research questions were different, but overlap may have resulted in overrepresentation in reporting of review characteristics (Table 2). Also, economic evaluations of some of the scaling programs may not have been included in reviews because such assessments are rarely published under the format of a knowledge synthesis and are rarely available other than as internal documents of the organization in question. Our search of the gray literature required that the reports be publicly available online. We assumed that economic evaluations would be no more detailed than in their organizations' internal reports, which would be excluded for not describing their methods and therefore not meeting our study criteria. We recommend an environmental scan for this topic. However, our results are based on a large number of reviews and indicate important knowledge gaps.

Conclusions

The importance of scaling health and social care interventions has been reinforced by the COVID-19 pandemic. A recent UN roadmap established implementation science as a key domain for COVID-19 recovery, highlighting the urgency of examining how interventions that have been successfully implemented in one setting can be adapted to different contexts in ways that are effective and equitable.²² This umbrella review, in which we review the evidence on the scaling of innovations in health and social care, is therefore timely. Our analysis shows where scaling is happening, which global regions are publishing on the topic, what scaling is focused on, and what scalars are using to guide their efforts.

The constant increase in publications confirms that scaling is a science in evolution. Nevertheless, the low quality of systematic reviews and the critical gaps found in most of their methodologies showed a need for better guidance and reporting in scaling research, or at least in review research.

Although LMICs have become experts in scaling to control infectious diseases and avert health problems related to economic and social crises, the authors of most publications were in HICs. Researchers must work not only toward true collaborations among patients, public, and stakeholders but also equitable partnerships between HICs and LMICs. As the science of scaling spreads across the world, it is essential to improve the power balance in research. All regions—both North and South—will suffer if this imbalance is not corrected.

We noted important evidence gaps regarding the type of innovations that are scaled. In LMICs, scaling efforts need to expand beyond controlling infectious diseases and averting humanitarian crises, whereas in HICs, scaling needs to move beyond the culture of perpetual pilot projects and address noncommunicable and chronic diseases. These gaps, combined with the advances in the science of scaling, suggest that it is time governments and funding agencies prioritize scaling not only health but also social innovations that address other complex conditions such those relating to aging and mental health. Programs in both LMICs and HICs also need to be scaled to reach more remote populations, such as Indigenous peoples, for prevention, screening, and treatment.

Above all, the aim of scaling health and social innovations must be to increase their impact including in reducing inequities and not just in reducing health care costs.^{7,28,30,98} The evidence of an increase in the engagement of patients and public in the scaling process and the emergence of ethical, legal, and societal considerations seem promising signals of this commitment.

References

1. Begin M, Eggertson L, Macdonald N. A country of perpetual pilot projects. *CMAJ*. 2009;180(12):1185, E88-E889. <https://doi.org/10.1503/cmaj.090808>
2. World Health Organization—ExpandNet. Nine steps for developing a scaling-up strategy. World Health Organization. 2010.

- Accessed August 23, 2022. <https://apps.who.int/iris/handle/10665/44432>
3. Fixsen D, Blase K, Metz A, Van Dyke M. Statewide implementation of evidence-based programs. *Except Child*. 2013;79(3):213-230. <https://doi.org/10.1177/001440291307900206>
 4. Sheridan DJ. Research: increasing value, reducing waste. *Lancet*. 2014;383(9923):1123. [https://doi.org/10.1016/S0140-6736\(14\)60556-0](https://doi.org/10.1016/S0140-6736(14)60556-0)
 5. Milat AJ, Newson R, King L, et al. A guide to scaling up population health interventions. *Public Health Res Pract*. 2016;26(1):e2611604. doi:10.17061/phrp2611604
 6. Ben Charif A, Zomahoun HTV, Gogovor A, et al. Tools for assessing the scalability of innovations in health: a systematic review. *Health Res Policy Syst*. 2022;20(1):34. <https://doi.org/10.1186/s12961-022-00830-5>
 7. McLean R, Gargani J. *Scaling Impact: Innovation for the Public Good*. 1st ed. Routledge; 2019.
 8. McLean R, Gargani J. In brief: scaling science. International Development Research Centre. 2020. Accessed June 15, 2021. <http://hdl.handle.net/10625/59067>
 9. Greenhalgh T, Papoutsi C. Spreading and scaling up innovation and improvement. *BMJ*. 2019;365:l2068. <https://doi.org/10.1136/bmj.l2068>
 10. Barker PM, Reid A, Schall MW. A framework for scaling up health interventions: lessons from large-scale improvement initiatives in Africa. *Implement Sci*. 2016;11:12. <https://doi.org/10.1186/s13012-016-0374-x>
 11. Aarons GA, Sklar M, Mustanski B, Benbow N, Brown CH. "Scaling-out" evidence-based interventions to new populations or new health care delivery systems. *Implement Sci*. 2017;12(1):111. <https://doi.org/10.1186/s13012-017-0640-6>
 12. Dube L, Jha S, Faber A, et al. Convergent innovation for sustainable economic growth and affordable universal health care: innovating the way we innovate. *Ann N Y Acad Sci*. 2014;1331(1):119-141. <https://doi.org/10.1111/nyas.12548>
 13. Kimble L, Massou MR. What do we mean by Innovation in Healthcare? *EMJ Innov*. 2017;1(1):89-91. doi:10.33590/emjinnov/10314103
 14. WHO Innovation Hub. World Health Organization. 2021. Accessed February 2, 2023. <https://www.who.int/teams/digital-health-and-innovation/who-innovation-hub#:~:text=The%20Innovation%20Hub%20focuses%20on,%2Dbased%2C%20or%20a%20combination>

15. World Health Organization—ExpandNet. Beginning with the end in mind: planning pilot projects and other programmatic research for successful scaling up. World Health Organization. 2011. Accessed August 23, 2022. <https://apps.who.int/iris/handle/10665/44708>
16. Roger EM. *Diffusion of Innovations*. 5th ed. Free Press; 2003.
17. Lebel J, McLean R. A better measure of research from the global south. *Nature*. 2018;559(7712):23-26. <https://doi.org/10.1038/d41586-018-05581-4>
18. Social Sciences and Humanities Research Council. Guidelines for effective knowledge mobilization. Government of Canada. 2019. Accessed August 23, 2022. https://www.sshrc-crsh.gc.ca/funding-financement/policies-politiques/knowledge_mobilisation-mobilisation_des_connaissances-eng.aspx
19. World Health Organization—R&D Blue Print. A coordinated global research roadmap: 2019 novel coronavirus. World Health Organization. 2020. Accessed August 23, 2022. <https://www.who.int/publications/m/item/a-coordinated-global-research-roadmap>
20. Gogovor A, Zomahoun HTV, Ben Charif A, et al. Essential items for reporting of scaling studies of health interventions (SUCCEED): protocol for a systematic review and Delphi process. *Syst Rev*. 2020;9(1):11. <https://doi.org/10.1186/s13643-019-1258-3>
21. Graham ID, Logan J, Harrison MB, et al. Lost in knowledge translation: time for a map? *J Contin Educ Health Prof*. 2006;26(1):13-24. <https://doi.org/10.1002/chp.47>
22. UN research roadmap for the COVID-19 recovery: leveraging the power of science for a more equitable, resilient and sustainable future. United Nations. 2020. Accessed March 30, 2022. <https://www.un.org/en/coronavirus/communication-resources/un-research-roadmap-covid-19-recovery>
23. Scaling an innovation during a pandemic. UNICEF. 2021 Accessed August 23, 2022. Accessed August 23, 2022. <https://www.unicef.org/supply/stories/scaling-innovation-during-pandemic>
24. Campbell D, Edwards B, Milat A, et al. NSW Health COVID-19 emergency response priority research program: a case study of rapid translation of research into health decision making. *Public Health Res Pract*. 2021;31(4):e3142124. doi:10.17061/phrp3142124
25. Freeman M, Savva N, Scholtes S. Economies of scale and scope in hospitals: an empirical study of volume spillovers. *Manag Sci*. 2021;67(2):673-697. <https://doi.org/10.1287/mnsc.2019.3572>
26. Kenton W. Economies of scale: what are they and how are they used? Investopedia. 2022. Accessed August 23, 2022. <https://www.investopedia.com/terms/e/economiesofscale.asp>

27. Simmons R, Fajans P, Ghiron L. Scaling up health service delivery: from pilot innovations to policies and programmes. World Health Organization. 2007. <https://apps.who.int/iris/handle/10665/43794>
28. Mangham LJ, Hanson K. Scaling up in international health: what are the key issues? *Health Policy Plan.* 2010;25(2):85-96. <https://doi.org/10.1093/heapol/czp066>
29. Hirschhorn LR, Talbot JR, Irwin AC, et al. From scaling up to sustainability in HIV: potential lessons for moving forward. *Global Health.* 2013;9:57. <https://doi.org/10.1186/1744-8603-9-57>
30. Zomahoun HTV, Ben Charif A, Freitas A, et al. The pitfalls of scaling up evidence-based interventions in health. *Glob Health Action.* 2019;12(1):1670449. <https://doi.org/10.1080/16549716.2019.1670449>
31. Bunge M. Mechanism and explanation. *Philos Soc Sci.* 1997;27(4):410-465. <https://doi.org/10.1177/004839319702700402>
32. Rycroft-Malone J, McCormack B, Hutchinson AM, et al. Realist synthesis: illustrating the method for implementation research. *Implement Sci.* 2012;7:33. <https://doi.org/10.1186/1748-5908-7-33>
33. Milat AJ, Bauman A, Redman S. Narrative review of models and success factors for scaling up public health interventions. *Implement Sci.* 2015;10:113. <https://doi.org/10.1186/s13012-015-0301-6>
34. Ben Charif A, Zomahoun HTV, LeBlanc A, et al. Effective strategies for scaling up evidence-based practices in primary care: a systematic review. *Implement Sci.* 2017;12(1):139. <https://doi.org/10.1186/s13012-017-0672-y>
35. Milat AJ, King L, Bauman A, Redman S. Scaling up health promotion interventions: an emerging concept in implementation science. *Health Promot J Austr.* 2011;22(3):238. <https://doi.org/10.1071/he11238>
36. Zakiyah N, van Asselt AD, Roijmans F, Postma MJ. Economic evaluation of family planning interventions in low and middle income countries: a systematic review. *PLoS One.* 2016;11(12):e0168447. <https://doi.org/10.1371/journal.pone.0168447>
37. Hempel S, O'Hanlon C, Lim YW, et al. Spread tools: a systematic review of components, uptake, and effectiveness of quality improvement toolkits. *Implement Sci.* 2019;14(1):83. <https://doi.org/10.1186/s13012-019-0929-8>
38. McCrabb S, Lane C, Hall A, et al. Scaling-up evidence-based obesity interventions: a systematic review assessing intervention

- adaptations and effectiveness and quantifying the scale-up penalty. *Obes Rev.* 2019;20(7):964-982. <https://doi.org/10.1111/obr.12845>
39. Bulthuis SE, Kok MC, Raven J, Dieleman MA. Factors influencing the scale-up of public health interventions in low- and middle-income countries: a qualitative systematic literature review. *Health Policy Plan.* 2020;35(2):219-234. <https://doi.org/10.1093/heapol/czz140>
 40. Carroll G, Safon C, Buccini G, et al. A systematic review of costing studies for implementing and scaling-up breastfeeding interventions: what do we know and what are the gaps? *Health Policy Plan.* 2020;35(4):461-501. <https://doi.org/10.1093/heapol/czaa005>
 41. Aromataris E, Fernandez R, Godfrey C, et al. Umbrella reviews. In: Aromataris E, Munn Z, eds. *Joanna Briggs Institute Manual for Evidence Synthesis*. Joanna Briggs Institute; 2020:ch. 10. doi:10.46658/JBIMES-20-11
 42. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ.* 2021;372:n71. <https://doi.org/10.1136/bmj.n71>
 43. Légaré F, Plourde KV, Charif AB, et al. Evidence on scaling in health and social care: protocol for a living umbrella review. *Syst Rev.* 2021;10(1):261. <https://doi.org/10.1186/s13643-021-01813-3>
 44. Grant MJ, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Info Libr J.* 2009;26(2):91-108. <https://doi.org/10.1111/j.1471-1842.2009.00848.x>
 45. Sutton A, Clowes M, Preston L, Booth A. Meeting the review family: exploring review types and associated information retrieval requirements. *Health Info Libr J.* 2019;36(3):202-222. <https://doi.org/10.1111/hir.12276>
 46. Evans D. The provision of health and social care services for older people by respite providers. *Contemp Nurse.* 2013;45(2):255-263. <https://doi.org/10.5172/conu.2013.45.2.255>
 47. Burge P, Netten A, Gallo F. Estimating the value of social care. *J Health Econ.* 2010;29(6):883-894. <https://doi.org/10.1016/j.jhealeco.2010.08.006>
 48. Strings attached: CADTH's database search filter. Canadian Agency for Drugs and Technologies in Health. 2020. Accessed July 2020. https://searchfilters.cadth.ca/list?q=systematic&p=1&ps=20&topic_facet=metaanalyses%20000000%7CMeta-analyses&topic_facet=systematic%20reviews%20000000%7CSystematic%20reviews

49. McGowan J, Sampson M, Salzwedel DM, et al. PRESS Peer Review of Electronic Search Strategies: 2015 guideline statement. *J Clin Epidemiol*. 2016;75:40-46. <https://doi.org/10.1016/j.jclinepi.2016.01.021>
50. Bramer WM, Giustini D, de Jonge GB, Holland L, Bekhuis T. De-duplication of database search results for systematic reviews in EndNote. *J Med Libr Assoc*. 2016;104(3):240-243. <https://doi.org/10.3163/1536-5050.104.3.014>
51. EndNote, version X9. Clarivate Analytics; 2018.
52. Covidence Systematic Review Software. Veritas Health Innovation; 2014. Accessed February 11, 2021. <https://www.covidence.org>
53. DeepL Translator Software. DeepL SE; 2017. Accessed May 7, 2020. <https://www.deepl.com/translator>
54. Yamey G. Scaling up global health interventions: a proposed framework for success. *PLoS Med*. 2011;8(6):e1001049. <https://doi.org/10.1371/journal.pmed.1001049>
55. Zamboni K, Schellenberg J, Hanson C, Betran AP, Dumont A. Assessing scalability of an intervention: why, how and who? *Health Policy Plan*. 2019;34(7):544-552. <https://doi.org/10.1093/heapol/czz068>
56. Wilson SE, Morris SS, Gilbert SS, et al. Scaling up access to oral rehydration solution for diarrhea: learning from historical experience in low- and high-performing countries. *J Glob Health*. 2013;3(1):010404. <https://doi.org/10.7189/jogh.03.010404>
57. Shrimpton R, du Plessis LM, Delisle H, et al. Public health nutrition capacity: assuring the quality of workforce preparation for scaling up nutrition programmes. *Public Health Nutr*. 2016;19(11):2090-2100. <https://doi.org/10.1017/S136898001500378X>
58. Johns B, Torres TT; WHO-CHOICE. Costs of scaling up health interventions: a systematic review. *Health Policy Plan*. 2005;20(1):1-13. <https://doi.org/10.1093/heapol/czi001>
59. Proctor E, Silmere H, Raghavan R, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Health*. 2011;38(2):65-76. <https://doi.org/10.1007/s10488-010-0319-7>
60. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ*. 2017;358:j4008. <https://doi.org/10.1136/bmj.j4008>
61. DAC list of ODA recipients. Organisation for Economic Co-operation and Development. 2022. Accessed May 10, 2022.

- <https://www.oecd.org/dac/financing-sustainable-development/development-finance-standards/DAC-List-of-ODA-Recipients-for-reporting-2022-23-flows.pdf>
62. Pollock M, Fernandes R, Becker L, Pieper D, Hartling L. Overviews of reviews. In: Higgins J, Thomas J, Chandler J, et al., eds. *Cochrane Handbook for Systematic Reviews of Interventions version 6.0*. 2020:ch. 5.
 63. Munk C, Portnoy A, Suharlim C, et al. Systematic review of the costs and effectiveness of interventions to increase infant vaccination coverage in low- and middle-income countries. *BMC Health Serv Res*. 2019;19(1):741. <https://doi.org/10.1186/s12913-019-4468-4>
 64. Perez-Escamilla R, Hromi-Fiedler AJ, Gubert MB, et al. Becoming Breastfeeding Friendly Index: development and application for scaling-up breastfeeding programmes globally. *Matern Child Nutr*. 2018;14(3):e12596. <https://doi.org/10.1111/mcn.12596>
 65. Milat A, Lee K, Conte K, et al. Intervention Scalability Assessment Tool: a decision support tool for health policy makers and implementers. *Health Res Policy Syst*. 2020;18(1):1. <https://doi.org/10.1186/s12961-019-0494-2>
 66. Moran AC, Kerber K, Pfitzer A, et al. Benchmarks to measure readiness to integrate and scale up newborn survival interventions. *Health Policy Plan*. 2012;27(suppl 3):iii29-iii39. <https://doi.org/10.1093/heapol/czs046>
 67. Simonet D. Healthcare reforms and cost reduction strategies in Europe: the cases of Germany, UK, Switzerland, Italy and France. *Int J Health Care Qual Assur*. 2010;23(5):470-488. <https://doi.org/10.1108/09526861011050510>
 68. Cahan EM, Kocher B, Bohn R. Why isn't innovation helping reduce health care costs? *Health Affairs Blog*. 2020. <https://doi.org/10.1377/forefront.20200602.168241>
 69. Policy recommendations for reducing health care costs. National Health Council. 2021. Accessed August 24, 2022. <https://nationalhealthcouncil.org/wp-content/uploads/2021/09/NHC-Health-Care-Costs-2021-Recommendations.pdf>
 70. Rodwin MA. The dark side of a consumer-driven health system. *Front Health Serv Manage*. 2003;19(4):31. <https://ssrn.com/abstract=1012147>
 71. Hagel J, Brown JS, Keith J, Hoversten S. A consumer-driven culture of health. Deloitte Insights. 2015. Accessed August 23, 2022. <https://www2.deloitte.com/us/en/insights/industry/health-care/future-of-us-health-care.html>

72. Gusmano MK, Maschke KJ, Solomon MZ. Patient-centered care, yes; patients as consumers, no. *Health Aff.* 2019;38(3):368-373. <https://doi.org/10.1377/hlthaff.2018.05019>
73. Johnson WG, Butler R, Harootunian G, Wilson B, Linan M. Registered nurses: the curious case of a persistent shortage. *J Nurs Schol-arsb.* 2016;48(4):387-396. <https://doi.org/10.1111/jnu.12218>
74. Britnell M. *Human: Solving the Global Workforce Crisis in Healthcare.* Oxford University Press; 2019.
75. Martin D. *Better Now: Six Big Ideas to Improve Health Care for All Canadians.* Allen Lane—Penguin Canada; 2017.
76. Research partnerships for scaling up mental health interventions in low-and middle-income countries (U19). National Institutes of Health. 2019. Accessed August 23, 2022. <https://grants.nih.gov/grants/guide/pa-files/PAR-16-174.html>
77. Canadian Institutes of Health Research—Institute of Health Services and Policy Research. Funding opportunity: spread and scale of community-based primary health care and ehealth innovations. Canadian Institutes of Health Research. 2019. Accessed August 23, 2022. <https://cihr-irsc.gc.ca/e/documents/ihspr-spread-scale-funding-en.pdf>
78. NIHR joins forces with Global Alliance for Chronic Diseases (GACD). National Institute for Health and Care Research. 2020. Accessed August 23, 2022. <https://www.nihr.ac.uk/news/nihr-joins-forces-with-global-alliance-for-chronic-diseases-gacd/24733>
79. Canadian Institutes of Health Research—Institute of Health Services and Policy Research. Strategic plan 2021–2026: accelerate health care system transformation through research to achieve the quadruple aim and health equity for all. Canadian Institutes of Health Research. 2021. Accessed August 23, 2022. <https://cihr-irsc.gc.ca/e/52481.html>
80. Marchione TJ. *Scaling Up, Scaling Down: Overcoming Malnutrition in Developing Countries.* 1st ed. Gordon and Breach; 1999.
81. Bhattacharyya O, Wu D, Mossman K, et al. Criteria to assess potential reverse innovations: opportunities for shared learning between high- and low-income countries. *Global Health.* 2017;13(1):4. <https://doi.org/10.1186/s12992-016-0225-1>
82. Hedt-Gauthier B, Airhihenbuwa CO, Bawah AA, et al. Academic promotion policies and equity in global health collaborations. *Lancet.* 2018;392(10158):1607-1609. [https://doi.org/10.1016/S0140-6736\(18\)32345-6](https://doi.org/10.1016/S0140-6736(18)32345-6)
83. Urassa M, Lawson DW, Wamoyi J, et al. Cross-cultural research must prioritize equitable collaboration. *Nature Hum*

- Behav.* 2021;5(6):668-671. <https://doi.org/10.1038/s41562-021-01076-x>
84. McLean R, Ofir Z, Etherington A, Acevedo M, Feinstein O. Research Quality Plus (RQ+) – evaluating research differently. International Development Research Centre. 2022. Accessed August 23, 2022. <http://hdl.handle.net/10625/60945>
 85. Chaminuka P, Llorente B, Rincón T, Syed A. Scaling the impact of innovation and research: a call to action from the Global South. International Development Research Centre. 2022. Accessed August 23, 2022. <http://hdl.handle.net/10625/60901>
 86. Indig D, Lee K, Grunseit A, Milat A, Bauman A. Pathways for scaling up public health interventions. *BMC Public Health.* 2017;18(1):68. <https://doi.org/10.1186/s12889-017-4572-5>
 87. Gray S, McKay H, Nettlefold L, et al. Physical activity is good for older adults-but is programme implementation being overlooked? A systematic review of intervention studies that reported frameworks or measures of implementation. *Br J Sports Med.* 2021;55(2):84-91. <https://doi.org/10.1136/bjsports-2020-102465>
 88. D'Agostina A, Wun J, Narayan A, Tharaney M, William T. Defining scale-up of nutrition projects. USAID. 2014. Accessed July 28, 2021. https://www.spring-nutrition.org/sites/default/files/publications/briefs/spring_scale_up_definition_working_paper.pdf
 89. Price-Kelly H, van Haeren L, McLean R. The scaling playbook. International Development Research Centre. 2020. Accessed July 24, 2020. <https://idl-bnc-idrc.dspacedirect.org/bitstream/handle/10625/58780/IDL-58780.pdf?sequence=2&isAllowed=y>
 90. Nundy S, Cooper LA, Mate KS. The quintuple aim for health care improvement: a new imperative to advance health equity. *JAMA.* 2022;327(6):521-522. <https://doi.org/10.1001/jama.2021.25181>
 91. Ben Charif A, Plourde KV, Guay-Belanger S, et al. Strategies for involving patients and the public in scaling-up initiatives in health and social services: protocol for a scoping review and Delphi survey. *Syst Rev.* 2021;10(1):55. <https://doi.org/10.1186/s13643-021-01597-6>
 92. Bradley EH, Curry LA, Taylor LA, et al. A model for scale up of family health innovations in low-income and middle-income settings: a mixed methods study. *BMJ Open.* 2012;2(4):e000987. <https://doi.org/10.1136/bmjopen-2012-000987>
 93. Miake-Lye I, Shekelle P. Scaling beyond early adopters: a systematic review and key informant perspectives. Evidence Synthesis Program Center. 2019. Accessed July 28, 2021. <https://www.hsrd.research.va.gov/publications/esp/scaling-REPORT.pdf>

94. Huaynoca S, Svanemyr J, Chandra-Mouli VC, Moreno Lopez DJ. Documenting good practices: scaling up the youth friendly health service model in Colombia. *Reprod Health*. 2015;12:90. <https://doi.org/10.1186/s12978-015-0079-7>
95. Ehrenkranz P, Grimsrud A, Rabkin M. Differentiated service delivery: Navigating the path to scale. *Curr Opin HIV AIDS*. 2019;14(1):60-65. <https://doi.org/10.1097/COH.0000000000000509>
96. Pallas SW, Minhas D, Perez-Escamilla R, et al. Community health workers in low- and middle-income countries: what do we know about scaling up and sustainability? *Am J Public Health*. 2013;103(7):e74-e82. <https://doi.org/10.2105/AJPH.2012.301102>
97. Global strategy on human resources for health: workforce 2030: reporting at seventy-fifth World Health Assembly. World Health Organization. 2022. Accessed February 2, 2023. <https://www.who.int/news/item/02-06-2022-global-strategy-on-human-resources-for-health-workforce-2030>
98. Knight R, Small W, Pakula B, Thomson K, Shoveller J. A scoping study to identify opportunities to advance the ethical implementation and scale-up of HIV treatment as prevention: priorities for empirical research. *BMC Med Ethics*. 2014;15:54. <https://doi.org/10.1186/1472-6939-15-54>

Funding/Support: This research project is funded by a Canadian Institutes of Health Research (CIHR) Foundation Grant (#FDN-159937). AG is funded by the CIHR Patient-Oriented Research Awards. FL holds a Tier 1 Canada Research Chair in Shared Decision Making and Knowledge Translation. LW is supported by a National Health and Medical Research Council Investigator Grant (APP1197022) and Centre for Research Excellence (APP1153479). The financial providers were not involved in the project.

Acknowledgments: We thank Frederic Bergeron, MLS, for the peer review and evaluation of the electronic search strategies and for performing the search strategy. We acknowledge the precious work of the reviewers for the review selection and data extraction: Talagbe Gabin Akpo, Gloria Ayivi-Vinz, Blanchard Conombo, Maya Fakhfakh, Lucas Gomes Souza, Felly Bakwa Kanyinga, Laura Langevin, Tania Lognon, and Yuxi Wang. We also thank Louisa Blair for her kind editorial help with the paper.

Address correspondence to: France Légaré, C.Q., BSc Arch, MD, MSc, PhD, CCFP, FCFP, VITAM. Centre de recherche en santé durable, 2480, Chemin de la Canardière, QC G1G 2G1, Canada (email: France.Legare@fmed.ulaval.ca).

Supplementary Material

Additional supporting information may be found in the online version of this article at [http://onlinelibrary.wiley.com/journal/10.1111/\(ISSN\)1468-0009](http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1468-0009):