

RESEARCH

Open Access



Using mixed methods and partnership to develop a program evaluation toolkit for organizations that provide physical activity programs for persons with disabilities

Sarah V. C. Lawrason^{1,2*}, Pinder DaSilva³, Emilie Michalovic³, Amy Latimer-Cheung^{4,5}, Jennifer R. Tomasone^{4,5}, Shane Sweet^{6,7}, Tanya Forneris¹, Jennifer Leo⁸, Matthew Greenwood⁹, Janine Giles¹⁰, Jane Arkell¹¹, Jackie Patatas¹², Nick Boyle¹⁰, Nathan Adams^{1,2} and Kathleen A. Martin Ginis^{1,2,13}

Abstract

Background The purpose of this paper is to report on the process for developing an online RE-AIM evaluation toolkit in partnership with organizations that provide physical activity programming for persons with disabilities.

Methods A community-university partnership was established and guided by an integrated knowledge translation approach. The four-step development process included: (1) identify, review, and select knowledge (literature review and two rounds of Delphi consensus-building), (2) adapt knowledge to local context (rating feasibility of outcomes and integration into online platform), (3) assess barriers and facilitators (think-aloud interviews), and (4) select, tailor, implement (collaborative dissemination plan).

Results Step 1: Fifteen RE-AIM papers relevant to community programming were identified during the literature review. Two rounds of Delphi refined indicators for the toolkit related to reach, effectiveness, adoption, implementation, and maintenance. Step 2: At least one measure was linked with each indicator. Ten research and community partners participated in assessing the feasibility of measures, resulting in a total of 85 measures. Step 3: Interviews resulted in several recommendations for the online platform and toolkit. Step 4: Project partners developed a dissemination plan, including an information package, webinars, and publications.

Discussion This project demonstrates that community and university partners can collaborate to develop a useful, evidence-informed evaluation resource for both audiences. We identified several strategies for partnership when creating a toolkit, including using a set of expectations, engaging research users from the outset, using consensus methods, recruiting users through networks, and mentorship of trainees. The toolkit can be found at et.cdpp.ca. Next steps include disseminating (e.g., through webinars, conferences) and evaluating the toolkit to improve its use for diverse contexts (e.g., universal PA programming).

Keywords Implementation science, Knowledge translation, Delphi technique

*Correspondence:

Sarah V. C. Lawrason

svcl11@mail.ubc.ca

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



© Crown 2024. **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.

Plain English summary

Organizations that provide sport and exercise programming for people with disabilities need to evaluate their programs to understand what works, secure funding, and make improvements. However, these programs can be difficult to evaluate due to lack of evidence-informed tools, low capacity, and few resources (e.g., money, time). For this project, we aimed to close the evaluation gap by creating an online, evidence-informed toolkit that helps organizations evaluate physical activity programs for individuals with disabilities. The toolkit development process was guided by a community-university partnership and used a systematic four-step approach. Step one included reviewing the literature and building consensus among partners and potential users about indicators related to the success of community-based programs. Step two involved linking indicators with at least one measure for assessment. Step three involved interviews with partners who provided several recommendations for the online toolkit. Step four included the co-creation of a collaborative plan to distribute the toolkit for academic and non-academic audiences. Our comprehensive toolkit includes indicators for the reach, effectiveness, adoption, implementation, and maintenance of physical activity programs for individuals with disabilities. This paper provides a template for making toolkits in partnership with research users, offers strategies for community-university partnerships, and resulted in the co-creation of an evidence-informed evaluation resource to physical activity organizations. Users can find the toolkit at et.cdpp.ca.

Background

Disability and physical activity

The United Nations Convention on the Rights of Persons with a Disability protects the rights of people living with disabilities to access full and effective participation in all aspects of life, including sports and other recreational forms of physical activity (PA) such as exercise and active play. But because of countless environmental, attitudinal and policy barriers [1], children, youth and adults with disabilities are the most physically inactive segment of society [2, 3]. Physical inactivity increases the risk that people with disabilities will experience physical and mental health conditions, social isolation, and stigma [4]. Systematic reviews have evaluated the effects of participation in PA programs among children, youth, and adults with physical, intellectual, mental, or sensory disabilities. Many, but not all, of these reviews have reported significant improvements in physical health, mental health, and social inclusion [2]. One reason for the inconsistent outcomes is that the PA participation experiences of people with disabilities are not universally positive [5].

Qualitative and quantitative research shows that people with disabilities often report negative PA experiences; for instance, being marginalized, excluded, and receiving sub-standard equipment, access, instruction, and opportunities to fully participate in PA [6–8]. Research and theorizing on quality PA participation and disability indicate that these low-quality PA experiences deter ongoing participation and undermine the potential physical and psychosocial benefits of PA for children and adults [5, 9]. These findings attest to the need for evaluation of existing PA programs to identify what is working, and where improvements are needed to achieve optimal participation and impact.

Evaluating community-based programs

Persons with disabilities increasingly participate in disability sport to be physically active, and disability sport is often delivered by community organizations [2]. Like many community-based and non-profit organizations, organizations that provide PA programming for persons with disabilities (herein referred to as ‘this sector’) are often expected to conduct evaluations. These evaluations are done to secure and maintain external funding, demonstrate impact to board members and collaborators, and understand capacity for growth [10]. Even though program evaluations are often required, real-world programs are difficult to evaluate [11] and organizations often lack capacity and resources to conduct evaluations effectively [12]. Programs may be difficult to evaluate due to program complexity (e.g., setting, target population, intended outcomes; [11], and evaluation priorities (e.g., differing partner needs and resources; [13]. Organizations may lack capacity in understanding and using appropriate evaluation methods and tools [14], determining what counts as evidence and its application [15], and the roles of researchers and practitioners in supporting real-world program evaluations [16].

Evaluation frameworks can be used to facilitate a guided, systematic approach to evaluation. A framework involves an overview or structure with descriptive categories, meaning they focus on describing phenomena and how they fit into a set of categories rather than providing explanations of how something is working or not working [17]. One evaluation framework that is commonly applied in PA and disability settings is the RE-AIM framework [18]. RE-AIM is comprised of five evaluation dimensions or categories: (a) Reach: the number, proportion, and representativeness of

individuals who engage in a program, (b) Effectiveness: the positive and negative outcomes derived from a program, (c) Adoption, the number, proportion, and representativeness of possible settings and staff participating in the program, (d) Implementation: the cost and extent to which the program was intended to be delivered, and (e) Maintenance: the assessment beyond six months at the individual and organizational levels. The RE-AIM framework is appropriate in this sector because it aligns with organizations' need to understand factors that influence PA participation at both individual and organizational levels and for process (formative) and outcome (summative) evaluations [19–23]. Additionally, the RE-AIM framework has demonstrated feasibility to evaluate programs in this sector [19, 21, 22]. The RE-AIM framework was developed to address the failures and delays of getting scientific research evidence into practice and policy [18].

Gaps between evaluation research and practice

There has been a growing body of evidence to suggest that one of the most effective ways to bridge the gap between research and practice is through integrated knowledge translation (IKT; [24]). IKT means that the right research users are meaningfully engaged at the right time throughout the research process [25]. IKT involves a paradigmatic shift from recognizing researchers as 'experts' to valuing the expertise of individuals with lived experience, programmers, and policymakers through their inclusion in the development of the research questions, methods, execution, and dissemination to ensure that the research is relevant, useful, and usable [25]. A commitment to IKT aligns with the "nothing about us without us" philosophy of the disability rights movement [26] and is therefore ideal for a toolkit development process for this sector.

To address the gaps of lack of evidence-informed resources and reduced organizational capacity to conduct program evaluations [12], our community partners (leaders from seven Canadian organizations in this sector) identified that a toolkit is needed. An evaluation toolkit is a collection of tools that includes materials that may be used individually or collectively, such as educational material, timelines, and assessment tools, and the tools may often be customized based on context, thus helping to bridge the translation gap between evidence and practice [27]. Toolkit development can be a multi-step process including literature reviews, interviewing partners, and using a Delphi approach [27]. Previous research with community-based disability PA organizations suggests that digital platforms can be an efficient way for participants and staff to provide evaluation access to evaluation tools [19, 23]. Together, this research culminated in our decision to (1) use RE-AIM for the toolkit's

framework, meaning the toolkit was organized using the five evaluation dimensions, and (2) to deliver the toolkit through interactive technology. The purpose of this paper is to report on a systematic, IKT-focused process for the design, development, and formulation of implementation considerations for an online RE-AIM evaluation toolkit for organizations that provide PA programming for persons with disabilities.

Methods

Research approach

A community-university partnership was established between seven Canadian disability PA organizations and three universities. A technology partner guided the back-end development of the online toolkit. Using an IKT approach [25], community partners were engaged before the research grant was written and submitted to ensure that the project was meaningful and focused on the appropriate tasks and outcomes. To guide our partnership, we agreed to adopt the IKT guiding principles for SCI research [25] which aim to provide a foundation for meaningful engagement between partners. An example of a guiding principle is partners share in decision-making [25]. The principles were presented at each bi-monthly team meeting and participants had the opportunity to share concerns if certain principles were not upheld. Partners had regular opportunities for sharing in decision making, provided financial contributions to accelerate the project, and benefitted from developing the toolkit to tailor indicators and measures relevant for disability PA organizations. Two community partner leaders also provided mentorship to academic trainees on community engagement in research, employment in non-academia, and project management, emphasizing the multi-directional nature of the partnership. To see the entire IKT process, see Appendix A in the supplemental file.

To maximize the likelihood that our toolkit is used in practice, our development process was guided by the Knowledge-to-Action (KTA) framework (see Fig. 1; [28]). The KTA framework was developed to help researchers with knowledge translation by identifying the steps in moving knowledge into action [28]. The KTA framework has two components: (a) knowledge creation and (b) action cycle. Our toolkit development process followed the steps of the action cycle, whereby existing knowledge is synthesized, applied, and mobilized. The problem to be addressed is a need for a program evaluation toolkit. To solve the problem, as shown with the yellow boxes in Fig. 1, the steps for developing the RE-AIM evaluation toolkit included: (1) identify, review, and select knowledge; (2) adapt the knowledge to the local context and users; (3) assess the barriers and facilitators to knowledge use; and (4) select, tailor, and implement the toolkit.

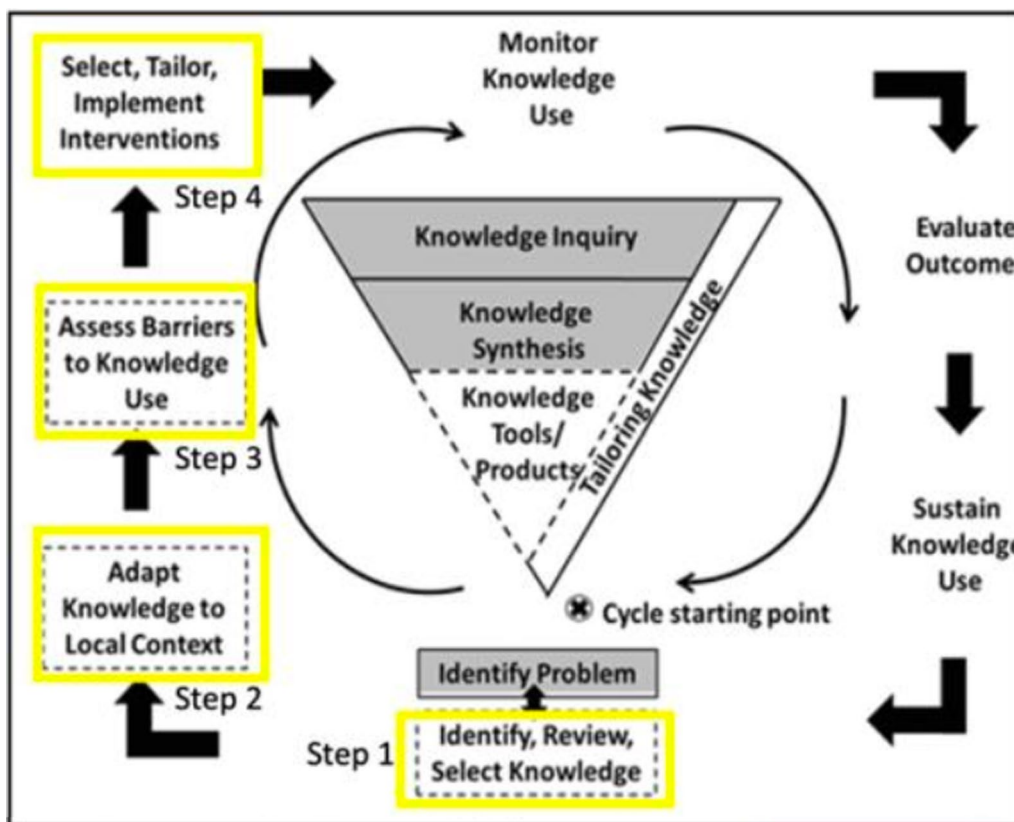


Fig. 1 Knowledge to action framework (adapted from [28])

To guide toolkit development, we ensured the methods aligned with recommendations from the COnsensus-based Standards/ Core Outcome Measures in Effectiveness Trials (COSMIN/COMET) groups for generating a set of core outcomes to be included in health intervention studies [29]. These guidelines state that developing a core outcome set requires finding existing outcome measurement instruments (see Step 1), quality assessment of instruments (see Step 2), and a consensus procedure to agree on the core outcome set (see Step 2) [29].

Step 1: Identify, review, and select knowledge

Literature review

The first step in identifying, reviewing, and selecting knowledge was to conduct a literature review. The literature review examined research using the RE-AIM framework to evaluate community-based and health-related programs. This was completed through a search of www.re-aim.org (which lists all RE-AIM evaluations) to identify indicators for each RE-AIM dimension within community-based and health-related contexts. Studies were included if they: used the RE-AIM framework to

evaluate a community-based health program or involved persons with disabilities, were published in English, and were peer reviewed. All study designs were included. The review also examined qualitative and quantitative studies of outcomes of community-based PA programs for people with disabilities (e.g., [9]) and outcomes our own partners have used in their own program evaluations. These papers and outcomes were discussed and chosen during early partnership meetings to initiate a list of indicators. Examples of community-based programs included peer support programs for individuals with spinal cord injuries in Quebec. Data extracted from papers included indicators (and their definitions) and associated measures used for evaluations.

Delphi process

The second part in identifying, reviewing, and selecting knowledge involves critically appraising the relevant literature identified, to determine its usefulness and validity for addressing the problem [28]. To determine usefulness and validity, a consensus-building outreach activity was used—an online Delphi method. Briefly, the Delphi method is used to arrive at a group decision by surveying a panel of experts [30, 31]. The experts each

respond to several rounds of surveys. Survey responses are synthesized and shared with the group after each round. The experts can adjust their responses in the next round based on their interpretations of the “group response.” The final response is considered a true consensus of the group’s opinion [30, 31]. Delphi was ideal for our partnership approach because it eliminates power dynamics from the consensus-building process and ensures every expert’s opinion is heard and equally valued. Previous research has demonstrated the utility of Delphi methods to generate consensus among disability organizations regarding the most important outcomes to measure in a peer-support program evaluation tool [32].

Delphi methodologies are considered a reliable means for achieving consensus when a minimum of six experts are included [33]. Therefore, we aimed to recruit a minimum of six participants from each target group (i.e., members of disability PA organizations and researchers). Partners were encouraged to invite members who may qualify and be interested in completing the Delphi process. Participants completed a two-round Delphi process and were asked to rate each RE-AIM indicator on a scale of 1 (not at all important) to 10 (one of the most important). An indicator was included if at least 70% of participants agreed it was “very important” (8 or above) [31]. Indicators that did not meet these criteria were removed from the list.

Retained indicators were then paired with at least one possible measure of that indicator (e.g., the ‘Positive Youth Development’ indicator was paired with the Out-of-School Time Observation instrument [34]). The partnership’s goal was to develop a toolkit comprised of valid and reliable measures. Therefore, the validity and reliability of each measure were critically appraised by the academic team-members using COSMIN/COMET criteria [29]. For some ‘Effectiveness’ indicators, published questionnaires were identified from the scientific literature. Measures were retained if they had high quality evidence of good content validity and internal consistency reliability [29] and were used in PA contexts and/or contexts involving participants with disabilities. The measures of all other indicators (where no published questionnaire measure was identified) were assessed by nine partners and modified to ensure that the measure was accurate and reliable for evaluation use in this sector.

Step 2: Adapt knowledge to local context

In the KTA framework, this phase involves groups making decisions about the value, usefulness, and appropriateness of knowledge for their settings and circumstances and customizing the knowledge to their particular situation [28]. Using Microsoft Excel, partners were sent a list of the selected indicators and measures in two phases (Phase

1: “RE” indicators and Phase 2: “AIM” indicators). Partners were asked to rate, on a scale of 0 to 2 the following categories for each measure: feasibility-time (not at all feasible to feasible), feasibility-complexity (not at all feasible to feasible), accuracy (not at all accurate to accurate), and unintended consequences (no, maybe, yes). They were also asked to provide additional feedback. This step only involved partners on the project with experience administering questionnaires (in research or evaluation settings) because the process required knowledge of how to administer measures to respondents. The median and mean of each category were calculated with community partner responses given double weighting/value relative to academic partner responses. Double weighting was given to community partner responses as the toolkit is anticipated to be used more frequently in community settings. The feedback was summarized. Results were presented to all partners during an online meeting, and team members discussed feedback to establish agreement on measures. The measures were sent out to partners again to provide any final feedback on included indicators and measures. The selected indicators and measures were compiled in an online program evaluation toolkit compliant with accessibility standards.

Step 3: Assess barriers and facilitators

In the KTA framework, this step involves identifying potential barriers that may limit knowledge uptake and supports or facilitators that can be leveraged to enhance uptake [28]. In Step 3, partners were invited to participate in an unstructured, think-aloud interview while they used the online program evaluation toolkit [35]. Interviews were conducted to collect detailed data about how users reacted to different parts of the toolkit content, format, and structure. Each interview was conducted over Zoom with one participant and two interviewers. The two-to-one interview format [36] supported the ability to take notes during the interview, ask questions from different perspectives, and reflect on common experiences to the two interviewers [36] with the website. Participants were also asked how the toolkit was used and any barriers to its use, and identified features of the toolkit that may need to be changed. In a separate group meeting, team members were asked for ideas on how to overcome potential barriers to using the toolkit and tips for its implementation. Data were analyzed using a content analysis approach [37] and recommendations were prioritized by the lead and senior authors using the MoSCoW method [38]. The MoSCoW method is a prioritization technique that has authors categorize recommendations using the following criteria: (a) “Must Have” (Mo), (b) “Should Have” (S), (c) “Could Have” (Co), and (d) “Won’t Have This Time” (W). These

recommendations were presented to all partners for further discussion. Based on the feedback, the toolkit content and technology were further iterated as needed. Information from this step was used to write brief user guides for toolkit users.

Step 4: Select, tailor, implement

In the KTA framework, this step involves planning and executing interventions to promote awareness and implementation of knowledge, and tailoring interventions to barriers and audiences [28]. In Step 4, during an online partnership meeting, a brainstorming activity was completed to discuss target audiences for the toolkit, barriers and facilitators to outreach, and dissemination ideas. Team members formulated a dissemination plan and identified promotional resources they need to tailor the dissemination of the toolkit to their sector networks.

Results

Step 1: Identify, review, and select knowledge

Literature Review

The initial searching process on the re-aim.org database identified 15 papers with relevant indicators for a RE-AIM toolkit. These papers and their citations are in Appendix B in the supplemental file. Additional resources identified by partners included: [2, 9, 39, 40], and partners’ previous experiences with evaluations to inform potential indicator choices. In total, 62 indicators were identified across all RE-AIM domains.

Delphi process

In round 1, 32 people participated in the exercise (two participants did not provide demographic information). In round 2, 28 people completed the questionnaire (four participants did not provide demographic information). Detailed participant demographics are presented in Table 1. The adaptation of indicators through the Delphi process can be found in Fig. 2. Given that nearly all indicators were deemed important from round 2, we agreed that a third round of the Delphi process was not needed. Based on the literature review, measures for each indicator were identified.

Step 2: Adapt knowledge to local context

Eight partners (*n*=3 academic, *n*=5 community) completed the rating process for the “RE” domains and 10 partners (*n*=3 academic, *n*=7 community) completed the rating process for the “AIM” domains (rating feasibility, complexity, accuracy, and unintended consequences; see Table 2). Respondent feedback was used to adapt and improve the measures to make them more feasible, less complex, and more accurate to reflect the indicators properly. Respondents also suggested that each

Table 1 Demographic details for Round 1 and Round 2 Delphi participants

Characteristics	Round 1 (n=32)	Round 2 (n=28)
<i>Role (n)</i>		
Researcher/professor	7	5
Coach, Coach Developer, or Athlete	4	1
Executive Director	4	3
Program Manager, Coordinator, or Evaluator	12	9
Physical Activity Consultant, Provider, or Therapist	3	6
<i>Primary activity of organization (n)</i>		
Post-secondary institution	7	5
Program/service delivery	20	15
Program/service planning or evaluation	3	4
Years of experience with organization, Mean ± SD	4.70 ± 5.88	3.04 ± 2.60
<i>Province (n)</i>		
Alberta	7	5
British Columbia	12	5
Maritimes and Quebec	2	2
Ontario	9	12
<i>Level of organization impact (%)*</i>		
Local	80	71
Provincial	60	54
Federal	13	21
<i>Types of disabilities served (%)*</i>		
Physical disabilities	100	100
Intellectual disabilities	67	63
Sensory disabilities	57	67
<i>Individual attributes (%)</i>		
Age (years), Mean ± SD	35.27 ± 9.43	34.54 ± 8.31
Women	83	88
Identify as person with disability	10	0
Have family member with disability	20	21
Caregiver for person with disability	17	21

*Participants can select more than one answer

measure should also include information boxes about the respondents, administrators, type of data collection, and time to complete data collection. The adaptation of indicators and measures from this process can be found in Fig. 2. The final list of indicators and measures can be found in Table 3.

Step 3: Assess barriers and facilitators

Six partners (community and academic partners) participated in unstructured think-aloud interviews, one of which was conducted jointly with two partners (*M_{time}*=43.37, *SD*±13.50 min). Across interviews, 45 unique recommendations were identified for improving

DEVELOPING A PROGRAM EVALUATION TOOLKIT

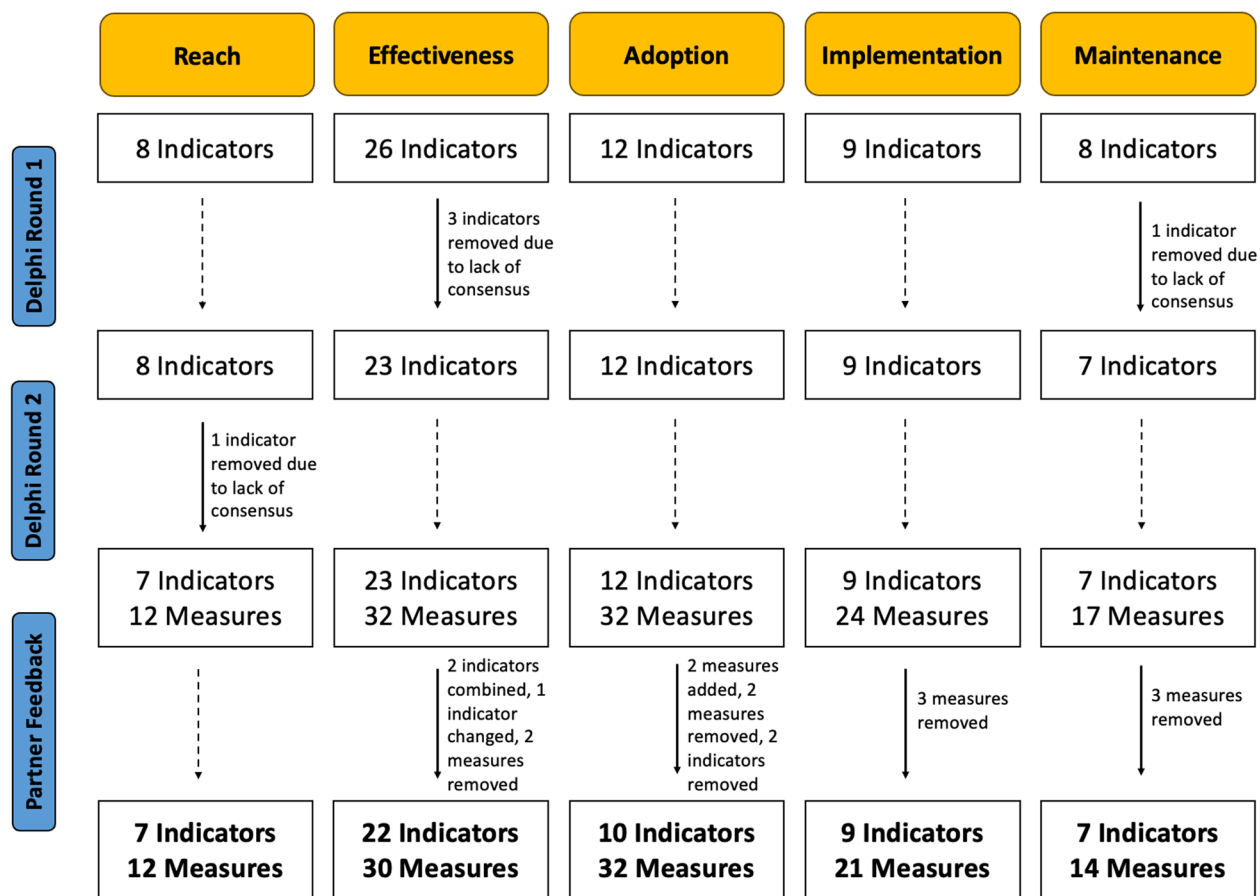


Fig. 2 Adaptation process for indicators and measures from the Delphi process and partner feedback during COSMIN/COMET rating

Table 2 Percentage of indicators with median ratings indicating feasibility (time, complexity), perceived accuracy, and no unintended consequences for RE-AIM measures

Domain	Number of measures assessed	Feasibility (time)	Feasibility (complexity)	Perceived accuracy	Unintended consequences
Reach	12	75	93	100	100
Effectiveness	32	91	94	100	100
Adoption	32	91	81	91	97
Implementation	23	96	100	100	100
Maintenance	17	71	88	76	100

the usability of the toolkit. These recommendations were sorted using the MoSCoW method, and prioritized based on budgetary constraints, team skillsets, and competing needs. Of the 45 recommendations, 30 were identified as ‘Must haves’, 6 as ‘Should haves’, 4 as ‘Could haves’, and 5 as ‘Won’t haves’ (see Appendix C in the supplemental file). All 30 ‘Must have’ recommendations were

implemented in collaboration with the technology partner, along with 2 ‘Should have’ recommendations.

Step 4: Select, tailor, implement

After all recommendations were executed by the technology partner, a final project meeting was held to discuss project updates, barriers and facilitators to outreach, and

Table 3 Indicators and their definitions and measures included in online toolkit (et.cdppp.ca)

Domain	Indicator	Definition	Measure(s)	Type of data collection
Reach	Target population	The individuals the program is trying to reach, or the individuals who are the target of the evaluation	Who is our target population? (e.g., athletes, coaches, volunteers)	Interviews/ Debriefs, Document Review
	Description of target population	Characteristics of the target population to describe the group	What are some specific characteristics of our target population? [Example response options: demographics (e.g., age, gender, ethnicity); disability-specific demographics (e.g., type, congenital vs. acquired, types of mobility aids); number of years in disability sport]	Document Review, Survey (Fixed)
Effectiveness	Community promotion	How the organization is promoting programs in the community	What kind of promotional material is [the organization] disseminating to reach the target audience? What types of events is [the organization] organizing to recruit or engage new participants?	Document Review, Interviews/ Debriefs
	Direct reach	How [the organization] is directly or physically reaching the target population	How many events has [the organization] hosted? How many individuals attended these events?	Document Review, Interviews/ Debriefs, Survey (Fixed)
	Awareness	Number of individuals aware of the organization or program. (Deliver this survey to the general population or target population)	Have you heard of [insert program or organization name]? If yes, how did you hear of [insert name]? [Provide response options: Email/Newsletter, social media, Word of mouth, etc.]	Survey (Fixed), Interviews/ Debriefs
	Willingness to participate	The number of individuals in target population who are willing to participate in the organization or program	According to Twitter, Instagram and Facebook, how many comments from users express interest in participating in the program or organization? Based on resources sent out to prospective participants, how many individuals expressed interest by responding and who were they?	Website Analytics, Document Review
	Population need	Understanding the documented gaps in the organization for reach	After reviewing previous documentation, internal reports, and above evaluated reach measures, who is [the organization] not reaching? What is missing in [the organization's] efforts?	Document Review
	Positive and negative outcomes for members	The positive and negative outcomes that members have experienced from participating in the program /organization	What has been your experience with [insert program]? What impact has [insert program] had on your life in general? What are some benefits that you have gained from participating in [insert program]? Have you had any negative outcomes or disadvantages from participating in [insert program]?	Interviews/ Debriefs
	Member belief in effectiveness	The percent of members who believe the program/organization has helped them improve in terms of program/ organization goals	Do you feel that [insert program] has improved [insert program goal]? Why or why not?	Survey (Fixed), Survey (Open)
	Service access	The number of members who access other programs or services because of their membership in your organization/ program. Understanding how members access these services, and how these supports change over time	Have you accessed other programs or services because of your membership in [insert program]? How have you accessed these services? Have these supports changed over time?	Survey (Fixed), Interviews/ Debriefs, Survey (Open)
	Member satisfaction	How satisfied members are with the program/ organization	Are you satisfied with [insert program]? Why or why not? Do you have any recommendations on how the program can improve?	Survey (Fixed), Survey (Open), Interviews/ Debriefs

Table 3 (continued)

Domain	Indicator	Definition	Measure(s)	Type of data collection
Psychological outcomes	Independence	One's ability to be self-sufficient	"I am more independent" (1-item)	Survey (Fixed)
	Life satisfaction	Feelings of satisfaction with one's own life, including self-satisfaction and satisfaction with different aspects of life	NIH Toolbox – General Life Satisfaction (5-items) [46]	Survey (Fixed)
	Meaning and purpose	Meaning in life refers to the feeling that one's life and experience make sense and matter. Purpose refers to the extent to which one experiences life as being directed, organized, and motivated by important goals	NIH Toolbox – Meaning and Purpose (7-items) [46]	Survey (Fixed)
	Confidence	Feeling self-assured about one's own qualities/capabilities	1-item measure: "I am _____ that I can accomplish most things I set out to do." On a scale from: Much less confident, to Much more confident [47]	Survey (Fixed)
	Self-efficacy	A person's feelings of control over their life, as well as their confidence in being able to manage their own functioning and deal effectively with situations and demands	<i>Domain or task specific self-efficacy</i> : 1-item measure: "How confident are you in your ability to..." followed by specific skill: 10-point Likert scale anchored by not at all confident and extremely confident. [48] <i>General Self-Efficacy</i> (4-items) [49]	Survey (Fixed)
	Resilience	The ability to bounce back and recover from difficult/ stressful situations	Brief Resilience Scale (6-items) [50]	Survey (Fixed)
	Quality of Life	An individual's perceptions of their position in life and in relation to their goals, expectations, standards, and concerns. This concept incorporates a person's physical health, psychological state, level of independence, social relationships, personal beliefs, and their relationships to salient features of the environment	WHO QoL BREF (26-items) [51]	Survey (Fixed)
Social outcomes	Social support	Perceived and received support that is intended to enhance the well-being and positive outcomes of the recipient	Social Support and Exercise Survey (13-items) [52]	Survey (Fixed)
	Sport and community inclusion	Making sure everybody has the same opportunities to participate in every aspect of sport and community to the best of their abilities and desires. Inclusion requires making sure that adequate policies and practices are in effect in a community or organization	Social Inclusion Measure (12-items) [53]	Survey (Fixed)
Physical health	Physical activity	Any bodily movement produced by skeletal muscles that requires energy expenditure. Physical activity refers to all movement including during leisure time, for transport to get to and from places, or as part of a person's work	International Physical Activity Survey: 7-items, estimates days/ week, hours/day, and minutes/day of intensity of physical activity and sitting time [54]	Survey (Fixed)
	Health	The overall evaluation of one's physical and mental health	Global Health Scale (7 to 10-items) [55, 56]	Survey (Fixed)
	Physical self-concept	A combination of individual self-perceptions related to physical appearance, athletic abilities, and physical capacities	Physical Self-Description Survey Short Form (47-items) [57]	Survey (Fixed)

Table 3 (continued)

Domain	Indicator	Definition	Measure(s)	Type of data collection
<i>Organizational outcomes</i>	Sport organization capacity	The overall capacity of an organization to produce the outputs and outcomes it desires. Partners will benefit from an oversight of the sport organization including: 1) The organization's capacity to deliver on various projects and to find the right fit, 2) It provides a measure of the sophistication of the organization	Organizational Capacity Variables (20-items) [58]	Survey (Fixed), Document Review
	Volunteer inspiration	Feeling of excitement and energy to take action to volunteer	Inspiration Scale (4-items) [59]	Survey (Fixed)
	Volunteer intention	The goal, reason, or purpose one has for volunteering	Intention to return to volunteering (5-items) [60]	Survey (Fixed)
<i>Program outcomes</i>	Quality participation	An athlete's broad subjective evaluation that their sport involvement is satisfying, enjoyable, and generates personally valued outcomes	Measure of Experiential Aspects of Participation (12-items) [61]	Survey (Fixed)
	Positive youth development	Extent to which youth participants enhance their skills, beliefs in the future, self-regulation, self-efficacy as well as social, emotional, cognitive, and behavioural competence	Out-of-School Time Observation Instrument (26-items) [34]	Observation (Fixed)
	Program quality	Refers to the structures and processes within a program that relate to participant outcomes. Program structures refer to an organization's capacity to deliver a program to youth (e.g., physical space, staffing, funding, community collaborations). Program processes refer to how the program is delivered (e.g., supportive relationships, opportunities for skill-building, autonomy)	Program Quality Assessment in Youth Sport (51-items) [62]	Observation (Fixed)
Adoption	Member-Type Adoption	The number of people from the target population who are members in the program/ organization, and how they are they involved (e.g., coach, volunteer, athlete)	How many members (according to member-type, e.g., athlete, staff) are in our program? How many people are estimated in your target population according to member type?	Document Review
	Engaged adoption	The total number of members that have become actively involved in a program/ initiative	How many members in your organization have adopted (or are actively participating in) a specific program/initiative?	Document Review
	Description of adopters	Demographics/description of members and staff who have adopted the program. This should be carefully considered about whether the information needs to be collected or not	10 Questions, Ex: What is your age? [Provide response options] What is your gender? [response options provided]	Survey (Fixed and Open)
	Accessibility and ease of adoption	Assess program /organization access to information and evaluated delivery of the program	What resources (delivery method e.g., pamphlets, mail, email) and information (content e.g., program information) is sent formally to new participants?	Document Review
	Dropouts	Dropouts from program or organization. Number of former program or organization members	How many people were registered in your program at the beginning of a season/program year? How many participants stayed in the program until the end of a season? How many participants registered for the next season?	Document Review

Table 3 (continued)

Domain	Indicator	Definition	Measure(s)	Type of data collection
	Attrition rate	The rate at which program members are no longer participating	Using your answers from "Dropouts", what percentage of individuals dropped out of the program? (Divide the number of dropouts by the number of people registered in the program?)	Document Review
	Usability	Staff satisfaction with organization or program	In an overall, general sense, how satisfied are you with the [insert program]? [Response options scale, 1–5] Why or why not? Do you have any recommendations on how the program can improve? [Provide response options]	Survey (Fixed and Open); Interviews/ Debriefs
	Professional development and contributions	The number and type of professional development opportunities available for staff	How many opportunities for professional development (e.g., workshops) are offered to staff and/or volunteers by the organization? How many people are eligible (i.e., qualified to participate) for each professional development opportunity?	Document Review, Survey (Fixed and Open), Interviews/ Debriefs
	Demand for Tools	The specific tools, resources, and/ or philosophies required for staff to aid their ability to promote program adoption	What kind of resources/tools have helped you in your ability in promoting program participation? What kind of philosophies and skills do you hold that aid in program participation?	Survey (Open); Interviews/ Debriefs
	Support and Commitment of Staff and Management	Staff perceptions about the organization's commitment and coordination by management on facilitating program adoption	What is the organization's commitment to facilitate program participation? Do you think it is enough? Why or why not? What is the organization's coordination to facilitate program participation? Do you think it is enough? Why or why not?	Survey (Open); Interviews/ Debriefs
Implementation	Adherence and commitment	The extent to which program or organizational goals are applied in programs. Measuring how well the organization adheres to implementing the program	1. How well do you think [insert goal] is applied in [insert program]? Why or why not? 2. How well does the organization adhere to program implementation?	Survey (Open); Interviews/ Debriefs
	Compatibility	Evaluate how well the actual implementation of the program compares to the intended implementation and organizational guidelines for the program	1. Can you discuss how the written goals of your program (mission statements, curriculums, etc.) are targeted regularly through the programming? 2. Does your program include the appropriate components according to the Program Quality Assessment in Youth Sport?	Document Review; Interviews/ Debriefs [Resource intensive], Connected to 'Program Quality Assessment in Youth Sport' measure
	Program Cost and funding	The cost of resources needed to run the program /organization and the funding received to subsidize cost	1. Using financial documents, what is the cost of resources (e.g., staff, equipment, total cost) required by all initiatives and programs? 2. What is the annual budget for your program? [Try to answer this question when budget information is available] 3. What is the total income received for program? (e.g., sponsorships, grants, yearly fees)	Document Review
	Delivery	The skills used by program implementers in order to deliver the program or attain organizational goals	1. What do you think makes the program implementation successful or not successful? Why? What do you think can be improved? 2. What are some of your skills that you use when delivering the program?	Survey (Open); Interviews/ Debriefs
	Contact	The type and amount of contact the program/ organization has with members. For instance, the number of direct contacts (e.g., phone calls, email, and personal communication) a program or organization has with members	1. How many phone contacts occur between the organization and members? 2. How many email contacts occur between the organization and members? (in a given time frame) 3. What other types of communication occur between the organization and members?	Document Review

Table 3 (continued)

Domain	Indicator	Definition	Measure(s)	Type of data collection
	Support from management	Whether management supports the implementation of goals in programs	<ol style="list-style-type: none"> 1. What support do you receive from management to implement the program? (e.g., time, resources, knowledge) 2. Do you think the support you receive is adequate? Why or why not? 3. How could management provide more support for program delivery? 	Survey (Open); Interviews/ Debriefs
	Staff skills	The training provided to staff, coaches, and volunteers to implement program goals	<ol style="list-style-type: none"> 1. What staff skills are fostered from staff engaging in training opportunities? Provide examples (leadership, professional coaching skills etc.) 	Survey (Open); Interviews/ Debriefs
	Time expenditure	The amount of time spent implementing programs	<ol style="list-style-type: none"> 1. What activities are involved in delivering [insert program] in [given time frame]? [Activities include lesson plans, implementing and planning program] 2. How much time do you spend on those activities? 	Document Review; Interviews/ Debriefs
	Cost benefit analysis	Whether various staff tasks have a benefit for the overall program or organization. The extent to which staff costs (e.g., staff training, salaries, number of staff) provide program benefits (e.g., recruitment, adoption) or participant benefits (e.g., improved performance or quality of life)	<ol style="list-style-type: none"> 1. What is the total cost of your program in [given time frame]? 2. Identify the most important positive outcomes for your program (adoption or see effectiveness measures, e.g., participation, quality of life). How many people experienced a positive change or received a certain score in that positive outcome? 3. What is the cost of each individual member who received the desired benefit listed in the previous question? (Divide total cost of program by number of members in relevant group, i.e., those who received benefit). Example: Total cost of program: \$5,000. Benefit = participation; number of members: 50. Benefit = gaining independence; number of members who improved independence = 30. Cost-benefit for participation: \$5,000/50 members = \$100 per member. Cost-benefit for gaining independence: \$5,000/30 members = \$166.67 	Document Review
Maintenance Individual	Continuity of members	Average number of years individuals have been members of program/ organization	<ol style="list-style-type: none"> 1. What is the average number of years an individual is a member? (Total membership years divided by total members) 	Document Review [May be difficult to collect this information]
	Outcome assessment	Evaluating how implementation of program/ organization goals changes over time. Comparing the initial program implementation with the current program implementation in terms of outcomes	<ol style="list-style-type: none"> How has your ability to implement program tools for program goals changed and compared over time? How have 'Effectiveness' outcomes changed over time (e.g., season per season, year per year)? 	Document Review; Interviews/ Debriefs
	Facilitators and Barriers	Members' experiences with the program and any factors that promote or inhibit their ability to participate	<ol style="list-style-type: none"> 1. What are your experiences with participation in [insert program]? 2. What has enabled your ability to participate in [insert program] over time? [Response options provided—family, transportation, fee support] 3. What has hindered your ability to participate in [insert program] over time? [Response options provided—family, transportation, fee support] 	Survey (Fixed and Open); Interviews/ Debriefs

Table 3 (continued)

Domain	Indicator	Definition	Measure(s)	Type of data collection
Organizational	Program Plan	Evaluate the program's/ organization's long-term plan for future success. For instance, understanding how the program will be maintained to achieve longevity in the community	1. Using documents, how have program plans been used to plan and achieve goals in the future? (e.g., 1–2 yrs, 2–5 yrs, 5–10 yrs) 2. How do you plan on maintaining the program within the community? What strategies will you use and why?	Interviews/ Debriefs, Document Review
	Cost	Evaluate the financial plan in place for maintaining program/ organizational costs. Understanding the financial sustainability of the program with respect to costs compared to income	1. Does the program receive ongoing funding? 2. Is the program based on a fee-for-service funding model to help offset the costs?	Document Review
	Capacity building	Evaluate staff training opportunities and the extent to which these opportunities improved staff confidence to implement programs, staff skills, or skills to work with members	1. How many staff training opportunities are there over time? 2. What staff skills arise from those training opportunities over time? 3. How many training opportunities on average achieve goal of changing staff skills? (Divide total number of times/sessions/years of training opportunity by number of times/sessions/years achieving staff skill goal)	Document Review
	Embeddedness in System	How the program /organization is incorporated with the rest of the sport system	1. Using documents to determine goals achieved, how do these programs achieve goals that are related to parent organizations? (e.g., PSOs, NSOs). 2. Using documents, how many partnerships does your program have with other organizations? 3. How well is your sport program embedded within an able-bodied sport/exercise/physical activity system? Why or why not?	Document Review [Can take time]; Survey (Fixed and Open)

ideas for dissemination. Barriers to outreach included lack of research or evaluation knowledge to use the toolkit, lack of funding to conduct evaluations, poor turnover from reaching users (i.e., users becoming aware of the toolkit) to receiving (i.e., users browse the toolkit website) to using the toolkit (i.e., users use the toolkit for an evaluation), and challenges connecting with hard-to-reach organizations. Facilitators to outreach included providing resources for evaluation support, connecting with trainees to support evaluations, having positive self-efficacy and attitude for conducting evaluation, building awareness on the benefits of the toolkit through a dissemination campaign, credibility in the toolkit development process, and reaching out to key funders for administration of toolkit as guidance.

The toolkit can be found at et.cdpp.ca and is intended to be used by community organizations and academic institutions that conduct program evaluations involving PA and disability (and inclusive integrated programming). This interactive toolkit allows users to customize to their program evaluation situation by selecting a) which RE-AIM dimensions they want to evaluate, and b) which indicators they want to measure within a particular RE-AIM dimension (e.g., self-efficacy and quality participation within the Effectiveness dimension). Based on users' selections, the toolkit program compiles the corresponding measures for each indicator into a customized, downloadable document that the user can then put in the format of their choosing (e.g., online survey, paper questionnaire) for their program evaluation. This design aligns with partner requests for a simple online interface that provides flexibility and tailoring to their program evaluation needs. The toolkit and user guides are made freely available (i.e., open access), to maximize accessibility to community organization and academic audiences.

A plan with dissemination and capacity building activities was created to ensure the supported uptake of the evaluation toolkit. Our priority was to create a knowledge translation and communications package (e.g., newsletter article, social media content) for community partner organizations to disseminate through their channels. This included disseminating information to other community organizations within their network and funding partners (e.g., Sport Canada, Canadian Tire Jumpstart, ParticipACTION, provincial ministries, and the Canadian Paralympic Committee). This package served as the official 'launch' of the evaluation toolkit on July 20, 2023. Through this package, other activities were listed as potential 'services' interested parties can use. These services include bookable time for 'office hours' whereby a one-on-one meeting on how to use the toolkit and conduct program evaluation can be arranged and a 1-h 'frequently asked questions' webinar/workshop. Other

activities included publishing an open-access manuscript, writing knowledge translation and media blogs about the manuscript, and delivering academic and community conference presentations.

Discussion

The purpose of this paper was to report on the process of developing an evaluation toolkit in partnership with organizations that provide PA programming for persons with disabilities. Informed by the RE-AIM framework [18] and the knowledge-to-action framework [28], the toolkit development process involved a literature review, Delphi process, and interviews to adapt indicators and measures. Recommendations from partners were implemented, and the final toolkit can be found at et.cdpp.ca. Partners collaborated to create a dissemination and capacity building plan to support the uptake of the toolkit across the target audience.

Community organizations struggle to conduct program evaluations and to use existing evaluation frameworks. A recent scoping review identified 71 frameworks used to evaluate PA and dietary change programs [41]. Despite access to many frameworks, Fynn et al. [41] found limited guidance and resources for using the frameworks. In response to these concerns, the toolkit acts as a resource for using the RE-AIM framework by facilitating the uptake of evidence-informed evaluation practices. The toolkit will help organizations overcome barriers to evaluation identified by previous research by increasing capacity to use appropriate methods and tools [14] and providing education on determining what counts as evidence and data [15]. This can facilitate better organizational direction, improved programming, and importantly, better quality PA experiences for individuals with disabilities. The toolkit also complied with accessibility standards, an important benchmark for our partnership and a necessary step when creating a product for organizations that serve persons with disabilities. Accessibility standards were relatively easy to achieve and should be customary in all IKT activities.

To the best of our ability, the toolkit was developed specifically for organizations that provide programming for people with disabilities by focussing the literature review, having program partners in the disability community participate in the Delphi process, and ensuring the validity and reliability of indicators in disability contexts. However, there is an enormous shortage of data related to PA and disability as most national health surveillance systems exclude or do not measure disability [2]. While this general limitation may affect the toolkit, it also means that the toolkit may be useful for universal PA organizations that are interested in evaluating programs with non-disabled

individuals. Additional research is needed to examine the effectiveness of the toolkit in diverse contexts.

This project provides a template for developing open-access, online evidence-informed toolkits using an IKT approach with community partners. There are few resources on how to develop toolkits for the health and well-being field informed by knowledge translation frameworks or that include perspectives of end-users (e.g., [42, 43]). The four-step mixed-methods approach was guided by the systematic use of frameworks to inform toolkit development. Our project utilized a rigorous, step-by-step process for creating toolkits and resources for this sector that centres the knowledge and expertise of research users. To centre the knowledge and expertise of research users, we employed several strategies identified by Hoekstra et al. [44] for building strong disability research partnerships. Important strategies for partnership when developing a toolkit include (1) using a set of norms, rules, and expectations, (2) engagement of research users in the planning of research, (3) using consensus methods (i.e., Delphi), and (4) recruiting research users via professional or community networks [44].

First, we used the IKT Guiding Principles [25] as the set of norms, rules, and expectations to guide our partnership. These principles were addressed throughout the partnership and provided criteria to understand the success of the partnership. Second, we engaged with community partners from the beginning of the research process. Working with community partners who were committed to developing a high-quality product was integral to the success of this project. Community partners were committed and highly engaged as the toolkit stemmed from a community-identified need, rather than solely a 'research gap'. Third, using consensus methods is an excellent strategy to avoid decision-making that is dominated by certain voices or interests in the partnership [45]. One way that our project allowed for multiple voices to be heard was through our anonymous Delphi processes, which encouraged partners to share their input in a non-confrontational and data-driven manner. Fourth, in our partnership, many individuals and organizations had longstanding working relationships and aligned priorities for the project. Building our partnership based on previous trusting, respectful relationships was essential and using the IKT guiding principles [25] ensured that we maintained similar values and priorities throughout the partnership.

We used an additional strategy that has not been previously mentioned in the IKT literature: mentorship of research trainees by community partners. Through

monthly meetings, two community partners provided mentorship sessions to three trainees. These sessions focused on how to close the research-to-practice gap and helped to facilitate strong relationships between researchers and research users. Mentorship was an important step for training the next generation of researchers to use IKT.

Limitations

This project has some limitations. First, an exhaustive systematic scoping review was not conducted to identify evaluation indicators. This may have limited the number of relevant evaluation indicators included in the Delphi surveys. However, given that only five indicators were removed, and none were added after two rounds of Delphi, we are confident that our search returned relevant indicators. In the future, it may be worthwhile to consider an in-person or video-conference-facilitated Delphi process to encourage discussion and differentiation of indicators. Second, we identified several barriers and facilitators for using the toolkit, but addressing these barriers meaningfully was beyond the scope of this paper. We are currently in the process of disseminating (e.g., social media campaigns, blogs, discussions with funders) and evaluating the toolkit (e.g., surveys, using data analytics). This data will be reported in a future paper. Third, the interviews revealed 45 unique recommendations for the website and toolkit, but only some of these recommendations could be implemented due to budgetary constraints (e.g., adding a search function and filtering indicators to the website could not be completed).

Conclusions

In summary, this paper reports on the development of an online, open-access program evaluation toolkit for the disability and PA sector. The toolkit is informed by the RE-AIM framework [18] and available at et.cdpp.ca. Our paper describes a four-step process guided by the KTA framework [28] and IKT principles [25] to work with community partners to ensure the toolkit is relevant, useful, and usable. The process included reviewing the literature, building consensus through two rounds of Delphi surveys, rating the feasibility and complexity of measures, assessing barriers and facilitators through think-aloud interviews, and crafting a dissemination and capacity-building plan. This paper provides a template for creating toolkits in partnership with research users, demonstrates strategies to enable successful community-university partnerships, and offers an evidence-informed evaluation resource to organizations that provide PA programming for persons with disabilities.

Abbreviations

IKT	Integrated knowledge translation
KTA	Knowledge-to-action
PA	Physical activity
RE-AIM	Reach, effectiveness, adoption, implementation, maintenance

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40900-024-00618-7>.

Additional file 1.

Additional file 2.

Acknowledgements

We would like to acknowledge Ava Neely and Kenedy Olsen for their contributions in assisting with this project. In memoriam of Jane Arkell who played an important role on this project and dedicated herself to improving the lives of individuals with disabilities.

Author contributions

KMG, PDS, EM, TF, JL, MG, JG, JA, JP, & NB made substantial contributions to the conception of the project. SVCL, PDS, EM, ALC, JRT, SS, TF, JL, and KMG designed the project. All authors were involved in acquiring the data through recruitment. SVCL, NA, and KMG analyzed the data. All authors were involved in interpreting the data. SVCL, NA, and KMG drafted the paper or substantively revised it. All authors have approved the submitted version of this paper. All authors have agreed both to be personally accountable for the author's own contributions and to ensure that questions related to the accuracy or integrity of any part of the work, even ones in which the author was not personally involved, are appropriately investigated, resolved, and the resolution documented in the literature.

Funding

This work was supported by a Social Sciences and Humanities Research Council Connection Grant.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations**Ethics approval and consent to participate**

Approval was waived as the project conducted under 'program evaluation' requirements for University of British Columbia.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹School of Health and Exercise Sciences, University of British Columbia, Kelowna, BC, Canada. ²International Collaboration on Repair Discoveries, Vancouver, BC, Canada. ³Abilities Centre, Whitby, ON, Canada. ⁴School of Kinesiology and Health Studies, Queen's University, Kingston, ON, Canada. ⁵Revved Up, Kingston, ON, Canada. ⁶Department of Kinesiology and Physical Education, McGill University, Montreal, QC, Canada. ⁷Center for Interdisciplinary Research in Rehabilitation of Greater Montreal (CRIR), Montreal, Canada. ⁸The Steadward Centre for Personal and Physical Achievement, University of Alberta, Edmonton, AB, Canada. ⁹Pickering Football Club, Pickering, ON, Canada. ¹⁰Rocky Mountain Adaptive, Canmore, AB, Canada. ¹¹Active Living Alliance, Ottawa, ON, Canada. ¹²BC Wheelchair Sports Association, Vancouver, BC, Canada. ¹³Department of Medicine, University of British Columbia, Vancouver, BC, Canada.

Received: 2 February 2024 Accepted: 23 July 2024

Published online: 02 September 2024

References

- Martin Ginis KA, Ma JK, Latimer-Cheung AE, Rimmer JH. A systematic review of review articles addressing factors related to physical activity participation among children and adults with physical disabilities. *Health Psychol Rev.* 2016;10(4):478–94. <https://doi.org/10.1080/17437199.2016.1198240>.
- Martin Ginis KA, van der Ploeg HP, Foster C, Lai B, McBride CB, Ng K, et al. Participation of people living with disabilities in physical activity: a global perspective. *Lancet.* 2021;398(10298):443–55.
- van den Berg-Emons RJ, Bussmann JB, Stam HJ. Accelerometry-based activity spectrum in persons with chronic physical conditions. *Arch Phys Med Rehabil.* 2010;91(12):1856–61.
- World Health Organization. *World report on disability.* Geneva; 2011.
- Evans MB, Shirazipour CH, Allan V, Zanhour M, Sweet SN, Martin Ginis KA, et al. Integrating insights from the parasport community to understand optimal experiences: the quality parasport participation framework. *Psychol Sport Exerc.* 2018;37:79–90.
- Martin Ginis KA, Gee CM, Sinden AR, Tomasone JR, Latimer-Cheung AE. Relationships between sport and exercise participation and subjective well-being among adults with physical disabilities: Is participation quality more important than participation quantity? *Psychol Sport Exerc.* 2024;70:102535.
- Allan V, Smith B, Côté J, Martin Ginis KA, Latimer-Cheung AE. Narratives of participation among individuals with physical disabilities: a life-course analysis of athletes' experiences and development in parasport. *Psychol Sport Exerc.* 2018;37:170–8.
- Orr K, Tamminen KA, Sweet SN, Tomasone JR, Arbour-Nicitopoulos KP. "I've had bad experiences with team sport": sport participation, peer need-thwarting, and need-supporting behaviors among youth identifying with physical disability. *Adapt Phys Activ Q.* 2018;35(1):36–56.
- Shirazipour CH, Latimer-Cheung AE. Understanding quality participation: exploring ideal physical activity outcomes for military veterans with a physical disability. *Qual Res Sport Exerc Health.* 2020;12(4):563–78. <https://doi.org/10.1080/2159676X.2019.1645037>.
- Patton M. *Qualitative research and evaluation methods.* London: Sage Publications; 2015.
- Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. *BMJ.* 2015 Mar 19 [cited 2022 Apr 18];350. Available from: <https://www.bmj.com/content/350/bmj.h1258>
- Lawrason S, Turnnidge J, Tomasone J, Allan V, Côté J, Dawson K, et al. Employing the RE-AIM framework to evaluate multisport service organization initiatives. *J Sport Psychol Action.* 2021;12(2):87–100.
- Habicht JP, Victora CG, Vaughan JP. Evaluation designs for adequacy, plausibility and probability of public health programme performance and impact. *Int J Epidemiol.* 1999;28(1):10–8. <https://doi.org/10.1093/ije/28.1.10>.
- Milstein B, Wetterhall S. A framework featuring steps and standards for program evaluation. *Health Promot Pract.* 2000;1(3):221–8. <https://doi.org/10.1177/15248399000100304>.
- Li V, Carter SM, Rychetnik L. Evidence valued and used by health promotion practitioners. *Health Educ Res.* 2015;30(2):193–205. <https://doi.org/10.1093/her/cyu071>.
- Lobo R, Petrich M, Burns SK. Supporting health promotion practitioners to undertake evaluation for program development. *BMC Public Health.* 2014;14(1):1315. <https://doi.org/10.1186/1471-2458-14-1315>.
- Nilsen P (2015) Making sense of implementation theories, models and frameworks. *Implementation Sci* 10:53. <https://doi.org/10.1186/s13012-015-0242-0>
- Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. *Am J Public Health.* 1999;89(9):1322–7.
- Shaw RB, Sweet SN, McBride CB, Adair WK, Martin Ginis KA. Operationalizing the reach, effectiveness, adoption, implementation, maintenance

- (RE-AIM) framework to evaluate the collective impact of autonomous community programs that promote health and well-being. *BMC Public Health*. 2019;19(1):803. <https://doi.org/10.1186/s12889-019-7131-4>.
20. Bean CN, Kendellen K, Halsall T, Forneris T. Putting program evaluation into practice: enhancing the Girls Just Wanna Have Fun program. *Eval Program Plann*. 2015;49:31–40.
 21. Gainforth HL, Latimer-Cheung AE, Athanasopoulos P, Martin Ginis KA. Examining the feasibility and effectiveness of a community-based organization implementing an event-based knowledge mobilization initiative to promote physical activity guidelines for people with spinal cord injury among support personnel. *Health Promot Pract*. 2014;16(1):55–62. <https://doi.org/10.1177/1524839914528210>.
 22. Sweet SN, Ginis KAM, Estabrooks PA, Latimer-Cheung AE. Operationalizing the RE-AIM framework to evaluate the impact of multi-sector partnerships. *Implement Sci*. 2014;9(1):74. <https://doi.org/10.1186/1748-5908-9-74>.
 23. Whitley MA, Forneris T, Barker B. The reality of evaluating community-based sport and physical activity programs to enhance the development of underserved youth: challenges and potential strategies. *Quest*. 2014;66(2):218–32. <https://doi.org/10.1080/00336297.2013.872043>.
 24. Graham ID, Kothari A, McCutcheon C, Angus D, Banner D, Bucknall T, et al. Moving knowledge into action for more effective practice, programmes and policy: protocol for a research programme on integrated knowledge translation. *Implement Sci*. 2018;13(1):22. <https://doi.org/10.1186/s13012-017-0700-y>.
 25. Gainforth HL, Hoekstra F, McKay R, McBride CB, Sweet SN, Martin Ginis KA, et al. Integrated knowledge translation guiding principles for conducting and disseminating spinal cord injury research in partnership. *Arch Phys Med Rehabil*. 2021;102(4):656–63.
 26. Charlton JI. *Nothing about us without us: disability oppression and empowerment*. Berkeley: University of California Press; 1998.
 27. Thoele K, Ferren M, Moffat L, Keen A, Newhouse R. Development and use of a toolkit to facilitate implementation of an evidence-based intervention: a descriptive case study. *Implement Sci Commun*. 2020;1(1):86. <https://doi.org/10.1186/s43058-020-00081-x>.
 28. Graham ID, Logan J, Harrison MB, Straus SE, Tetroe J, Caswell W, et al. Lost in knowledge translation: time for a map? *J Contin Educ Heal Prof*. 2006;26(1):13–24. <https://doi.org/10.1002/chp.47>.
 29. Prinsenc CAC, Vohra S, Rose MR, Boers M, Tugwell P, Clarke M, et al. How to select outcome measurement instruments for outcomes included in a “Core Outcome Set”—a practical guideline. *Trials*. 2016;17(1):449. <https://doi.org/10.1186/s13063-016-1555-2>.
 30. Hsu CC, Sandford BA. Minimizing non-response in the Delphi Process: how to respond to non-response. *Practical Assessment, Research and Evaluation*. 2007;12(17).
 31. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. *J Adv Nurs*. 2000;32(4):1008–15.
 32. Shi Z, Michalovic E, McKay R, Gainforth HL, McBride CB, Clarke T, et al. Outcomes of spinal cord injury peer mentorship: a community-based Delphi consensus approach. *Ann Phys Rehabil Med*. 2023;66(1):101678.
 33. Black N, Murphy M, Lamping D, McKee M, Sanderson C, Askham J, et al. Consensus development methods: a review of best practice in creating clinical guidelines. *J Health Serv Res Policy*. 1999;4(4):236–48. <https://doi.org/10.1177/135581969900400410>.
 34. Birmingham J, Pechman EM, Russell CA, Mielke M. Shared features of high-performing after-school programs: a follow-up to the TASC evaluation. Washington, DC; 2005.
 35. van den Haak MJ, de Jong MDT, Schellens PJ. Evaluation of an informational Web site: three variants of the think-aloud method compared. *Tech Commun*. 2007;54(1):58–71.
 36. Monforte J, Úbeda-Colomer J. Tinkering with the two-to-one interview: reflections on the use of two interviewers in qualitative constructionist inquiry. *Methods Psychol*. 2021;5:100082.
 37. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. *Qual Health Res*. 2005;15(9):1277–88. <https://doi.org/10.1177/1049732305276687>.
 38. Hatton S. Early Prioritisation of Goals. In: Hainaut JL, Rundensteiner EA, Kirchberg M, Bertolotto M, Brochhausen M, Chen YPP, et al., editors. *Advances in conceptual modeling—foundations and applications*. Berlin, Heidelberg: Springer; 2007. p. 235–44.
 39. Williams TL, Ma JK, Martin Ginis KA. Participant experiences and perceptions of physical activity-enhancing interventions for people with physical impairments and mobility limitations: a meta-synthesis of qualitative research evidence. *Health Psychol Rev*. 2017;11(2):179–96. <https://doi.org/10.1080/17437199.2017.1299027>.
 40. Shirazipour CH, Tennant EM, Aiken AB, Latimer-Cheung AE. Psychosocial aspects of physical activity participation for military personnel with illness and injury: a scoping review. *Mil Behav Health*. 2019;7(4):459–76. <https://doi.org/10.1080/21635781.2019.1611508>.
 41. Fynn JF, Hardeman W, Milton K, Jones AP. A scoping review of evaluation frameworks and their applicability to real-world physical activity and dietary change programme evaluation. *BMC Public Health*. 2020;20(1):1000. <https://doi.org/10.1186/s12889-020-09062-0>.
 42. Hildebrand J, Lobo R, Hallett J, Brown G, Maycock B. My-peer toolkit [1.0]: developing an online resource for planning and evaluating peer-based youth programs. *Youth Studies Australia*. 2012;31(2):53–61.
 43. Buckingham S, Anil K, Demain S, Gunn H, Jones RB, Kent B, et al. Telerehabilitation for people with physical disabilities and movement impairment: development and evaluation of an online toolkit for practitioners and patients. *Disabil Rehabil*. 2023;45(11):1885–92. <https://doi.org/10.1080/09638288.2022.2074549>.
 44. Hoekstra F, Trigo F, Sibley KM, Graham ID, Kennefick M, Mrklas KJ, et al. Systematic overviews of partnership principles and strategies identified from health research about spinal cord injury and related health conditions: a scoping review. *J Spinal Cord Med*. 2023;46(4):614–31. <https://doi.org/10.1080/10790268.2022.2033578>.
 45. Oliver K, Kothari A, Mays N. The dark side of coproduction: do the costs outweigh the benefits for health research? *Health Res Policy Syst*. 2019;17(1):33. <https://doi.org/10.1186/s12961-019-0432-3>.
 46. Salsman JM, Lai JS, Hendrie HC, Butt Z, Zill N, Pilkonis PA, et al. Assessing psychological well-being: self-report instruments for the NIH Toolbox. *Qual Life Res*. 2014;23(1):205–15. <https://doi.org/10.1007/s11136-013-0452-3>.
 47. Bandura A. *Self-Efficacy Beliefs of Adolescents*. Pajares F, Urdan T, editors. Greenwich: information age publishing; 2006.
 48. Bandura A. *Social foundations of thought and action: a social cognitive theory*. Englewood Cliffs: Prentice Hall; 1986.
 49. Salsman JM, Schalet BD, Merluzzi TV, Park CL, Hahn EA, Snyder MA, et al. Calibration and initial validation of a general self-efficacy item bank and short form for the NIH PROMIS®. *Qual Life Res*. 2019;28(9):2513–23. <https://doi.org/10.1007/s11136-019-02198-6>.
 50. Smith BW, Dalen J, Wiggins K, Tooley E, Christopher P, Bernard J. The brief resilience scale: assessing the ability to bounce back. *Int J Behav Med*. 2008;15(3):194–200. <https://doi.org/10.1080/1070550080222972>.
 51. The WHO QoL Group. Development of the World Health Organization WHOQOL-BREF quality of life assessment. *Psychol Med*. 1998;05/01. 1998;28(3):551–8.
 52. Sallis JF, Grossman RM, Pinski RB, Patterson TL, Nader PR. The development of scales to measure social support for diet and exercise behaviors. *Prev Med (Baltimore)*. 1987;16(6):825–36.
 53. Asunta P, Rintala P, Pochstein F, Lyyra N, McConkey R. The development and initial validation of a short, self-report measure on social inclusion for people with intellectual disabilities: a transnational study. *Int J Environ Res Public Health*. 2021;18(5):2540.
 54. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc*. 2003;35(8):1381–95.
 55. Forrest CB, Bevans KB, Pratiwadi R, Moon J, Teneralli RE, Minton JM, et al. Development of the PROMIS® pediatric global health (PGH-7) measure. *Qual Life Res*. 2014;23(4):1221–31. <https://doi.org/10.1007/s11136-013-0581-8>.
 56. Hays RD, Bjorner JB, Revicki DA, Spritzer KL, Cella D. Development of physical and mental health summary scores from the patient-reported outcomes measurement information system (PROMIS) global items. *Qual Life Res*. 2009;18(7):873–80. <https://doi.org/10.1007/s11136-009-9496-9>.
 57. Marsh HW, Martin AJ, Jackson S. Introducing a short version of the physical self description questionnaire: new strategies, short-form evaluative criteria, and applications of factor analyses. *J Sport Exerc Psychol*. 2010;32(4):438–82.

58. Wicker P, Breuer C. Exploring the organizational capacity and organizational problems of disability sport clubs in Germany using matched pairs analysis. *Sport Manag Rev*. 2014;17(1):23–34.
59. Thrash TM, Elliot AJ. Inspiration: core characteristics, component processes, antecedents, and function. *J Pers Soc Psychol*. 2004;87(6):957–73.
60. Bakhsh JT, Lachance EL, Thompson A, Parent MM. Outcomes of the sport event volunteer experience: examining demonstration effects on first-time and returning volunteers. *Int J Event Festiv Manag*. 2021;12(2):168–83. <https://doi.org/10.1108/IJEFM-09-2020-0057>.
61. Caron JG, Martin Ginis KA, Rocchi M, Sweet SN. Development of the measure of experiential aspects of participation for people with physical disabilities. *Arch Phys Med Rehabil*. 2019;100(1):67–77.
62. Bean C, Kramers S, Camiré M, Fraser-Thomas J, Forneris T. Development of an observational measure assessing program quality processes in youth sport. *Cogent Soc Sci*. 2018;4(1):1467304. <https://doi.org/10.1080/23311886.2018.1467304>.

Publisher's Note

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