

REVIEW

Open Access



# Tools for assessing the scalability of innovations in health: a systematic review

Ali Ben Charif<sup>1</sup> , Hervé Tchala Vignon Zomahoun<sup>2,3,4</sup>, Amédé Gogovor<sup>5,6,7,8</sup>, Mamane Abdoulaye Samri<sup>5,6,7</sup>, José Massougbodji<sup>9</sup>, Luke Wolfenden<sup>10,11,12</sup>, Jenny Ploeg<sup>13</sup>, Merrick Zwarenstein<sup>14</sup>, Andrew J. Milat<sup>15,16</sup>, Nathalie Rheault<sup>5,8</sup>, Youssoufa M. Ousseine<sup>17</sup>, Jennifer Salerno<sup>13</sup>, Maureen Markle-Reid<sup>13,18</sup> and France Légaré<sup>5,6,7,8,19\*</sup>

## Abstract

**Background:** The last decade has seen growing interest in scaling up of innovations to strengthen healthcare systems. However, the lack of appropriate methods for determining their potential for scale-up is an unfortunate global handicap. Thus, we aimed to review tools proposed for assessing the scalability of innovations in health.

**Methods:** We conducted a systematic review following the COSMIN methodology. We included any empirical research which aimed to investigate the creation, validation or interpretability of a scalability assessment tool in health. We searched Embase, MEDLINE, CINAHL, Web of Science, PsycINFO, Cochrane Library and ERIC from their inception to 20 March 2019. We also searched relevant websites, screened the reference lists of relevant reports and consulted experts in the field. Two reviewers independently selected and extracted eligible reports and assessed the methodological quality of tools. We summarized data using a narrative approach involving thematic syntheses and descriptive statistics.

**Results:** We identified 31 reports describing 21 tools. Types of tools included criteria (47.6%), scales (33.3%) and checklists (19.0%). Most tools were published from 2010 onwards (90.5%), in open-access sources (85.7%) and funded by governmental or nongovernmental organizations (76.2%). All tools were in English; four were translated into French or Spanish (19.0%). Tool creation involved single (23.8%) or multiple (19.0%) types of stakeholders, or stakeholder involvement was not reported (57.1%). No studies reported involving patients or the public, or reported the sex of tool creators. Tools were created for use in high-income countries (28.6%), low- or middle-income countries (19.0%), or both (9.5%), or for transferring innovations from low- or middle-income countries to high-income countries (4.8%). Healthcare levels included public or population health (47.6%), primary healthcare (33.3%) and home care (4.8%). Most tools provided limited information on content validity (85.7%), and none reported on other measurement properties. The methodological quality of tools was deemed inadequate (61.9%) or doubtful (38.1%).

**Conclusions:** We inventoried tools for assessing the scalability of innovations in health. Existing tools are as yet of limited utility for assessing scalability in health. More work needs to be done to establish key psychometric properties of these tools.

*Trial registration* We registered this review with PROSPERO (identifier: CRD42019107095)

\*Correspondence: [france.legare@mfa.ulaval.ca](mailto:france.legare@mfa.ulaval.ca)

<sup>5</sup> VITAM – Centre de recherche en santé durable, Université Laval, Quebec City, QC, Canada

Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

**Keywords:** Scalability, Scaling up, Scaling, Spread, Health innovations, Patient and public involvement, Scalability assessment tool, Systematic review

## Background

Various innovations have been developed and successfully piloted to strengthen healthcare systems in low-, middle- or high-income countries [1–3]. A health innovation refers to a set of behaviours, routines and ways of working that are perceived as new; that aim to improve health outcomes, administrative efficiency, cost-effectiveness or user experience; and that are implemented through planned action [4–6]. But there is a global delivery gap between innovations for which evidence of effectiveness has been established and those that actually reach the people who could benefit [7, 8]. Thus, the last decade has seen growing interest in the scaling up of health innovations. Scaling up, or expanding the impact and reach of effective innovations, could reduce waste and inequalities in health settings and improve outcomes [7–9]. For example, up to 85% of all maternal, neonatal and child deaths in low- or middle-income countries could potentially be averted through scaling up of successfully piloted innovations [10]. The science of knowledge mobilization, or moving knowledge into action (also known variously as knowledge translation and implementation science), can be a key instrument for closing this gap by taking evidence-based innovations and testing strategies to move them into wider practice [11–13]. Thus, there is a need for tools to help identify evidence-based innovations that could be successfully expanded or scaled up to reach more patients in healthcare systems.

There are various definitions of scaling up [14], ranging from an increase in the number of beneficiaries, organizations or geographic sites, to more complex definitions in which expanding the variety, equity and sustainability of an innovation is also considered [1, 6, 15]. Some innovations are implemented at scale before ever going through a pilot trial or small-scale introduction [16]. This was the case with the coronavirus disease 2019 (COVID-19) vaccines in Canada, for example, which were developed elsewhere through clinical research and then introduced simultaneously nationwide at the local level. In some situations, scale-up is transnational; for example, innovations adopted first in a low- or middle-income country are then transferred or scaled up to a high-income country [17, 18]. Scale-up can be nonlinear, and is inherently complex and often political [19]. Scalability is defined as the “ability of a health innovation shown to be efficacious on a small scale and/or under controlled conditions to be expanded under real-world conditions to reach a greater proportion of the eligible population,

while retaining effectiveness” [20]. Here, we consider scalability broadly as also including assessing whether the innovation can be replicated, transferred or sustained [6, 21].

Among other considerations in preparing for scale-up, decision-makers need to assess the more technical scalability components of an innovation [2, 3]. In 2003, Everett Rogers identified key innovation characteristics relevant for assessing scalability: relative advantage (which includes effectiveness), compatibility, complexity, comprehensibility (to the user), trialability, observability and potential re-invention (i.e. adaptation) [5]. Since then, others have adapted and added to these characteristics [6]. Milat’s scalability assessment tool [22], for example, based on existing frameworks, guides and checklists, is a recent and comprehensive effort to select and summarize essential components of a scale-up preparedness plan [1, 15, 21, 23]. In spite of these advances, however, scalability assessments are still often overlooked by those responsible for developing and delivering innovations in health [1, 16].

Thus, scalability assessments target certain key components or properties that are critical for scale-up. For example, many health innovations are scaled up in the absence of evidence of beneficial impact [16], a scalability component that is an essential predictor of successful scale-up [3, 6, 21]. Scalability assessments should also anticipate known pitfalls of scale-up, that is, elements that have compromised the success of scaling up, such as the replicating of harms at scale [24]. While few studies focus on scale-up failures, studies that do so can throw into relief gaps that otherwise might be overlooked [25]. Failing to involve patients and the public, especially those who may be socially excluded owing to age, ethnicity, or sex and gender, may also result in poor programmatic outcomes, as scale-up could overlook the concerns of its intended beneficiaries [14, 15, 26].

In addition to the complex strategic, political and environmental considerations surrounding scale-up, end-users (e.g. policy-makers, implementers) lack theoretical, conceptual and practical tools for guiding scalability assessments in health settings [27]. In Canada, many innovation teams have expressed the need for a validated tool for scalability assessment in primary healthcare [2, 3]. No previous knowledge synthesis has been conducted on the measurement properties (i.e. quality aspects such as reliability, validity and responsiveness) of scalability assessment tools. Thus, we aimed

to review existing tools for assessing scalability of health innovations, describing how the tools were created and validated, and describing the scalability components they target. Our research question was as follows: “What tools are available for assessing the scalability of innovations in health, how were they created, what are their measurement properties, and what components do they target?”

## Methods

### Design

We performed a systematic review with a comprehensive overview of the components targeted by scalability assessment tools and their measurement properties. We adapted and followed the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) methodology for systematic reviews [28]. We reported the review according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines [29] and the COSMIN reporting recommendations [28]. In this manuscript, the noun “report” refers to a document (paper or electronic) supplying information about a study, and the noun “record” refers to the title or abstract of a report indexed in a database or website [29]. We registered this review in the International Prospective Register of Systematic Reviews (PROSPERO) on 2 May 2019 (registration identifier: CRD42019107095) [30].

### Eligibility criteria

Following the COSMIN approach, we used the following eligibility criteria.

- **Construct:** We included any tool aiming to assess or measure scalability of innovations in health. WHO defines health as “a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity”. According to the International Classification of Health Interventions, types of health innovations could include management, prevention, therapeutic, diagnostic, other (i.e. not classified elsewhere) or unspecified [3, 31].
- **Population:** We included any type of stakeholder or end-user. Stakeholders refer to persons who were involved in the conception, creation or validation of the tool [32]. End-users refer to individuals such as policy-makers who are likely to use the tool to make decisions about scaling up an innovation [33]. Stakeholders can also be end-users, and both can include patients and the public, healthcare providers, policy-makers, investigators, trainees and funders [14]. End-users can be involved in the creation or validation process of the tool, and the level of their involvement may vary from minimal (i.e. receiving information

about it, but with no contributing role) to coproducing the tool (i.e. participating as an equal member of the research team) [14, 34, 35].

- **Instrument:** We included any tool containing items proposed for assessing the scalability of an innovation in health. A tool refers to a structured instrument such as a guide, framework, questionnaire, factors, facilitators or barriers. Items refer to individual elements of the tool such as questions or statements that were mapped to targeted components.
- **Measurement properties:** We included any reports presenting (1) creation of a scalability assessment tool, (2) assessment of one or more measurement properties of the tool or (3) assessment of the interpretability of the tool. A measurement property is defined as a quality aspect of a tool, i.e. reliability, validity and responsiveness [28]. We included any of the following nine measurement properties: content validity, structural validity, internal consistency, cross-cultural validity or measurement invariance, reliability, measurement error, criterion validity, hypotheses testing for construct validity, and responsiveness. We excluded any study protocol and any editorial material, defined as an article that gives the opinions of a person, group or organization (e.g. editorials, commentaries and letters).

In other words, we included any empirical research which aimed to investigate the creation, validation or interpretability of a scalability assessment tool in health settings (Table 1).

### Literature search

Overall, we performed a comprehensive search to identify records through both electronic databases of peer-reviewed literature and secondary searches, including hand searching relevant websites, screening reference lists of included or relevant reports, and consulting experts in the field of scale-up. There was no restriction regarding language, date or country of publication, or type of reports.

First, we searched Embase via embase.com, MEDLINE via Ovid, CINAHL via EBSCO, Web of Science, PsycINFO via Ovid, the Cochrane Library, and ERIC via EBSCO from their inception to 20 March 2019. An information specialist with the Unité de soutien SSA Québec [36] (NR) drafted the preliminary version of the search strategy for Ovid MEDLINE. The search terms were based on previous works to reflect three concepts: scalability [1], tool [37] and health [38]. The preliminary search strategy was reviewed by eight international experts (ABC, HTVZ, LW, JP, MZ, AJM, JS and MMR), and then by a second information specialist in the Faculty

**Table 1** Criteria for considering records or reports for this review

Criteria	Inclusion	Exclusion	Question related to the criteria
Type of report	<input checked="" type="checkbox"/> Original paper <input checked="" type="checkbox"/> Research report <input checked="" type="checkbox"/> Evaluation report <input checked="" type="checkbox"/> Knowledge synthesis <input checked="" type="checkbox"/> Government document	<input checked="" type="checkbox"/> Editorial <input checked="" type="checkbox"/> Commentary <input checked="" type="checkbox"/> Opinion letter <input checked="" type="checkbox"/> Protocol	<i>Is this empirical research using quantitative or qualitative methods?</i>
Aim of study	<input checked="" type="checkbox"/> Development of a tool (e.g. a guide, framework, questionnaire, factors, facilitators or barriers) <input checked="" type="checkbox"/> Assessment of one or more measurement (or psychometric) properties of a tool <input checked="" type="checkbox"/> Assessment of the interpretability of a tool	<input checked="" type="checkbox"/> The study did <b>not</b> present or describe a guide, framework, questionnaire, factors, facilitators or barriers (hereafter referred to as "tool")	<i>Does the study present the development, validation or interpretability of a tool?</i>
Aim of the tool	<input checked="" type="checkbox"/> Tool aiming to assess the scalability of an innovation (i.e. potential or readiness for scale-up, for spread, for transfer, for diffusion or for system wide implementation)	<input checked="" type="checkbox"/> The tool is <b>not</b> intended to be used for assessing the scalability of an innovation	<i>Does the tool aim to evaluate the scalability of an innovation?</i>
Setting	<input checked="" type="checkbox"/> Any health context	<input checked="" type="checkbox"/> The tool is <b>not</b> intended to be used for an innovation in the field of health	<i>Is the tool intended to be used for innovation in the field of health?</i>

of Medicine at Université Laval (F. Bergeron) using the Peer Review of Electronic Search Strategies (PRESS) guideline [39]. The experts were university-based investigators (from Benin, Togo, Comoros, Australia, and Canada) and experts in knowledge mobilization, health services research, health research methodology and scaling up. We resolved any disagreements through a consensus meeting between the two information specialists and a third party (ABC and HTVZ). The search terms were adapted to the above-mentioned databases by removing search terms related to the concept of health in all biomedical databases—the difference in the number of records found in MEDLINE when removing health-related terms was minimal (104 records out of a total of 2528). Details of the search strategy in each electronic database can be found in the appendix (Additional file 1).

*Second*, we identified other records by searching relevant websites, screening reference lists of included or relevant reports, and consulting experts in the field of scale-up. This approach is promoted as a way of reducing publication bias [40]. We consulted Google Scholar, Google web search, and the websites of a list of 24 Canadian and international organizations in both English and French from 10 October to 20 December 2019 (Additional file 2). In French, we used the following keywords: “potentiel de mise à l’échelle”, “potentiel de passage à grande échelle”, “transférabilité”, “mise à l’échelle”, “passage à grande échelle”, “accroissement d’échelle”, “passage à l’échelle” and “diffusion”. In English, we used terms related to the concept of scalability including scalability, transferability, readiness, scale, scaling, upscaling, upscaling, and spread (Additional file 1). We also established

a list of experts in the field of scale-up and asked them via email about documentation of tools they had created or knew about, from 5 to 29 May 2020 (Additional file 3). The list of experts included authors of reports included in this systematic review, authors of reports included in our previous systematic review [1], members of the 12 Canadian Institutes of Health Research (CIHR)-funded Community-Based Primary Health Care (CBPHC) teams [2, 41], and members of the Research on Patient-Oriented Scaling-up (RePOS) network [14].

### Selection process

*First*, we operationalized eligibility criteria using questions with the following responses: “met”, “not met” and “unclear”. Five author reviewers (ABC, AG, MAS, JM and YMO) independently screened a random sample of 5% of records identified with our literature search. We discussed the results of this pilot and reviewed the eligibility criteria. *Second*, two senior end-users and experts in scaling up (JP and MZ) independently screened five records and suggested a minor change in wording to clarify eligibility criteria. *Third*, the same five reviewers independently piloted the selection of another random sample of 5% of the remaining records. We calculated inter-reviewer agreement between these five reviewers using the weighted Cohen’s kappa [42] and considered it substantial when we reached a value of at least 0.60 [43]. *Fourth*, the five reviewers (ABC, AG, MAS, JM and YMO) independently screened all remaining records. We detail the records assignment and kappa calculation in the appendix (Additional file 4). *Fifth*, two reviewers (ABC and MAS) assessed all potentially relevant reports

to identify reports meeting the eligibility criteria. For all ineligible reports, we documented the main reason for exclusion. *Finally*, in all steps, we resolved all disagreements through consensus among reviewers in face-to-face meetings and, when required, with the project leader (ABC). Records that referred to the same report were considered duplicates, but records that referred to reports that were merely similar were considered unique [29]. We used EndNote X9 software to identify duplicates and an Excel form for the selection process.

#### Data collection process

We developed an Excel form to guide extraction of variables based on the COSMIN manual [28]. Six reviewers (ABC, HTVZ, AG, MAS, JM and YMO) performed a calibration exercise to ensure that the form captured all relevant data. Then two reviewers (ABC and MAS) independently extracted data using the Excel form. The following information was extracted from each included unique report:

- characteristics of included tools (e.g., type, date of issue or publication, funding support, language, stakeholder, open-access source, name, scalability components targeted, content and pitfall predictions);
- intended context of use (e.g. income level of country, healthcare level, focus area, end-user and aim); and
- data that could be considered sources of validity for measurement properties. For example, data regarding the tool's content validity could include test blueprint, representativeness of items in relation to the scalability component, logical or empirical relationship of content tested to scalability component, strategies to ensure appropriate content representation, item writer qualifications, and analyses by experts regarding how adequately items represent the content of the scalability component [44].

All disagreements were resolved through consensus between ABC and MAS in face-to-face and virtual meetings. We used Microsoft Teams for the virtual meetings.

#### Quality assessment of tools

We used the COSMIN Risk of Bias checklist to assess the methodological quality of included tools [28]. This checklist contains one box with standards for assessing the tool's methodological quality and nine boxes for assessing the methodological quality of studies that reported measurement properties for tools. In this review, because there were very limited data on content validity and no data on other measurement properties, we assessed the methodological quality of tool creation only, which is

also part of the content validity. Two reviewers (ABC and MAS) independently assessed the quality of all included tools after a pilot using a sample of two tools. We resolved all disagreements through consensus between ABC and MAS in virtual meetings using Microsoft Teams.

The COSMIN standards for tool creation consist of 35 items divided into two parts [45]: Part A addresses the quality of the design and Part B the quality of the pilot study. Part A includes a concept elicitation study performed with end-users to identify relevant items for a new tool, and a clear description of the construct and how it relates to the theory or conceptual framework from which it originates. Part B includes a pilot study performed with end-users to evaluate comprehensiveness and comprehensibility. Each standard is scored on a four-point rating scale: "very good", "adequate", "doubtful" or "inadequate". A standard is rated as "doubtful" if it is doubtful whether the quality aspect is adequate (i.e. minor methodological flaws), and "inadequate" when evidence is provided that the quality aspect is not adequate (i.e. important methodological flaws) [28]. Where a score for a standard was not requested, the option "not applicable" was available. Total scores are determined separately for concept elicitation and pilot test. A total score per tool is obtained by taking the lowest rating of any item (i.e. worst score counts).

#### Data analysis

We analysed and summarized extracted data using a narrative approach involving framework and content analysis [46]. We created an integrated framework of categories for the purpose of this study based on recent work on scaling up. All classification was carried out independently by two reviewers (ABC and MAS) and all disagreements were resolved through consensus in virtual meetings using Microsoft Teams. We used the PRISMA 2020 flowchart to describe the process of tool selection [29]. We summarized the main characteristics of tools, including components targeted by the tools and their methodological quality, in a tabular display using SAS 9.4 software.

*First*, we classified each tool using the three types: (1) scale, (2) checklist or (3) set of criteria. To be considered a scale, each item within the tool had to have a numeric score attached to it so that an overall summary score could be calculated. To be considered a checklist, the tool had to include multiple items to observe for scalability criteria to be met. To be considered "criteria", the tool had to include a list of items (questions or statements) with no proposed responses. *Second*, we mapped each item of each tool to the following 12 possible components targeted by the tool: (C<sub>1</sub>) health problem addressed by the innovation; (C<sub>2</sub>) development process of the innovation;

(C<sub>3</sub>) innovation characteristics; (C<sub>4</sub>) strategic, political or environmental context of the innovation; (C<sub>5</sub>) evidence available for effectiveness of the innovation; (C<sub>6</sub>) innovation costs and quantifiable benefits; (C<sub>7</sub>) potential for implementation fidelity and adaptation of the innovation; (C<sub>8</sub>) potential reach and acceptability to the target population; (C<sub>9</sub>) delivery setting and workforce; (C<sub>10</sub>) implementation infrastructure required for scale-up; (C<sub>11</sub>) sustainability (i.e. longer-term outcomes of the scale-up); and (C<sub>Other</sub>) other components. This classification was based on Milat's 10-component framework [22], to which we added items related to the development process of the innovation such as the use of a theoretical, conceptual or practical framework (C<sub>2</sub>) [2, 3], which is the primary stage of scale-up [16]. *Third*, we determined whether each tool included items related to eight potential pitfalls to be anticipated when planning scale-up of the innovation. Six of those pitfalls were based on a rapid review of points of concern regarding the success or failure of scale-up efforts [24]. To these six pitfalls we added patient and public involvement and sex and gender. These were demonstrations that development or piloting of the innovation had not excluded its targeted beneficiaries (e.g. excluding women in a programme about women's health) [1, 14, 15, 26]. The expanded pitfalls thus consisted of the following: (P<sub>1</sub>) sex and gender considerations; (P<sub>2</sub>) patient and public involvement; (P<sub>3</sub>) the difficulty of cost-effectiveness estimates; (P<sub>4</sub>) the production of health inequities; (P<sub>5</sub>) scaled-up harm; (P<sub>6</sub>) ethics (e.g. informed consent at scale); (P<sub>7</sub>) top-down approaches (i.e. the needs, preferences and culture of beneficiaries of the innovation may be forgotten when scale-up is directed from above); and (P<sub>8</sub>) context (e.g. difficulty in adapting the innovation to certain contexts). *Finally*, we adopted a previous rating system to quantify the extent to which sources of validity evidence for measurement properties of the tools were reported: 0="no discussion or data presented as a source of validity evidence"; 1="data that weakly support the validity evidence"; 2="some data (intermediate level) that support the validity evidence, but with gaps"; and 3="multiple sets of data that strongly support the validity evidence" [44].

## Results

### Study selection

Our electronic search identified 11 299 potentially relevant records. Of these, 2805 were duplicates, leaving 8494 records. Of these, 8422 did not meet the review criteria. With the second random sample of 5% of the 8494 records, we found substantial pair inter-reviewer agreements for decisions regarding inclusion, with kappa values ranging from 0.66 to 0.89 across all reviewers (Additional file 4). Finally, we reviewed a total of 72

reports, retained 13 [2, 47–58] and excluded 59 [59–117] (Additional file 5). In addition, our secondary searches led to the inclusion of 18 additional reports [3, 6, 20–22, 118–130]. Overall, we included a total of 31 reports from all sources [2, 3, 6, 20–22, 47–58, 118–130], which described a total of 21 unique tools (Fig. 1). We included the following tools: the Innovation Scalability Self-administered Questionnaire (ISSaQ) [2, 3], the AnalySe de la Transférabilité et accompagnement à l'Adaptation des Interventions en pRomotion de la santE (ASTAIRE) [53, 54], the Process model for the assessment of transferability (PIET-T) [55], the CORRECT attributes [6, 121, 122], the scalability assessment framework [57], the Intervention Scalability Assessment Tool (ISAT) [22], the Readiness to Spread Assessment Scoring Sheet [125], the Readiness to Receive Assessment Scoring Sheet [126], the Applicability and Transferability of Evidence Tool (A&T Tool) [118, 119], the Scalability Assessment and Planning (SAP) Toolkit [130] and the Scalability Checklist [127–129]. We did not find names for 10 of the tools [20, 21, 47–52, 56, 58, 120, 123, 124].

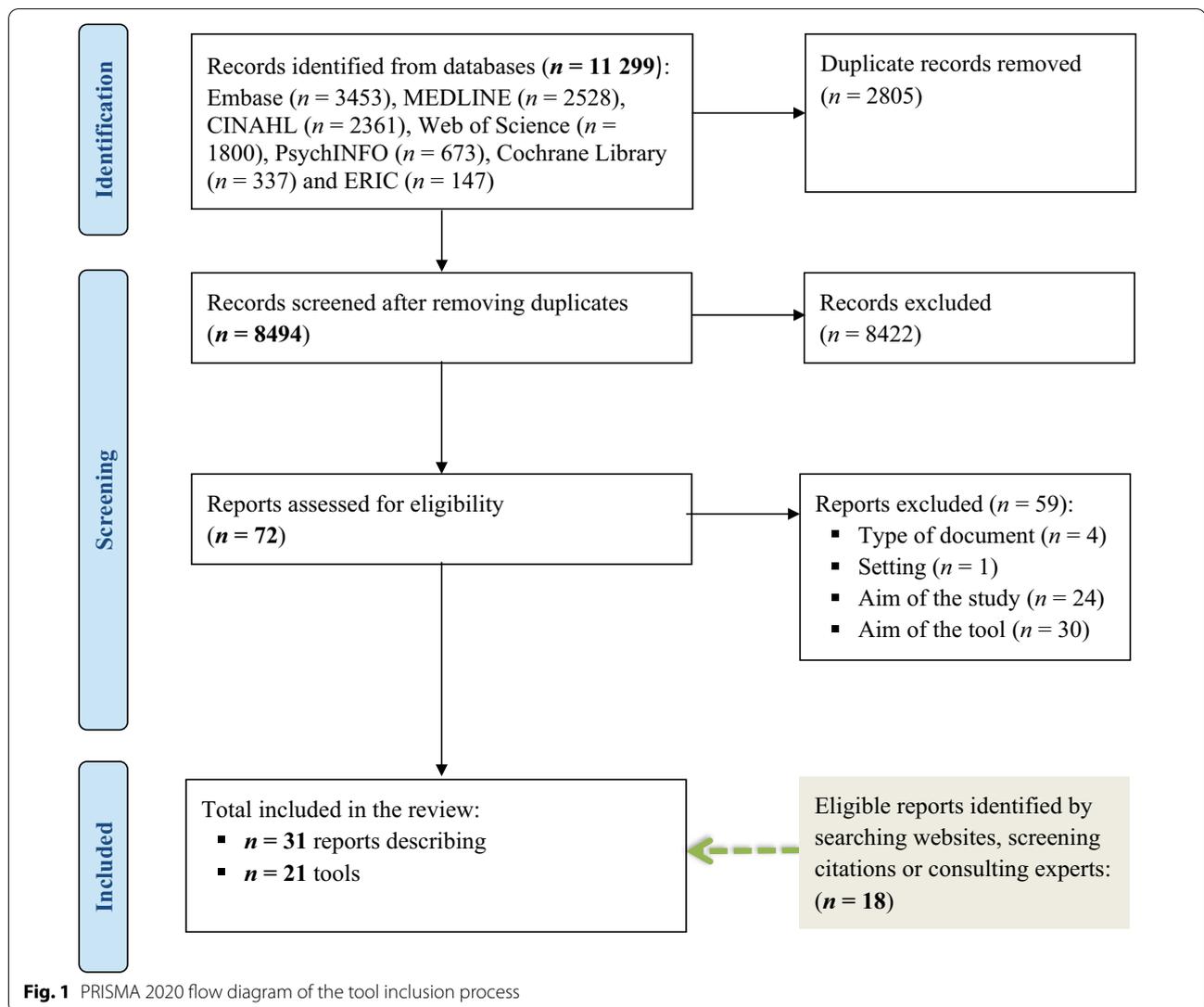
### Characteristics of tools

Characteristics of included tools are outlined in Table 2.

Type and source of tools: most tools were criteria ( $n=10$ , 47.6%), followed by scales ( $n=7$ , 33.3%) and checklists ( $n=4$ , 19.0%). Included tools were created or published from 2005 onwards and the majority since 2010 ( $n=19$ , 90.5%). Their creation was funded by governmental or nongovernmental organizations ( $n=16$ , 76.2%). All tools were in English; three were translated into French only (14.3%) and one into French and Spanish (4.8%). Most tools were available through open-access peer-reviewed journals, ResearchGate or organizational websites ( $n=18$ , 85.7%).

Scalability components: all tools targeted multiple components. The most frequently targeted components were potential implementation fidelity and adaptation (81.0%), delivery setting and workforce (81.0%), and implementation infrastructure (81.0%). The three least frequently targeted were health problems addressed by the innovation (57.1%), sustainability (47.6%), and development process of innovations (28.6%) (Table 2).

Content of tools: tools contained a total of 320 items (e.g. questions, statements) mapping to targeted components (Additional file 6). There was a median of 16 items per tool (interquartile range: 13 items). In 286 items, just one scalability component was targeted; in 27 items, two scalability components were targeted; in five items, three scalability components were targeted; and in two items, four scalability components were targeted. Most items covered delivery setting and workforce (68 items), reach and acceptability for the target population (62 items), and



evidence available for effectiveness of the innovation (42 items). Components least covered by items were problem addressed by the innovation (19 items), development process of the innovation (16 items), and sustainability (12 items).

**Pitfall predictions:** most tools included items that considered contextual pitfalls (90.5%) and cost-effectiveness estimation pitfalls (71.4%). Pitfalls least considered were scaled-up harms (14.3%) and health inequities (4.8%) (Table 2).

**Stakeholder involvement:** no information on stakeholder involvement in tool creation or validation was found for 12 out of the 21 tools (57.1%) (Table 2). No studies reported involving patients or the public, for example, or reported on the sex of tool creators. Tool creation involved single ( $n = 5$ , 23.8%) or multiple ( $n = 4$ , 19.0%) types of stakeholders, including clinicians,

policy-makers, researchers and civil society organizations (Table 2).

#### Intended context of use

Eight tools did not report the income levels of countries for which they were created (38.1%) (Table 3). Six tools were reported as created for use in high-income countries (28.6%), four in low- or middle-income countries (19.0%), two in both (9.5%), and one for transnational transfers from low- or middle-income to high-income contexts (4.8%).

Seven tools did not report which healthcare levels they were created for (33.3%) (Table 3). The largest proportion of tools for which this information was reported were created for public or population health (47.6%), primary healthcare (33.3%) or home care (4.8%) initiatives.

**Table 2** Characteristics of included tools

Name (abbreviation) <sup>a</sup> [References]	Type and source		Language	Type of stakeholder	Open-access source	Scalability component targeted by tools <sup>b</sup>											Number of items	Pitfall predictions <sup>c</sup>	
	Type	Year of issue or publication				Source of funding	C <sub>1</sub>	C <sub>2</sub>	C <sub>3</sub>	C <sub>4</sub>	C <sub>5</sub>	C <sub>6</sub>	C <sub>7</sub>	C <sub>8</sub>	C <sub>9</sub>	C <sub>10</sub>			C <sub>11</sub>
Innovation Scalability Self-administered Questionnaire (ISSAQ) [2, 3]	Checklist	2017	Governmental organization	Researcher	Peer-reviewed journal	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	16	P <sub>3,8</sub>
AnalySe de la Transférabilité et accompagnement à l'Adaptation des Interventions en promotion de la santé (ASTAIRE) [53, 54]		2013	Governmental organization	Researcher	Peer-reviewed journal	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	23	P <sub>2,3,7,8</sub>
WHO/Expend-Net [123, 124]		2011	Governmental and nongovernmental organizations	Not found	Organizational website	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	23	P <sub>1,3,4,7,8</sub>
Scalability Checklist [127–129]		2016	Nongovernmental organization	Not found	Organizational website	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	7	P <sub>3,8</sub>



**Table 2** (continued)

Name (abbreviation) <sup>a</sup> [References]	Type and source		Year of issue or publication	Source of funding	Language	Type of stakeholder	Open-access source	Scalability component targeted by tools <sup>b</sup>											Number of items	Pitfall predictions <sup>c</sup>
	Type	Scale						C <sub>1</sub>	C <sub>2</sub>	C <sub>3</sub>	C <sub>4</sub>	C <sub>5</sub>	C <sub>6</sub>	C <sub>7</sub>	C <sub>8</sub>	C <sub>9</sub>	C <sub>10</sub>	C <sub>11</sub>		
Bhattacharyya et al. [49]	Scale		2017	Governmental and nongovernmental organizations	English	Policy-maker	Peer-reviewed journal	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	8	P <sub>2,3,8</sub>	
Scalability assessment framework [57]			2018	Nongovernmental organization	English	Not found	Not found	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	16	P <sub>3,7,8</sub>	
Intervention Scalability Assessment Tool (ISAT) [22]			2019	Governmental organization	English	Clinician, policy-maker, researcher	Organizational website	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	19	P <sub>3,8</sub>	
Readiness to Spread Assessment Scoring Sheet [125]			2013	Not found	English	Not found	Organizational website	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	4	P <sub>5</sub>	
Readiness to Receive Assessment Scoring Sheet [126]			2013	Not found	English	Not found	Organizational website	✓										4	P <sub>8</sub>	
Applicability and Transferability of Evidence Tool (A&T Tool) [118, 119]			2007	Governmental organization	English, French	Not found	Organizational website	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	21	P <sub>3,7,8</sub>	
Scalability Assessment and Planning (SAP) Toolkit [130]			2018	Not found	English	Researcher	Organizational website					✓	✓	✓	✓	✓	✓	5	Not found	

<sup>a</sup> We did not find names for 10 of the tools, in which case we indicated names of authors or organizations

<sup>b</sup> Scalability components targeted: (C<sub>1</sub>) health problem addressed by the innovation; (C<sub>2</sub>) development process of the innovation; (C<sub>3</sub>) innovation characteristics; (C<sub>4</sub>) strategic, political or environmental context of the innovation; (C<sub>5</sub>) evidence available for effectiveness of the innovation; (C<sub>6</sub>) innovation costs and quantifiable benefits; (C<sub>7</sub>) potential for implementation fidelity and adaptation of the innovation; (C<sub>8</sub>) potential reach and acceptability to the target population; (C<sub>9</sub>) delivery setting and workforce; (C<sub>10</sub>) implementation infrastructure required for scale-up; and (C<sub>11</sub>) sustainability (i.e. longer-term outcomes of the scale-up)

<sup>c</sup> Pitfalls of scale-up relate to: (P<sub>1</sub>) sex and gender considerations; (P<sub>2</sub>) patient and public involvement; (P<sub>3</sub>) the difficulty of cost-effectiveness estimates; (P<sub>4</sub>) the production of health inequities; (P<sub>5</sub>) scaled-up harm; (P<sub>6</sub>) ethics (e.g. informed consent at scale); (P<sub>7</sub>) top-down approaches (i.e. the needs, preferences and culture of beneficiaries of the innovation may be forgotten when scale-up is directed from above); and (P<sub>8</sub>) context (e.g. difficulty in adapting the innovation to certain contexts)

<sup>d</sup> CORRECT attributes: C—credible in that they are based on sound evidence or advocated by respected persons or institutions; O—observable to ensure that potential users can see the results in practice; R—relevant for addressing persistent or sharply felt problems; R—relative advantage over existing practices so that potential users are convinced the costs of implementation are warranted by the benefits; E—easy to install and understand rather than complex and complicated; C—compatible with the potential users' established values, norms and facilities; fit well into the practices of the national programme; and T—testable so that potential users can see the innovation on a small scale prior to large-scale adoption

✓ is a checkmark for the item

**Table 3** Characteristics of intended context of use of included tools

Name (abbreviation) <sup>a</sup> [References]	Income level context	Healthcare level	Focus area	Sex or gender of beneficiaries of the targeted innovations	End-user of tool	Aim of tool	Degree of report of validity evidence for content validity <sup>b</sup>	Methodological quality of tools <sup>c</sup>
Innovation Scalability Self-administered Questionnaire (SSaQ) [2, 3]	High-income country	Primary healthcare	Not found	Female, male	Clinician, policy-maker, researcher	Assess the scalability of innovations in primary healthcare	2	Inadequate
Analyse de la Transférabilité et accompagnement à l'Adaptation des Interventions en promotion de la santé (ASTAIRE) [53, 54]	High-income country	Not found	Health promotion	Not found	Not found	Assess transferability and adaptation of health promotion innovations	3	Doubtful
WHO/ExpendNet [123, 124]	Not found	Not found	Not found	Not found	Researcher, policy-maker, programme manager, funder	Assess the scalability of programmatic research; provide a quick assessment of how easy or difficult it will be to scale up a project that is being planned or proposed or is in the process of implementation	2	Doubtful
Scalability Checklist [127–129]	Not found	Primary healthcare	Reproductive, maternal, newborn, child or adolescent health	Not found	Not found	Prioritize alternatives and identify actions that can be taken to simplify the scaling-up process	2	Doubtful
Baker et al. [47]	Not found	Primary healthcare, home care, public or population health	Reproductive, maternal, newborn, child or adolescent health	Female	Not found	Assess the applicability and transferability of innovations to the Aboriginal and Torres Strait Islander setting	2	Inadequate
Bennett et al. [48]	Low- or middle-income country	Primary healthcare	Reproductive, maternal, newborn, child or adolescent health	Not found	Not found	Explore feasibility and effectiveness of health innovations	1	Inadequate
Burchett et al. [50]	Not found	Public or population health	Not found	Not found	Not found	Assist in the assessment of applicability and transferability	2	Inadequate

**Table 3** (continued)

Name (abbreviation) <sup>a</sup> [References]	Income level context	Healthcare level	Focus area	Sex or gender of beneficiaries of the targeted innovations	End-user of tool	Aim of tool	Degree of report of validity evidence for content validity <sup>b</sup>	Methodological quality of tools <sup>c</sup>
Burchett et al. [51]	Low- or middle-income country	Primary healthcare, public or population health	Reproductive, maternal, newborn, child or adolescent health	Not found	Not found	Assess a study's applicability and transferability	2	Inadequate
Cambon et al. [52]	Low-, middle- or high-income country	Public or population health	Health prevention or promotion	Not found	Not found	Guide and assess transferability	2	Inadequate
Process model for the assessment of transferability (PIET-T) [55]	Not found	Not found	Health prevention or promotion	Not found	Not found	Accompany the steps for determining transferability	2	Inadequate
Spicer et al. [56]	Low- or middle-income country	Primary healthcare	Reproductive, maternal, newborn, child or adolescent health	Not found	Researcher	Increase the prospects of government adoption and community uptake of innovations at scale	1	Inadequate
Wang et al. [58]	Not found	Public or population health	Not found	Not found	Not found	Assess applicability and transferability from a study setting to a local setting using evidence about both the local setting and the public health innovation of interest	1	Inadequate
Milat et al. [20, 21, 120]	High-income country	Public or population health	Health prevention or promotion	Not found	Policy-maker, researcher	Explore whether an innovation is scalable	2	Inadequate
CORRECT attributes <sup>d</sup> [6, 121, 122]	Low- or middle-income country	Primary healthcare, public or population health	Reproductive, maternal, newborn, child or adolescent health	Not found	Researcher, manager, funder	Assess the attributes that determine the scalability of the innovation and identify needed actions	2	Doubtful

**Table 3** (continued)

Name (abbreviation) <sup>a</sup> [References]	Income level context	Healthcare level	Focus area	Sex or gender of beneficiaries of the targeted innovations	End-user of tool	Aim of tool	Degree of report of validity evidence for content validity <sup>b</sup>	Methodological quality of tools <sup>c</sup>
Bhattacharyya et al. [49]	Transnational	Not found	Not found	Not found	Policy-maker, funder	Assess promising low- or middle-income country innovations for adaptation in high-income countries and identify those with high potential for more in-depth review and evaluation	3	Doubtful
Scalability assessment framework [57]	Low-, middle- or high-income country	Public or population health	Education, nutrition, sanitation, hygiene or international development more broadly	Female, male	Not found	Expand or replicate as part of a planned scaling-up process	2	Inadequate
Intervention Scalability Assessment Tool (ISAT) [22]	High-income country	Public or population health	Not found	Not found	Clinician, policy-maker	Assist practitioners, programme managers, and researchers to determine the scalability of a discrete health programme	3	Doubtful
Readiness to Spread Assessment Scoring Sheet [125]	High-income country	Not found	Not found	Not found	Manager	Help programme champions and leadership understand whether a promising practice is ripe for successful spread across organizations	1	Inadequate
Readiness to Receive Assessment Scoring Sheet [126]	High-income country	Not found	Not found	Not found	Manager	Help a site determine its readiness to receive an effective practice from elsewhere	1	Inadequate

**Table 3** (continued)

Name (abbreviation) <sup>a</sup> [References]	Income level context	Healthcare level	Focus area	Sex or gender of beneficiaries of the targeted innovations	End-user of tool	Aim of tool	Degree of report of validity evidence for content validity <sup>b</sup>	Methodological quality of tools <sup>c</sup>
Applicability and Transferability of Evidence Tool (A&T Tool) [118, 119]	Not found	Public or population health	Not found	Not found	Manager	Assist public health managers and planners in decision-making about programme priorities for their community	2	Doubtful
Scalability Assessment and Planning (SAP) Toolkit [130]	Not found	Not found	Reproductive, maternal, newborn, child or adolescent health	Not found	Not found	Guide scaling and assessment planning with corrective actions to strengthen or enable scale-up	1	Doubtful

<sup>a</sup> We did not find names for 10 of the tools, in which case we indicate names of authors or organizations

<sup>b</sup> We found no discussion or data presented as a source of validity evidence for the eight other measurement properties. 1 = Only a limited amount of data (e.g. simply listing items without justification); 2 = listing items with some references and justifications, limited description of the process for creating the tool; 3 = well-defined process for developing tool content, including both an explicit theoretical, conceptual or practical basis for the tool items and systematic item review by experts

<sup>c</sup> According to COSMIN definitions, a standard is rated as “doubtful” if it is doubtful whether the quality aspect is adequate (i.e. minor methodological flaws), and “inadequate” when evidence is provided that the quality aspect is not adequate (i.e. important methodological flaws)

<sup>d</sup> CORRECT attributes: C—credible in that they are based on sound evidence or advocated by respected persons or institutions; O—observable to ensure that potential users can see the results in practice; R—relevant for addressing persistent or sharply felt problems; R—relative advantage over existing practices so that potential users are convinced the costs of implementation are warranted by the benefits; E—easy to install and understand rather than complex and complicated; C—compatible with the potential users’ established values, norms and facilities; fit well into the practices of the national programme; and T—testable so that potential users can see the innovation on a small scale prior to large-scale adoption

Nine tools did not report on the focus area (42.9%) (Table 3). The largest proportion of tools for which this information was reported were created for innovations related to reproductive, maternal, newborn, child or adolescent health ( $n=7$ , 33.3%).

We found no information about intended end-users for 11 tools (52.4%) (Table 3). Tools for which this information was reported were intended for researchers, policy-makers, programme managers, healthcare providers or funders ( $n=10$ , 47.6%). No tool was created for lay end-users including patients or the public.

### Measurement properties of tools

All tools presented information for content validity, but most tools ( $n=18$ , 85.7%) provided limited information (e.g. simply listing items without justification, limited description of the process for creating the tool). Only three tools (14.3%) provided multiple sets of information that strongly supported content validity, such as descriptions and origins of constructs, or comprehensibility and comprehensiveness of items. No tool reported on the other measurement properties.

### Methodological quality of tools

According to COSMIN standards, the methodological quality of tools was deemed inadequate in 61.9% of cases ( $n=13$ ) and doubtful in 38.1% of cases ( $n=8$ ) (Table 3). The main reason was that design requirements were not met: for example, there was no clear description of the target population, context of use, or the tool's evaluative or predictive purpose.

### Discussion

We reviewed tools proposed for assessing the scalability of innovations in health. Altogether, identified tools targeted 11 scalability components and predicted eight pitfalls of scale-up. All included tools were created or published since 2005, but their methodological quality was inadequate or doubtful. No studies reported that patients or the public were involved in the creation or validation process of tools, and there was limited information on how the tools were intended to be used or on their intended end-users. These findings lead us to make the following observations.

*First*, all items found in the included tools were covered by our 11 defined scalability components, confirming that these classifications come close to reflecting the full range identified by others [22], and were enriched by items contributing to avoiding identified pitfalls such as replication of harms. Scalability assessment should ensure that innovations do not replicate social inequities when implemented at scale [15, 24, 131, 132]. For example, if the design of an innovation to be scaled up

was based on the male body as the norm [131], its scale-up could reproduce harmful outcomes at scale. This is the case with the conventional seat belt: Seat belts are not tested with pregnant women, and their design has undergone almost no changes since they were first patented in 1958 [133]. Yet car crashes are the main cause of foetal deaths related to maternal trauma. The forces of the seat belt against a pregnant woman's abdomen leads to placental abruption, causing foetal death [133]. More scalability assessments should also involve patients and the public [1, 14]. For example, members of the advisory committee, together with patient representatives and other stakeholders, could visit actual or potential sites to review arrangements for the project and to assess the potential for scale-up if the innovation proves successful. Discussion with providers, programme managers and community members could provide insights into how the project will be implemented on the ground and possible challenges and opportunities for scaling up, and could inspire reflection on possible adjustments to enhance its scalability [15, 124]. Certain scalability components could be less relevant for some innovations depending on the political circumstances, or on whether they are outcome evaluations under ideal circumstances (efficacy) or real-world circumstances (effectiveness) [2, 3, 15]. In addition, epidemics (e.g. COVID-19) have highlighted how dramatically scalability considerations can change when the world changes [1, 15].

*Second*, included tools were created or published since 2005, had inadequate or doubtful methodological quality, and most were of the "criteria" type. As key psychometric properties of these tools are yet to be established, for many of the tools there is still insufficient evidence to justify their claims. Future reviews involving the use of included tools should begin at the year 2005. Our results suggest that scalability assessment tools for health are still in their infancy. Previous studies confirm this, particularly in high-income countries [1, 22, 27, 134]. Indeed, the sophistication of our included tools varied from a simple list of items (i.e. criteria) to elaborate scales [135], although none of these had been validated [22, 27]. There were also important limitations in terms of sample representativity in the creation or validation of tool content. Intended context of use, for example, and content validity, the primary measurement property, were not fully addressed in most of the included tools [135]. However, we believe that content validation may increase over time as we learn more about the notion of scalability [136]. Nevertheless, for end-users wanting to adopt an existing tool or create a new one, we propose a useful inventory of items (Additional file 6). We also hope to create a repository of existing items whose language is accessible to

lay end-users, including patients and the public. This will contribute to increasing patient and public involvement in the science and practice of scale-up in health and social services [14].

*Third*, we noticed an absence of patient and public involvement in the creation of the scalability assessment tools. Patient perspectives are not only essential in innovation development; they are also important in the creation of scalability assessment tools [14, 15], asking the right questions and providing suggestions regarding items to include [135]. Although researchers, clinicians and policy-makers may be well positioned to describe the nature, scope and impact of a health problem that is being addressed, only those who experience the issues can report on the more subjective elements [135]. When appropriate, innovation teams have a responsibility to work with target patients to anticipate potential benefits and risks associated with scaling up, and to learn what risks they are willing to accept at each step of scale-up [15]. In practice, however, involving multiple stakeholders including patients and the public in the scalability assessments is a highly complex process [14, 15]. We have established the RePOS network to build patient-oriented research capacity in the science and practice of scaling up and ensure that patients, the public and other stakeholders are meaningfully and equitably engaged [14]. This international network will undertake the next phase of this review, conducting a multi-stakeholder consensus exercise to propose patient-oriented scalability assessment tools.

*Finally*, we acknowledge that our findings should be interpreted with caution. First, the interpretability criteria for what constitutes a useful item are not met by all items listed in our inventory (e.g. reading level, lack of ambiguity, asking only a single question) [135, 137]. However, at this early stage in the creation of scalability assessment tools, our interest is in creating an item pool. We aimed to be as inclusive as possible, even to the point of being overinclusive, as nothing can be done after the fact to compensate for items we neglected to include. Indeed, our research findings can be used to detect and weed out poor items using interpretability criteria proposed in the literature for item selection (Additional file 7) [135, 137, 138]. Second, characteristics of the innovation are important in scalability assessments, but there are other important, equally relevant assessments. Examples include comparing effects over time, namely at different stages of scale-up, so that innovations can be refined as coverage expands [27], and taking into account ongoing interactions between the innovation and its potential contexts [21, 23].

## Conclusions

We reviewed and inventoried tools proposed for assessing the scalability of innovations in health and described the scalability components they targeted. Overall, the included tools covered many components of scalability and helped predict the pitfalls of scale-up in health such as the replication of harms at scale. However, our findings show that these tools are still at an early stage of creation and their key psychometric properties are yet to be established. Scalability is a new concept, and as our understanding of this construct evolves, we will often need to revise tools accordingly. Our review may aid future investigators in weighting or prioritizing where planning and actions for scale-up should focus. Future studies could further compare and contrast the identified tools to illuminate the many perspectives on scale-up and the diverse approaches needed. Further analyses of our identified tools could also deepen understanding of how implementers, including patient partners, evaluate scalability components and how tools differ in their incorporation of evidence about acceptability. We also need to identify further scalability components, nuances of components already identified, and precisely how each scalability component contributes to the scale-up process.

## Abbreviations

ASTAIRE: Analyse de la Transférabilité et accompagnement à l'Adaptation des Interventions en pRomotion de la santé; A&T Tool: Applicability and Transferability of Evidence Tool; CBPHC: Community-Based Primary Health Care; CIHR: Canadian Institutes of Health Research; CORRECT: C—Credible in that they are based on sound evidence or advocated by respected persons or institutions; O—observable to ensure that potential users can see the results in practice; R—relevant for addressing persistent or sharply felt problems; R—relative advantage over existing practices so that potential users are convinced the costs of implementation are warranted by the benefits; E—easy to install and understand rather than complex and complicated; C—compatible with the potential users' established values, norms and facilities; fit well into the practices of the national programme; and T—testable so that potential users can see the innovation on a small scale prior to large-scale adoption; COSMIN: COnsensus-based Standards for the selection of health Measurement INstruments; COVID-19: Coronavirus disease 2019; ISAT: Intervention Scalability Assessment Tool; ISSaQ: Innovation Scalability Self-administered Questionnaire; PIET-T: Process model for the assessment of transferability; PRESS: Peer Review of Electronic Search Strategies; PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses; PROSPERO: International Prospective Register of Systematic Reviews; RePOS: Research on Patient-Oriented Scaling-up; SAP: Scalability Assessment and Planning.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12961-022-00830-5>.

**Additional file 1.** Search strategy.

**Additional file 2.** List of relevant websites used to identify potential eligible records.

**Additional file 3.** Email sent to experts to identify potential eligible records.

**Additional file 4.** Records selection process.

**Additional file 5.** List of excluded reports with reason for exclusion.

**Additional file 6.** List of identified items.

**Additional file 7.** Interpretability criteria for selecting items.

### Acknowledgements

We wish to acknowledge the following persons for their dedicated assistance with various aspects of this systematic review: Dr Arlene Bierman, M. Frédéric Bergeron, Dr Louisa Blair and Dr Annie LeBlanc. Also, we thank Dr Louisa Blair, English-language editor, for her kind help with the manuscript.

### Authors' contributions

ABC and FL participated in the conception of the study. ABC, HTVZ, AG, JM, LW, JP, MZ, AJM, NR, JS, MMR, FL and A. Bierman participated in the design of the study. ABC, HTVZ, LW, JP, MZ, AJM, NR, JS, MMR, L. Blair and F. Bergeron contributed to the development of the search strategy. ABC, AG, MAS, JM and YMO contributed to the selection of records or reports. ABC, HTVZ, AG, MAS, JM and YMO contributed to the data extraction. ABC and MAS contributed to the data analysis. ABC, HTVZ, AG, MAS, JM, YMO, FL and A. LeBlanc contributed to the interpretation of data. ABC drafted the manuscript. All authors revised the manuscript critically for important intellectual content and agreed to be accountable for all aspects of the study. All authors read and approved the final manuscript.

### Funding

Our systematic review was funded by the following Canadian Institutes of Health Research (CIHR) grants: (1) the Unité de soutien SSA Québec (#SU1-139759), (2) a Catalyst Grant (#PAO-169411) and (3) a Foundation Grant (#FDN-159931). ABC was supported by CubecXpert and the Fonds de recherche en santé du Québec—Santé (FRQ-S). AG was supported by the CIHR. LW is funded by a NHMRC Career Development and Heart Foundation Future Leaders Fellowship. MMR holds a Canada Research Chair in Person-Centred Interventions for Older Adults with Multimorbidity and their Caregivers. FL holds a Tier 1 Canada Research Chair in Shared Decision Making and Knowledge Translation. The funding agreement ensured the authors' independence in designing the study, interpreting the data, and writing and publishing the article. The information provided or views expressed in this article are the responsibility of the authors alone.

### Availability of data and materials

Please send all requests for study data or materials to Dr Ali Ben Charif (ali.bencharif@gmail.com) or Dr France Légaré (france.legare@mfa.ulaval.ca).

### Declarations

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>CubecXpert, Quebec City, QC, Canada. <sup>2</sup>Department of Social and Preventive Medicine, Université Laval, Quebec City, QC, Canada. <sup>3</sup>Faculty of Medicine and Health Science, School of Physical and Occupational Therapy, McGill University, Montreal, QC, Canada. <sup>4</sup>Institut national d'excellence en santé et en services sociaux (INESSS), Quebec City, QC, Canada. <sup>5</sup>VITAM – Centre de recherche en santé durable, Université Laval, Quebec City, QC, Canada. <sup>6</sup>Tier 1 Canada Research Chair in Shared Decision Making and Knowledge Translation, Université Laval, Quebec City, QC, Canada. <sup>7</sup>Department of Family Medicine and Emergency Medicine, Université Laval, Quebec City, QC, Canada. <sup>8</sup>Unité de soutien SSA Québec, Université Laval, Quebec City, QC, Canada. <sup>9</sup>Institut national de santé publique du Québec (INSPQ), Quebec City, QC, Canada. <sup>10</sup>School of Medicine and Public Health, University of Newcastle, Callaghan, NSW, Australia. <sup>11</sup>Hunter Medical Research Institute, New Lambton Heights, NSW, Australia. <sup>12</sup>Hunter New England Population Health, Wallsend, NSW,

Australia. <sup>13</sup>Aging, Community and Health Research Unit, School of Nursing, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada. <sup>14</sup>Department of Family Medicine, Centre for Studies in Family Medicine, Schulich School of Medicine and Dentistry, Western University, London, ON, Canada. <sup>15</sup>School of Public Health, University of Sydney, Sydney, NSW, Australia. <sup>16</sup>Centre for Epidemiology and Evidence, NSW Ministry of Health, Sydney, Australia. <sup>17</sup>Santé publique France, Paris, France. <sup>18</sup>Canada Research Chair in Person Centred Interventions for Older Adults with Multimorbidity and their Caregivers, McMaster University, Hamilton, ON, Canada. <sup>19</sup>Population Health and Practice-Changing Research Group, CHU de Québec Research Centre, Quebec City, QC, Canada.

Received: 26 May 2021 Accepted: 16 February 2022

Published online: 24 March 2022

### References

1. Ben Charif A, Zomahoun HTV, LeBlanc A, Langlois L, Wolfenden L, Yoong SL, et al. Effective strategies for scaling up evidence-based practices in primary care: a systematic review. *Implement Sci.* 2017;12:139.
2. Ben Charif A, Hassani K, Wong ST, Zomahoun HTV, Fortin M, Freitas A, et al. Assessment of scalability of evidence-based innovations in community-based primary health care: a cross-sectional study. *CMAJ Open.* 2018;6:E520–7.
3. Ben Charif A, Zomahoun HTV, Massougbodji J, Khadhraoui L, Pilon MD, Boulanger E, et al. Assessing the scalability of innovations in primary care: a cross-sectional study. *CMAJ Open.* 2020;8:E613–8.
4. Greenhalgh T, Robert G, Bate P, Macfarlane F, Kyriakidou O. Diffusion of innovations in health service organisations: a systematic literature review. Chichester: Wiley; 2005. <https://doi.org/10.1002/9780470987407>.
5. Rogers EM. Diffusion of innovations. 5th ed. New York: Free Press; 2003.
6. World Health Organization (WHO). Nine steps for developing a scaling-up strategy. 2010. [http://www.who.int/reproductivehealth/publications/strategic\\_approach/9789241500319/en/](http://www.who.int/reproductivehealth/publications/strategic_approach/9789241500319/en/). Accessed 26 Apr 2018.
7. Bégin M, Eggertson L, Macdonald N. A country of perpetual pilot projects. *CMAJ.* 2009;180(1185):E88–9.
8. Fixsen D, Blase K, Metz A, Van Dyke M. Statewide implementation of evidence-based programs. *Except Child.* 2013;79:213–30.
9. Sheridan DJ. Research: increasing value, reducing waste. *Lancet.* 2014;383:1123.
10. Whitworth J, Sewankambo NK, Snewin VA. Improving implementation: building research capacity in maternal, neonatal, and child health in Africa. *PLOS Med.* 2010;7: e1000299.
11. Bauer MS, Kirchner J. Implementation science: what is it and why should I care? *Psychiatry Res.* 2020;283: 112376.
12. Straus S, Tetroe J, Graham ID. Knowledge translation in health care: moving from evidence to practice. 2nd ed. Chichester: Wiley-Blackwell; 2013.
13. Haynes A, Rychetnik L, Finegood D, Irving M, Freebairn L, Hawe P. Applying systems thinking to knowledge mobilisation in public health. *Health Res Policy Syst.* 2020;18:134.
14. Ben Charif A, Plourde KV, Guay-Bélanger S, Zomahoun HTV, Gogovor A, Straus 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;10:55.
15. McLean R, Gargani J. Scaling impact: innovation for the public good. 1st ed. London: Routledge; 2019.
16. Indig D, Lee K, Grunseit A, Milat A, Bauman A. Pathways for scaling up public health interventions. *BMC Public Health.* 2017;18:68.
17. DePasse JW, Lee PT. A model for 'reverse innovation' in health care. *Glob Health.* 2013;9:40.
18. Hadengue M, de Marcellis-Warin N, Warin T. Reverse innovation: a systematic literature review. *Int J Emerg Mark.* 2017;12:142–82.
19. Shaw J, Tepper J, Martin D. From pilot project to system solution: innovation, spread and scale for health system leaders. *BMJ Lead.* 2018. <https://doi.org/10.1136/leader-2017-000055>.
20. Milat AJ, King L, Bauman AE, Redman S. The concept of scalability: increasing the scale and potential adoption of health promotion interventions into policy and practice. *Health Promot Int.* 2013;28:285–98.

21. Milat AJ, Newson R, King L, Rissel C, Wolfenden L, Bauman A, et al. A guide to scaling up population health interventions. *Public Health Res Pract.* 2016;26: e2611604.
22. Milat AJ, Lee K, Conte K, Grunseit A, Wolfenden L, van Nassau F, et al. Intervention scalability assessment tool: a decision support tool for health policy makers and implementers. *Health Res Policy Syst.* 2020;18:1.
23. Greenhalgh T, Papoutis C. Spreading and scaling up innovation and improvement. *BMJ.* 2019;365: I2068.
24. Zomahoun HTV, Ben Charif A, Freitas A, Garvelink MM, Meneer M, Dugas M, et al. The pitfalls of scaling up evidence-based interventions in health. *Glob Health Action.* 2019;12:1670449.
25. Glassman A, Temin M. Millions saved: new cases of proven success in global health. Baltimore: Center for Global Development; 2016.
26. Rottach E, Hardee K, Jolivet R, Kiesel R. Integrating gender into the scale-up of family planning and maternal neonatal and child health programs. Washington, DC: Futures Group, Health Policy Project; 2012. <https://www.semanticscholar.org/paper/Integrating-gender-into-the-scale-up-of-family-and-Rottach-Hardee/073567bc408feabcedcb06ce433c8e3248962f7c>.
27. Zamboni K, Schellenberg J, Hanson C, Betran AP, Dumont A. Assessing scalability of an intervention: why, how and who? *Health Policy Plan.* 2019;34:544–52.
28. Mokkink LB, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, de Vet HCW, et al. COSMIN methodology for systematic reviews of patient-reported outcome measures (PROMs). Amsterdam: VU University Medical Center; 2018. p. 78.
29. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ.* 2021;372:n71.
30. Ben Charif A, Zomahoun HTV, Gogovor A, Rheault N, Ploeg J, Wolfenden L, et al. Tools for assessing the scalability of health innovations: a systematic review. PROSPERO. 2019;CRD42019107095. [http://www.crd.york.ac.uk/PROSPERO/display\\_record.php?ID=CRD42019107095](http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42019107095).
31. World Health Organization (WHO). International classification of health interventions (ICHI). Geneva: WHO. <http://www.who.int/classifications/ichi/en/>. Accessed 26 Apr 2018.
32. Hovland I. Successful communication: a toolkit for researchers and civil society organisations. 2005. <https://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/192.pdf>. Accessed 1 Mar 2021.
33. Canadian Institutes of Health Research (CIHR). Knowledge user engagement. 2012. <https://cihr-irsc.gc.ca/e/49505.html>. Accessed 1 Mar 2021.
34. Pomey M-P, Flora L, Karazivan P, Dumez V, Lebel P, Vanier M-C, et al. The Montreal model: the challenges of a partnership relationship between patients and healthcare professionals. *Sante Publique.* 2015;27:541–50.
35. Carman KL, Dardess P, Maurer M, Sofaer S, Adams K, Bechtel C, et al. Patient and family engagement: a framework for understanding the elements and developing interventions and policies. *Health Aff.* 2013;32:223–31.
36. Unité de soutien SRAP du Québec. Unité soutien SRAP. <http://unitesoutiensrapqc.ca/>. Accessed 19 Aug 2019.
37. Moore ZE, Cowman S. Risk assessment tools for the prevention of pressure ulcers. *Cochrane Database Syst Rev.* 2014. <https://doi.org/10.1002/14651858.CD006471.pub3/full>.
38. Otten R. Health. In: Search blocks/Zoekblokken. Biomedical information of the Dutch Library Association (KNVI). 2012. <https://blocks.bmi-online.nl/catalog/141>. Accessed 3 Feb 2021.
39. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS peer review of electronic search strategies: 2015 guideline statement. *J Clin Epidemiol.* 2016;75:40–6.
40. Agency for Healthcare Research and Quality (AHRQ). Methods guide for effectiveness and comparative effectiveness reviews. Rockville: Agency for Healthcare Research and Quality (US); 2008. <http://www.ncbi.nlm.nih.gov/books/NBK47095/>. Accessed 28 Aug 2019.
41. Wong ST, Langton JM, Katz A, Fortin M, Godwin M, Green M, et al. Promoting cross-jurisdictional primary health care research: developing a set of common indicators across 12 community-based primary health care teams in Canada. *Prim Health Care Res Dev.* 2018;20. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6476395/>. Accessed 28 Nov 2019.
42. Cohen J. Weighted kappa: nominal scale agreement provision for scaled disagreement or partial credit. *Psychol Bull.* 1968;70:213–20.
43. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159–74.
44. Ghaderi I, Manji F, Park YS, Juul D, Ott M, Harris I, et al. Technical skills assessment toolbox: a review using the unitary framework of validity. *Ann Surg.* 2015;261:251–62.
45. Terwee CB, Prinsen CA, Chiarotto A, de Vet H, Bouter LM, Alonso J, et al. COSMIN methodology for assessing the content validity of PROMs. Amsterdam: VU University Medical Center; 2018. p. 73.
46. Patrick DL, Burke LB, Gwaltney CJ, Leidy NK, Martin ML, Molsen E, et al. Content validity—establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: part 1—eliciting concepts for a new PRO instrument. *Value Health.* 2011;14:967–77.
47. Baker PRA, Shipp JJ, Wellings SH, Priest N, Francis DP. Assessment of applicability and transferability of evidence-based antenatal interventions to the Australian indigenous setting. *Health Promot Int.* 2012;27:208–19.
48. Bennett S, Mahmood SS, Edward A, Tetui M, Ekirapa-Kiracho E. Strengthening scaling up through learning from implementation: comparing experiences from Afghanistan, Bangladesh and Uganda. *Health Res Policy Syst.* 2017;15:108.
49. Bhattacharyya O, Wu D, Mossman K, Hayden L, Gill P, Cheng Y-L, et al. Criteria to assess potential reverse innovations: opportunities for shared learning between high- and low-income countries. *Glob Health.* 2017;13:4.
50. Burchett H, Umoquit M, Dobrow M. How do we know when research from one setting can be useful in another? A review of external validity, applicability and transferability frameworks. *J Health Serv Res Policy.* 2011;16:238–44.
51. Burchett HED, Mayhew SH, Lavis JN, Dobrow MJ. When can research from one setting be useful in another? Understanding perceptions of the applicability and transferability of research. *Health Promot Int.* 2013;28:418–30.
52. Cambon L, Minary L, Ridde V, Alla F. Transferability of interventions in health education: a review. *BMC Public Health.* 2012;12:497.
53. Cambon L, Minary L, Ridde V, Alla F. A tool to analyze the transferability of health promotion interventions. *BMC Public Health.* 2013;13:1184.
54. Cambon L, Minary L, Ridde V, Alla F. Un outil pour accompagner la transférabilité des interventions en promotion de la santé: ASTAIRE, vol. 26. *Sante Publique. S.F.S.P.*; 2014. p. 783–6.
55. Schloemer T, Schröder-Bäck P. Criteria for evaluating transferability of health interventions: a systematic review and thematic synthesis. *Implement Sci.* 2018;13:88.
56. Spicer N, Bhattacharya D, Dimka R, Fanta F, Mangham-Jefferies L, Schellenberg J, et al. "Scaling-up is a craft not a science": catalysing scale-up of health innovations in Ethiopia, India and Nigeria. *Soc Sci Med.* 2014;121:30–8.
57. Vaughan-Lee H, Moriniere LC, Bremaud I, Turnbull M. Understanding and measuring scalability in disaster risk reduction. *Disaster Prevent Manag Int J.* 2018;27:407–20.
58. Wang S, Moss JR, Hiller JE. Applicability and transferability of interventions in evidence-based public health. *Health Promot Int.* 2006;21:76–83.
59. Spicer N, Hamza YA, Berhanu D, Gautham M, Schellenberg J, Tadesse F, et al. 'The development sector is a graveyard of pilot projects!' Six critical actions for externally funded implementers to foster scale-up of maternal and newborn health innovations in low and middle-income countries. *Glob Health.* 2018;14:74.
60. Spicer N, Berhanu D, Bhattacharya D, Tilley-Gyado RD, Gautham M, Schellenberg J, et al. 'The stars seem aligned': a qualitative study to understand the effects of context on scale-up of maternal and newborn health innovations in Ethiopia, India and Nigeria. *Glob Health.* 2016;12:75.
61. Simpson DD. A conceptual framework for transferring research to practice. *J Subst Abuse Treat.* 2002;22:171–82.
62. Vanderkruik R, McPherson ME. A contextual factors framework to inform implementation and evaluation of public health initiatives. *Am J Eval.* 2017;38:348–59.

63. Houck JM, Manuel JK, Pyeatt CJ, Moyers TB, Montoya VL. A new scale to assess barriers to adopting motivational interviewing. *Alcohol Clin Exp Res*. 2009;33:137A.
64. Renju J, Andrew B, Nyalali K, Kishamawe C, Kato C, Changanlucha J, et al. A process evaluation of the scale up of a youth-friendly health services initiative in northern Tanzania. *J Int AIDS Soc*. 2010;13:32–32.
65. Davies BM, Smith J, Rikabi S, Wartolowska K, Morrey M, French A, et al. A quantitative, multi-national and multi-stakeholder assessment of barriers to the adoption of cell therapies. *J Tissue Eng*. 2017;8:2041731417724413.
66. Huang TT-K, Grimm B, Hammond RA. A systems-based typological framework for understanding the sustainability, scalability, and reach of childhood obesity interventions. *Children's Health Care*. 2011;40:253–66.
67. Toobert DJ, Glasgow RE, Strycker LA, Barrera M Jr, King DK. Adapting and RE-AIMing a heart disease prevention program for older women with diabetes. *Transl Behav Med*. 2012;2:180–7.
68. Flynn R, McDonnell P, Mackenzie I, Wei L, MacDonald T. An efficient and inexpensive system for the distribution and tracking of investigational medicinal products in streamlined safety trials. *Pharmacoepidemiol Drug Saf*. 2012;21:124–124.
69. Grooten L, Borgermans L, Vrijhoef H. An instrument to measure maturity of integrated care: a first validation study. *Int J Integr Care*. 2018;18:10.
70. Leon N, Schneider H, Daviaud E. Applying a framework for assessing the health system challenges to scaling up mHealth in South Africa. *BMC Med Inform Decis Mak*. 2012;12:123.
71. Grooten L, Vrijhoef HJM, Calciolari S, Ortiz LGG, Janečková M, Minkman MMN, et al. Assessing the maturity of the healthcare system for integrated care: testing measurement properties of the SCIROCCO tool. *BMC Med Res Methodol*. 2019;19:63.
72. Berger KB, Otto-Salaj LL, Stoffel VC, Hernandez-Meier J, Gromoske AN. Barriers and facilitators of transferring research to practice: an exploratory case study of motivational interviewing. *J Soc Work Pract Addict*. 2009;9:145–62.
73. Pérez-Escamilla R, Hromi-Fiedler AJ, Gubert MB. Becoming baby friendly: a complex adaptive systems toolbox for scaling up breastfeeding programs globally. *FASEB J*. 2017;31:165–8.
74. Pérez-Escamilla R, Hromi-Fiedler AJ, Gubert MB, Doucet K, Meyers S, dos Santos Buccini G. Becoming breastfeeding friendly index: development and application for scaling-up breastfeeding programmes globally. *Matern Child Nutr*. 2018;14: e12596.
75. Rosas SR, Behar LB, Hydaker WM. Community readiness within systems of care: the validity and reliability of the system of care readiness and implementation measurement scale (SOC-RIMS). *J Behav Health Serv Res*. 2016;43:18–37.
76. Schloemer T, Schroder-Back P. Criteria for evaluating transferability of child health interventions: a systematic review. *Eur J Public Health*. 2017;27:501–501.
77. Carr H, de Lusignan S, Liyanage H, Liaw S-T, Terry A, Rafi I. Defining dimensions of research readiness: a conceptual model for primary care research networks. *BMC Fam Pract*. 2014;15:1–10.
78. Nguyen K, Schmidt T, Ozcelik S, Bell J, Whitney RL. Development of a serious illness care model implementation framework. *JCO*. 2017;35:161–161.
79. Diwan S. Differences in the scalability of formal and informal in-home care of urban elderly. *Home Health Care Serv Q*. 1994;14:105–16.
80. Brink SG, Basen-Engquist KM, O'Hara-Tompkins NM, Parcel GS, Gottlieb NH, Lovato CY. Diffusion of an effective tobacco prevention program. Part I: evaluation of the dissemination phase. *Health Educ Res*. 1995;10:283–95.
81. Aamir J, Ali SM, Kamel Boulos MN, Anjum N, Ishaq M. Enablers and inhibitors: a review of the situation regarding mHealth adoption in low- and middle-income countries. *Health Policy Technol*. 2018;7:88–97.
82. Sako B, Leerlooijer JN, Lelisa A, Hailemariam A, Brouwer ID, Brown AT, et al. Exploring barriers and enablers for scaling up a community-based grain bank intervention for improved infant and young child feeding in Ethiopia: a qualitative process evaluation. *Matern Child Nutr*. 2018;14: e12551.
83. Ezeanochie NP. Exploring the adoption and scale of mobile health solutions: antenatal mobile application data exchange in Nigeria [Ph.D. Thesis]. University of Colorado at Denver; 2017.
84. Maar M, Yeates K, Barron M, Hua D, Liu P, Moy Lum-Kwong M, et al. I-RREACH: an engagement and assessment tool for improving implementation readiness of researchers, organizations and communities in complex interventions. *Implement Sci*. 2015;10:64.
85. Agbakoba R, McGee-Lennon M, Bouamrane MM, Watson N, Mair FS. Implementation factors affecting the large-scale deployment of digital health and well-being technologies: a qualitative study of the initial phases of the "Living-It-Up" programme. *Health Inform J*. 2016;22:867–77.
86. Southon FCG, Sauer C, Dampney CNG. Information technology in complex health services: organizational impediments to successful technology transfer and diffusion. *J Am Med Inform Assoc*. 1997;4:112–24.
87. Gylmark M, Lampe K, Ruof J, Pöhlmann J, Hebborn A, Kristensen FB. Is the EUnetHTA HTA Core Model® fit for purpose? Evaluation from an industry perspective. *Int J Technol Assess Health Care*. 2018;34:458–63.
88. Carlford S, Lindberg M, Bendtsen P, Nilsen P, Andersson A. Key factors influencing adoption of an innovation in primary health care: a qualitative study based on implementation theory. *BMC Fam Pract*. 2010;11:60.
89. Kliche T, Post M, Pftzner R, Plaumann M, Dubben S, Nöcker G, et al. Knowledge transfer methods in German disease prevention and health promotion. A survey of experts in the federal prevention research program. *Gesundheitswesen*. 2012;74:240–9.
90. Forkuo-Minka A. Knowledge transfer: theoretical framework to systematically spread best practice. *Br J Sch Nurs*. 2018;13:26–35.
91. Perla RJ, Bradbury E, Gunther-Murphy C. Large-scale improvement initiatives in healthcare: a scan of the literature. *J Healthc Qual*. 2013;35:30–40.
92. Grooten L, Borgermans L, Vrijhoef H. Measuring maturity of integrated care: a first validation study. *Int J Integr Care*. 2017;17:A412.
93. Joly BM, Booth M, Mittal P, Shaler G. Measuring quality improvement in public health: the development and psychometric testing of a QI maturity tool. *Eval Health Prof*. 2012;35:119–47.
94. Fogerty R, Sankey C, Kenyon K, Sussman S, Sigurdsson S, Klinger A. Pilot of a low-resource, EHR-based protocol for sepsis monitoring, alert, and intervention. *J Gen Intern Med*. 2016;31:902–3.
95. Fischer S, Itoh M, Inagaki T. Prior schemata transfer as an account for assessing the intuitive use of new technology. *Appl Ergon*. 2015;46(Pt A):8–20.
96. Shamu S, Rusakaniko S, Hongoro C. Prioritizing health system and disease burden factors: an evaluation of the net benefit of transferring health technology interventions to different districts in Zimbabwe. *Clinicoecon Outcomes Res*. 2016;8:695–705.
97. Egeland KM, Ruud T, Ogden T, Lindstrøm JC, Heiervang KS. Psychometric properties of the Norwegian version of the evidence-based practice attitude scale (EBPAS): to measure implementation readiness. *Health Res Policy Syst*. 2016;14:47.
98. Hill LG, Cooper BR, Parker LA. Qualitative comparative analysis: a mixed-method tool for complex implementation questions. *J Primary Prevent*. 2019;40:69–87.
99. Nguyen DTK. Scaling up [to] a population health intervention: a readiness assessment framework. *Dissertation Abstracts International: Section B: The Sciences and Engineering*. 2017;78.
100. Sako B, Leerlooijer JN, Lelisa A, Hailemariam A, Brouwer I, Tucker Brown A, et al. Scaling up a community-based grainbank intervention for improved infant and young child feeding (IYCF) in Ethiopia. *Ann Nutr Metab*. 2017;71:834–5.
101. Renju J, Nyalali K, Andrew B, Kishamawe C, Kimario M, Remes P, et al. Scaling up a school-based sexual and reproductive health intervention in rural Tanzania: a process evaluation describing the implementation realities for the teachers. *Health Educ Res*. 2010;25:903–16.
102. Pednekar MS, Nagler EM, Gupta PC, Pawar PS, Mathur N, Adhikari K, et al. Scaling up a tobacco control intervention in low resource settings: a case example for school teachers in India. *Health Educ Res*. 2018;33:218–31.
103. Werfel Th, Hellings P, Agache I, Muraro A, Bousquet J. Scaling up strategies of the chronic respiratory disease programme of the European innovation partnership on active and healthy ageing—executive summary. *Alergologia Polska/Pol J Allergol*. 2017;4:2–6.

104. Bacci JL, Coley KC, McGrath K, Abraham O, Adams AJ, McGivney MS. Strategies to facilitate the implementation of collaborative practice agreements in chain community pharmacies. *J Am Pharm Assoc*. 2003;2016(56):257–265.e2.
105. Kastner M, Sayal R, Oliver D, Straus SE, Dolovich L. Sustainability and scalability of a volunteer-based primary care intervention (Health TAP-ESTRY): a mixed-methods analysis. *BMC Health Serv Res*. 2017;17:514.
106. Rosella L, Bowman C, Pach B, Morgan S, Fitzpatrick T, Goel V. The development and validation of a meta-tool for quality appraisal of public health evidence: meta quality appraisal tool (MetaQAT). *Public Health*. 2016;136:57–65.
107. Upton D, Upton P, Scurlock-Evans L. The reach, transferability, and impact of the evidence-based practice questionnaire: a methodological and narrative literature review. *Worldviews Evid Based Nurs*. 2014;11:46–54.
108. Sani NSBM, Arshad NIB. Towards a framework to measure knowledge transfer in organizations. New York; 2015.
109. Goeree R, He J, O'Reilly D, Tarride J-E, Xie F, Lim M, et al. Transferability of health technology assessments and economic evaluations: a systematic review of approaches for assessment and application. *Clinicoecon Outcomes Res*. 2011;3:89–104.
110. Essers BAB, Seferina SC, Tjan-Heijnen VCG, Severens JL, Novák A, Pompen M, et al. Transferability of model-based economic evaluations: the case of trastuzumab for the adjuvant treatment of HER2-positive early breast cancer in the Netherlands. *Value Health*. 2010;13:375–80.
111. Kaló Z, Landa K, Doležal T, Vokó Z. Transferability of national institute for health and clinical excellence recommendations for pharmaceutical therapies in oncology to Central-Eastern European countries. *Eur J Cancer Care*. 2012;21:442–9.
112. O'Hara BJ, Phongsavan P, King L, Develin E, Milat AJ, Eggins D, et al. "Translational formative evaluation": critical in up-scaling public health programmes. *Health Promot Int*. 2014;29:38–46.
113. Rohrbach LA, Grana R, Sussman S, Valente TW. Type II translation: transporting prevention interventions from research to real-world settings. *Eval Health Prof*. 2006;29:302–33.
114. Marshall M, Mountford J, Gamet K, Gungor G, Burke C, Hudson R, et al. Understanding quality improvement at scale in general practice: a qualitative evaluation of a COPD improvement programme. *Br J Gen Pract*. 2014;64:e745–751.
115. Morgenthaler TI, Lovely JK, Cima RR, Berardinelli CF, Fedraw LA, Wallicher TJ, et al. Using a framework for spread of best practices to implement successful venous thromboembolism prophylaxis throughout a large hospital system. *Am J Med Qual*. 2012;27:30–8.
116. Madden L, Bojko MJ, Farnum S, Mazhnaya A, Fomenko T, Marcus R, et al. Using nominal group technique among clinical providers to identify barriers and prioritize solutions to scaling up opioid agonist therapies in Ukraine. *Int J Drug Policy*. 2017;49:48–53.
117. Miller J. Winning big. *Pharm Technol*. 2010;34:52–4.
118. Buffett C, Ciliska D, Thomas H. Can I use this evidence in my program decision? Assessing applicability and transferability of evidence. Hamilton, ON, Canada; 2007. <https://www.nccmt.ca/uploads/media/media/0001/01/110008a2754f35048bb7e8ff446117133b81ab13.pdf>. Accessed 21 Aug 2020.
119. Buffett C, Ciliska D, Thomas H. Évaluation de l'applicabilité et de la transférabilité des données probantes: Puis-je utiliser ces données probantes dans mes décisions de programmes? Hamilton, ON, Canada; 2007. <https://www.nccmt.ca/uploads/media/media/0001/01/ea0f35a0458f84bce52deabc21c4a57ff6a818f6.pdf>. Accessed 21 Aug 2020.
120. Milat AJ, Newson R, King L. Increasing the scale of population health interventions: a guide. Evidence and Evaluation guidance series, population and public health division. Sydney: NSW Ministry of Health, Centre for Epidemiology and Evidence; 2014. <https://www.health.nsw.gov.au/research/Pages/scalability-guide.aspx>. Accessed 19 Aug 2020.
121. Organización mundial de la salud (OMS). Nueve pasos para formular una estrategia de ampliación a escala. Geneva: World Health Organization; 2011. [https://www.who.int/reproductivehealth/publications/strategic\\_approach/9789241500319/es/](https://www.who.int/reproductivehealth/publications/strategic_approach/9789241500319/es/). Accessed 19 Aug 2020.
122. Organisation mondiale de la santé (OMS). Neuf étapes pour élaborer une stratégie de passage à grande échelle. 2011. [https://www.who.int/reproductivehealth/publications/strategic\\_approach/9789241500319/fr/](https://www.who.int/reproductivehealth/publications/strategic_approach/9789241500319/fr/). Accessed 19 Aug 2020.
123. Organisation mondiale de la santé (OMS). Avoir le but à l'esprit dès le début : la planification des projets pilotes et d'autres recherches programmatiques pour un passage à grande échelle réussi. World Health Organization, Department of Reproductive Health and Research—ExpandNet; 2013. [https://www.who.int/reproductivehealth/publications/strategic\\_approach/9789241502320/fr/](https://www.who.int/reproductivehealth/publications/strategic_approach/9789241502320/fr/). Accessed 19 Aug 2020.
124. World Health Organization (WHO). Beginning with the end in mind: planning pilot projects and other programmatic research for successful scaling up. Geneva: World Health Organization; 2011. [https://www.who.int/reproductivehealth/publications/strategic\\_approach/9789241502320/en/](https://www.who.int/reproductivehealth/publications/strategic_approach/9789241502320/en/). Accessed 19 Aug 2020.
125. Canadian Foundation for Healthcare Improvement (CFHI). Readiness to spread assessment. 2013. [https://www.cfhi-fcass.ca/docs/default-source/itr/tools-and-resources/cfhi-readiness-to-spread-assessment-tool-e.pdf?sfvrsn=d2c69479\\_2](https://www.cfhi-fcass.ca/docs/default-source/itr/tools-and-resources/cfhi-readiness-to-spread-assessment-tool-e.pdf?sfvrsn=d2c69479_2). Accessed 21 Aug 2020.
126. Canadian Foundation for Healthcare Improvement (CFHI). Readiness to receive assessment. 2013. <https://www.cfhi-fcass.ca/docs/default-source/itr/tools-and-resources/cfhi-readiness-to-receive-assessment-tool-e.pdf>. Accessed 21 Aug 2020.
127. Cooley L, Kohl R. Scaling up—from vision to large-scale change: a management framework for practitioners. 1st ed. Washington, DC: Management Systems International; 2006.
128. Cooley L, Ved RR. Scaling up—from vision to large-scale change: a management framework for practitioners. 2nd ed. Washington, DC: Management Systems International; 2012.
129. Cooley L, Kohl R, Ved RR. Scaling up—from vision to large-scale change: a management framework for practitioners. 3rd ed. Washington, DC: Management Systems International; 2016.
130. Morinière LC, Turnbull M, Bremaud I, Vaughan-Lee H, Xaxa V, Farheen SA. Toolkit: scalability assessment and planning (SAP) (including workshop guidance). 2018. <https://resourcecentre.savethechildren.net/library/scalability-assessment-and-planning-sap-toolkit>. Accessed 21 Aug 2020.
131. Attwood S, van Sluijs E, Sutton S. Exploring equity in primary-care-based physical activity interventions using PROGRESS-Plus: a systematic review and evidence synthesis. *Int J Behav Nutr Phys Act*. 2016;13:60.
132. O'Neill J, Tabish H, Welch V, Petticrew M, Pottie K, Clarke M, et al. Applying an equity lens to interventions: using PROGRESS ensures consideration of socially stratifying factors to illuminate inequities in health. *J Clin Epidemiol*. 2014;67:56–64.
133. Briones Panadero H. Analysis of the design of a car seatbelt : a study of the invention and a proposal to minimize the risk of injuries during pregnancy [Thesis]. Massachusetts Institute of Technology; 2020. <https://hdl.handle.net/1721.1/132804>. Accessed 23 Dec 2021.
134. Milat AJ, Bauman A, Redman S. Narrative review of models and success factors for scaling up public health interventions. *Implement Sci*. 2015;10:113.
135. Streiner DL, Norman GR, Cairney J. Health measurement scales: a practical guide to their development and use. Oxford: Oxford University Press; 2014. <https://doi.org/10.1093/med/9780199685219.001.0001/med-9780199685219>.
136. Haynes SN, Richard DCS, Kubany ES. Content validity in psychological assessment: a functional approach to concepts and methods. *Psychol Assess*. 1995;7:238–47.
137. Patrick DL, Burke LB, Gwaltney CJ, Leidy NK, Martin ML, Molsen E, et al. Content validity—establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: part 2—assessing respondent understanding. *Value Health*. 2011;14:978–88.
138. Peasgood T, Mukuria C, Carlton J, Connell J, Brazier J. Criteria for item selection for a preference-based measure for use in economic evaluation. *Qual Life Res*. 2020. <https://doi.org/10.1007/s11366-020-02718-9>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.