


RESEARCH ARTICLE

Open Access



A pediatric virtual care evaluation framework and its evolution using consensus methods

Catherine Dulude^{1,2}, Stephanie Sutherland², Shelley Vanderhout^{2,3}, W. James King^{1,2,4}, Caroline Zuijdwijk^{1,2,4}, Nathalie Major^{1,2,4}, Tobey Audcent^{1,2,4}, Heather Howley², Paula Cloutier^{1,2}, Melanie Buba^{1,2,4,5}, Raagini Jain^{1,4}, Justyna Litwinska¹, Leanne Findlay⁵, Claudia Malic^{1,2,4}, Karen Macaulay¹, Kasey Parker¹, Christine Kouri¹ and Ellen B. Goldbloom^{1,2,4*} 

Abstract

Background The use of virtual care has increased dramatically in response to the COVID-19 pandemic, yet evidence is lacking regarding the impact of virtual care on patient outcomes, particularly in pediatrics. A standardized evaluation approach is required to support the integration of virtual care into pediatric health care delivery programs. The objective of this work was to develop a comprehensive and structured framework for pediatric virtual care evaluation. This framework is intended to engage and guide care providers, health centres, and stakeholders towards the development of a standardized approach to the evaluation of pediatric virtual care.

Methods We brought together a diverse multidisciplinary team, including pediatric clinicians, researchers, digital health leads and analysts, program leaders, a human factors engineer, a family advisor and our manager of health equity and diversity. The team reviewed the literature, including published evaluation frameworks, and used a consensus-based method to develop a virtual care evaluation framework applicable to a broad spectrum of pediatric virtual care programs. We used an iterative process to develop framework components, including domains and sub-domains, examples of evaluation questions, measures, and data sources. Team members met repeatedly over seven months to generate and provide feedback on all components of the framework, making revision as needed until consensus was reached. The framework was then applied to an existing virtual care program.

Results The resulting framework includes four domains (health outcomes, health delivery, individual experience, and program implementation) and 19 sub-domains designed to support the development and evaluation of pediatric virtual care programs. We also developed guidance on how to use the framework and illustrate its utility by applying it to an existing pediatric virtual care program.

Conclusions This virtual care evaluation framework expands on previously developed frameworks by providing additional detail and a structure that supports practical application. It can be used to evaluate a wide range of pediatric virtual care programs in a standardized manner. Use of this comprehensive yet easy to use evaluation framework will inform appropriate implementation and integration of virtual care into routine practice and support its sustainability and continuous improvement.

Keywords Virtual care, Evaluation, Pediatrics, Framework

*Correspondence:

Ellen B. Goldbloom
egoldbloom@cheo.on.ca

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



© The Author(s) 2023. **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.

Contributions to literature

- The proposed methodology structured within this framework supports and expands on previously published virtual care evaluation frameworks by providing additional detail including evaluation sub-domains and a structure that supports practical application.
- The framework presented herein is widely applicable and can be used to evaluate a range of pediatric virtual care programs in a standardized manner.
- Application of this comprehensive evaluation framework will inform appropriate implementation and integration of virtual care into routine practice and support its sustainability and continuous improvement.

Background

The global SARS CoV-2 (COVID-19) pandemic accelerated the growth and “normalization” of virtual health care, [1–4] which encompasses any interaction between patients and/or members of their circle of care, using a range of technologies and applications, that support synchronous and asynchronous care delivered at a distance. Examples of virtual care include telehealth, telemedicine, eMental Health and eHealth [4]. Virtual care has been shown to offer benefits to patients, families, and healthcare providers, such as convenience, improved care accessibility, and higher patient and family satisfaction [5]. However, there is limited evidence, particularly for pediatrics, of the impact of virtual care on patient health outcomes and care experiences.

Patients, families and healthcare providers have expressed enthusiasm for virtual care and expect it to continue beyond the COVID-19 pandemic [6]. To ensure appropriate integration into safe, effective, equitable, and cost-effective care, there is a need for evaluation, performance measurement and monitoring of virtual care delivery models [7]. Virtual care is not interchangeable with in-person care. In some cases, the need for in-person care is indisputable; for example, a new patient requiring a physical exam or procedure. In other cases, it is clear that virtual or hybrid in-person/virtual care has benefits; for example, when sub-specialists provide care virtually to remote communities to support assessments or treatments that are not locally available. Yet in many cases, competing factors such as patient or provider preference, access to appropriate technology, and prohibitive travel costs, make direct comparisons between virtual and in-person care difficult. In these cases, establishing whether virtual care is

more, or less, appropriate requires evaluation of many factors. A lack of standardized data definitions and evaluation approaches makes evidence synthesis challenging [7, 8]. In the absence of data-driven recommendations and protocols, expert groups have convened to provide guidance on how to safely incorporate virtual care into their healthcare delivery models [9–13].

Evaluation frameworks can be used to guide the evaluation process [14–16] and consist of the following elements: evaluation questions and associated measures or indicators of success, data sources, data collection strategies, and bases of comparison (internal and/or external, standards, etc.); see Additional file 1. A structured framework can be used to support ongoing assessment of health care delivery, determine the value of virtual care for key stakeholders, and facilitate appropriate integration of virtual care into service delivery [10, 17]. While several different evaluation frameworks [10, 18, 19] and measurement standards have been introduced to guide assessment of virtual care, a standardized approach is needed to yield meaningful evidence to guide implementation and integration of virtual care in pediatric settings. The American Academy of Pediatrics (AAP) Section on Telehealth Care’s SPROUT (Supporting Pediatric Research on Outcomes and Utilization of Telehealth) recognized this need and published the SPROUT Telehealth Evaluation and Measurement (STEM) framework [10] with an intention to welcome collaboration towards its continued evolution [20]. The STEM framework synthesizes quality-focused evaluation frameworks and guidelines developed by the National Quality Forum (NQF), [21]. World Health Organization (WHO) [19] and Agency for Healthcare Research and Quality (AHRQ) [22] into a single outcomes-focused framework organized into four measurement domains: health outcomes, health delivery (quality and cost), experience, and program implementation and key performance indicators (KPIs). It is a helpful guideline and clearly aligned with the quadruple aim (improving patient experience, health of populations and provider satisfaction while reducing costs) [23, 24].

We saw an opportunity to build upon the STEM framework and provide additional detail that would support clinical or program leaders in planning and conducting their own evaluations of virtual care. Using the STEM framework [10] as a guide, our aim was to develop a comprehensive framework to support a standardized approach to virtual care evaluation across a variety of programs. We further aimed to create a virtual care evaluation framework that was flexible for use, in both research or quality improvement contexts, at various stages of program development, from

planning through implementation, monitoring and improvement.

Methods

Setting and team

At the onset of the COVID-19 pandemic, our pediatric academic centre quickly adapted to virtual care delivery in almost all clinical areas with a focus on safety, access, and sustainability. For example, we now deliver approximately 40% of annual outpatient visits virtually, with program-level virtual care evaluation projects, and separate centre-wide initiatives to collect feedback from patients, families, and healthcare providers [25]. However, a comprehensive evaluation approach has not been standardized across the organization.

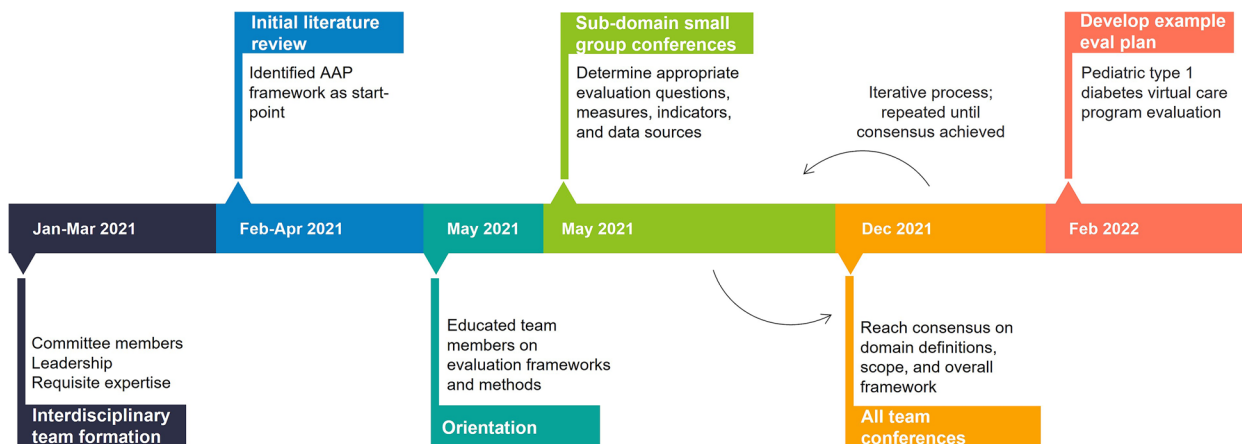
Thus, we formed a multidisciplinary team of virtual care leaders and stakeholders to develop a unified approach to virtual care evaluation. This initiative was sponsored by the centre’s executive leadership and research institute. The team consisted of 17 clinicians, researchers, operational leaders and stakeholders from diverse clinical and operational specialties: pediatric mental health (PC); pediatric complex care (NM); pediatric surgery (CM); pediatric emergency medicine (RJ); pediatric medicine (MB, WJK, TA); pediatric endocrinology (EG, CZ); autism/behaviour services (JL); quality improvement (CM, MB); epidemiology (WJK); information technology & informatics (KM, WJK, EG); human factors (CD); business intelligence and reporting (KP); program evaluation and qualitative methodology (SS); research development (HH); patient engagement (EG, MB, LF); family voice (LF); equity, diversity inclusion & Indigeneity (CK); social

determinants of health (CZ, TA) and medical education (CZ, TA, RJ).

Approach

We used consensus methodology, including consensus development panels, to produce a pediatric virtual care evaluation framework. The framework development timeline is presented in Fig. 1. A consensus development panel is one of the three approaches (nominal group process, consensus development panels, Delphi technique) to conducting consensus methodology research. We selected consensus development panels as it was most amenable to our objectives by allowing the flexibility to have as many consensus building rounds as necessary. Consensus development panels (also known as consensus development conferences) are organized meetings of between eight and twelve experts from a given field, or combination of fields [26]. However, panel composition can be modified to accommodate larger or smaller groups of experts. In this case teams were comprised of multidisciplinary experts. Consensus development panels are useful for achieving consensus in health care because it is an evidence-based and multidisciplinary approach for problem solving and policy development [26–29].

The STEM framework [10] was critically appraised by all team members and approved as a starting point to populate the domains and sub-domains of our framework. Strengths and weaknesses of the STEM framework for application and use within our context were discussed. To achieve a shared vision of a more extensive, comprehensive, applicable and usable framework, members also used other healthcare evaluation frameworks, including: Ontario Health Quality, [18] COMET



Note: Kickoff meeting March 18, 2021

Fig. 1 Framework development timeline

Outcomes Classification, [30] Canada Health Infoway Benefits Evaluation Framework (CHIBEF), [31] and other work on the evaluation of telehealth usability [18, 19, 32].

Consensus on evaluation framework domains and sub-domains

Starting with the four AAP framework domains (health outcomes, health delivery, experience, program implementation and KPIs), multi-disciplinary sub-groups were convened to generate evaluation questions, measures, and data sources for each. Consensus development panels were composed based on interest and expertise, while ensuring representation from diverse clinical areas (e.g., the health outcomes group was composed of a clinical researcher in mental health (PC), research development manager (HH), complex care physician and clinical lead (NM), and a surgeon and quality improvement lead (CM)). Consensus development panels met regularly over a period of eight months. An iterative process, alternating between meeting in smaller panels to develop framework details and as a larger group to discuss and develop consensus on domain definitions and scope, was guided by an expert in evaluation methodology (SS) and a human factors specialist with experience in virtual care development and evaluation (CD).

Each panel met repeatedly, typically three to five one-hour meetings, until they reached consensus on the scope and examples within their domain. Then, the entire evaluation framework was reviewed and revised by all team members until consensus was reached on final content.

Target audience and application

The team discussed and reached consensus on appropriate settings and contexts for which the framework could be used and clarified the intended audience. As a final step, we applied the framework to an existing virtual care program (pediatric type 1 diabetes) to assess real-world applicability, consistency, and illustrate practical use. Two pediatric endocrinologists used the framework to specify and contextualize the questions, measures/indicators, and aligned data sources and comparators within each domain and sub-domain to create an evaluation plan relevant for the patient population. A formal evaluation of the implementation of the framework is not presented in this manuscript.

Results

We developed a pediatric virtual care evaluation framework (Fig. 2) based on the four domains proposed by STEM [10] and other established evaluation frameworks and tools [17, 29, 31]. Team members agreed the framework should support comprehensive evaluation of virtual

care programs and appropriate integration of virtual care based on patient factors (e.g., age, diagnosis, socioeconomic status), timing in the care pathway, and mode of delivery. We also agreed to use the term "virtual care" [3] rather than telehealth or telemedicine to highlight the applicability of this framework to all types of virtual care regardless of technology or synchronicity. We believe this framework can be applied in a variety of contexts ranging from community-based practice to multi-site health centres, and oversight or funding agencies (e.g., government health ministry).

Framework structure

Our framework has four measurement domains that parallel STEM: health outcomes, health delivery, individual experience, and program implementation. Under these four domains, we have specified 19 sub-domains. In addition to the sub-domains referenced in the STEM domain descriptions and examples, we added the following: Behavioural to the Mental sub-domain under Health Outcomes, Privacy under Health Delivery, Usefulness under Individual Experience, and Leadership under Program Implementation. Under Program Implementation, we also sub-divided System Changes into Resources, Training and Support, and Infrastructure and Technology. With these modifications, we provide a structure that can readily support development of organization-wide or program-specific evaluation questions to illustrate scope, and associated measures, data sources, and bases of comparison for each (see Tables 1, 2, 3 and 4 for examples and Additional file 2 for the full framework). We therefore removed Key Performance Indicators (KPIs) from the Program Implementation domain title used by STEM, as the development of indicators is embedded in the framework structure across all domains. Using this tool, other teams can apply the framework to a variety of patient care settings; articulate their own evaluation questions under each subdomain, and identify measures, data sources and bases of comparison that can be used to answer each question.

Framework domains and sub-domains

The first domain, Health Outcomes, includes clinical measures of individual and population health within three sub-domains: physical, mental and behavioural, and quality of life (QoL). This domain aligns with the ultimate goal of helping children and youth achieve their best life by going beyond direct measures of health to include the impact of health status on the quality of life of patients and caregivers. We added "behavioural" to the mental health sub-domain to ensure that our framework is relevant for programs delivering developmental and behavioural health services (e.g., applied behaviour

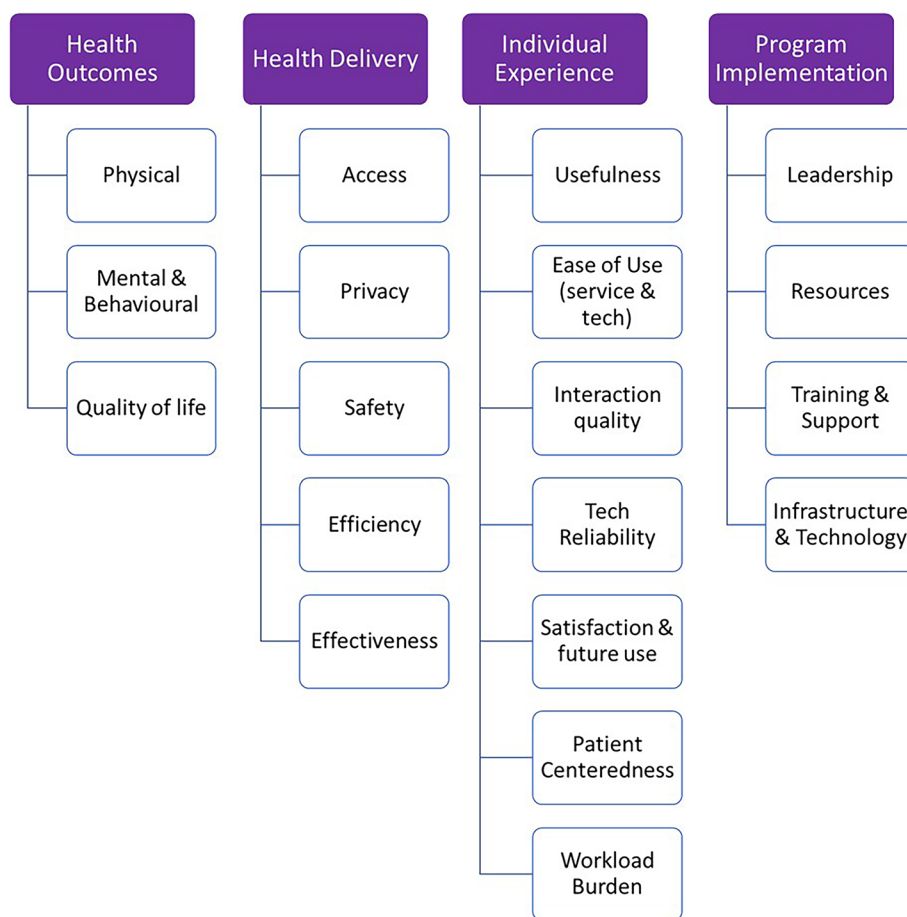


Fig. 2 Pediatric Virtual Care Evaluation Framework

analysis, complex care) and provides a holistic view of the impact of virtual care on pediatric health outcomes. Data sources are primarily patient charts, assuming patient and caregiver self-reports and other measures of QoL are included therein. Comparators include internal or external databases or clinical practice guidelines (depending on the evaluation questions). Table 1 provides example evaluation questions for the physical health outcomes sub-domain along with associated measures, data sources and bases of comparison.

The second domain, Health Delivery, includes five sub-domains: access to care, privacy, safety, efficiency, and effectiveness of care delivery. We have defined the scope of these sub-domains as follows. Access to care considers both timeliness (e.g., time to first visit, wait times) and equity (e.g., to what extent the program provides equitable access to care or addresses access or equity gaps present in other programs). Privacy examines the patient environment and the extent to which the tools and processes used to deliver virtual care adhere to privacy and confidentiality standards. Safety considers the potential

for virtual care to introduce or address safety risks (e.g., adverse event reporting or risk assessments). Efficiency considers both time and cost (e.g., the financial and time–cost-difference for patients to travel to the hospital compared to attending a virtual visit). Finally, effectiveness considers how well the virtual care program delivers care as intended in terms of quality and quantity, and can be linked back to program objectives and compared to in-person care as appropriate. While organized slightly differently, four of these five sub-domains are consistent across AAP STEM, [10], Ontario Health Quality [18] and CHIBE [31] frameworks. We added privacy because it is a theme that appears regularly in telehealth experience surveys, especially when evaluating eMental Health programs [33–35] and can be impacted by virtual models of care delivery. Examples of evaluation questions, measures, and data sources for the privacy sub-domain are presented in Table 2.

The third domain, Individual Experience, includes seven sub-domains: usefulness, ease of use, interaction quality, technology reliability, satisfaction and future use,

Table 1 Example for the Health Outcomes domain, Physical health outcomes sub-domain: questions, measures, and data specifics

Health Outcomes				
Example evaluation questions	Potential measures	Data sources	Collection strategy	Basis of comparison
Physical health outcomes (i.e., measures of physiological function, signs and symptoms, laboratory (and other scientific) measures relating to the function of major organ systems, the extent of disease or disability and the provision and/or response to therapeutic interventions)				
What is the impact of virtual care on individual physiological/clinical measures of health?	Diabetes patient’s Hemoglobin A1C (e.g., value, % within target), BMI (e.g., within normal range), Mortality, SSI (Surgical Site Infection) (e.g., SSI infection occurrence, frequency, rate) Complex care pain scores (child or parent proxy, e.g., Linear VAS), feeding/swallowing performance	Patient charts	Quantitative, surveys	Data registries, literature
What is the impact of virtual care on population level physiological/clinical measures of health?	Children/Youth with diabetes: Hemoglobin A1C (e.g., % of patients with A1C within target), BMI (% of patients with BMI SDS > 3; % of patients with increase in BMI SDS), Surgical patients: SSI (% of patients with SSI), mortality	Patient charts	Chart review, reporting tools, quantitative	Regional or provincial data registries (e.g., Canadian Institute of Health Information), established patient registries, literature

BMI Body Mass Index
 CIHI Canadian Institute of Health Information
 SDS Standard Deviation Score
 SSI Surgical Site Infection
 VAS Visual Analog Scale

Table 2 Example for the Health Delivery domain, Privacy sub-domain: evaluation questions, measures, and data specifics

Health Delivery				
Example evaluation questions	Potential measures	Data sources	Collection strategy	Basis of comparison
Privacy				
To what extent do patients feel comfortable sharing openly with their provider during virtual visits? (e.g., as a function of the patient or provider’s home/work environment)	Patient and provider perception	Patient/provider surveys	Post-visit questionnaire	Literature review, other virtual care programs (e.g., mental health), in-person surveys
To what extent can we adhere to privacy and confidentiality standards?	Reported privacy breaches, privacy impact assessments, provider surveys	Safety reporting system, privacy breaches, privacy office audit/ assessments	Quantitative	National/ organizational privacy standards

Table 3 Example for the Individual Experience domain, Ease of use sub-domain: evaluation questions, measures, and data specifics

Individual Experience (patient, caregiver, provider, support staff)				
Example evaluation questions	Potential measures	Data sources	Collection strategy	Basis of comparison
Ease of use				
To what extent is the process for booking virtual care services easy to use?	Proportion of participants who felt booking virtual care was easy to use, # of deviations from standard booking process	Post-encounter questionnaires/ patient/provider/ clerk experience surveys, support calls, electronic booking system	Qualitative & quantitative	Satisfaction with in-person booking process
To what extent is the method of connecting (to patient/provider) easy to use?	Participant perceptions (patients, caregivers, providers), # of reported errors, # of support calls	Post-encounter questionnaires/ patient experience surveys	Qualitative & quantitative	Literature, established/ comparable telehealth programs

Table 4 Example for the Program Implementation domain, Infrastructure and technology subdomain: evaluation questions, measures, and data specifics

Program Implementation				
Example evaluation questions	Potential measures	Data sources	Collection strategy	Basis of comparison
Infrastructure & Technology (functionality, performance, security)^a				
To what extent has cybersecurity been reviewed and considered? (e.g., has a threat risk assessment (TRA) been conducted?)	# of access issues, adherence to cybersecurity standards, # of threats or security breaches	Cyber-review by 3rd party, TRA, cybersecurity monitoring tools	Quantitative	Applicable regional or national security guidance or directives (e.g., Canadian Centre for Cybersecurity guidance & directives)
Do providers and/or patients have appropriate access to information required to support care delivery? (e.g., access to whole patient chart, ability to communicate with other providers, or provide follow up information to patients)	Surveys to providers, staff, patients/families, safety reporting system, # of release of information requests	Surveys	Qualitative & quantitative	Other similar programs
To what extent does the technology support virtual care service delivery?	% of planned/unplanned downtime, # of requests for change/optimization, volume of technical issues, needs assessment/requirements evaluation	HelpDesk tickets/IT reports, project intake/enhancement requests, post-encounter questionnaires	Quantitative	Other similar programs

TRA Threat Risk Assessment

^a Note privacy is under health delivery

patient-centeredness, and workload burden. To ensure a holistic evaluation of any virtual care program (implemented or envisioned), patient, family/caregiver, provider, and support staff experiences are all considered essential components of high-quality care [7, 24]. To support consistent and comprehensive evaluation of the individual experience, the defining components of usability and themes that appear in validated survey tools (e.g., Telehealth Usability Questionnaire [36]) and CHIBE [31] are represented within the sub-domains. Usefulness is the degree to which a system or tool supports a desired function or goal (e.g., perception of how well the virtual care program supports access to care, clinical outcomes or reduces cost), [37] while ease of use speaks to how easy it is to use, regardless of utility. Interaction quality is the quality of the interpersonal interaction between patient/caregiver and provider that is facilitated by the virtual care system/tool, for example how well participants felt they were able to see, hear, and express themselves [36]. In alignment with the quadruple aim, [23, 24] sub-domains also include patient centeredness (e.g., whether patients should receive care in-person or virtually based on need or personal choice) and individual workload burden (e.g., patient and/or caregiver, provider workload or wellness). Table 3 presents example evaluation questions for the ease-of-use sub-domain along with associated measures, data sources and bases of comparison.

The fourth domain, Program Implementation, highlights key factors impacting system change and sustainability, and includes four sub-domains: leadership engagement and structure (institutional buy-in, provision of policy or guidance), resources (human and financial), training and support (availability and appropriateness), and infrastructure and technology (functionality, performance, and security). The last two sub-domains reflect the CHIBE framework [31]. Here, infrastructure and technology considerations relate to information systems and device capabilities rather than a technology usability evaluation, which falls under the individual experience domain. These four sub-domains are considered important to facilitate planning for implementation and to support decision-making regarding additions or changes to the reach and depth of virtual care programs over time. This domain does not emphasize the identification of KPIs and value definitions as in STEM, [10] because our framework is designed to identify measures and bases of comparison for each sub-domain. Once measures have been identified, these can be used as a repository from which to identify program-specific KPIs. Example evaluation questions, measures, data sources and bases of comparison for the infrastructure and technology sub-domain are presented in Table 4.

How to use this framework

The first step in using this evaluation framework is to define the objectives of the virtual care program. Excellent resources are available outlining how to develop program objectives. For example, program leaders could consider developing SMART aims following quality improvement methodology [38, 39]. The next step is to develop evaluation questions for each sub-domain that speak directly to the objectives of the virtual care program. For example, when selecting health outcomes evaluation questions, consider whether health outcome targets and impacts should be defined at the individual or population level, or both. For each evaluation question, identify measures that can be used to answer the question (e.g., hemoglobin A1C value as an outcome measure for children and youth with diabetes). Measures may come from encounter records (e.g., observations documented by healthcare providers) or patient-reported outcomes. The operational definitions of each measure should be described so data retrieved by different operators, processes and organizations are consistent and reproducible. Identify existing data sources to determine what data are already available versus need to be collected (e.g., electronic health record data or patient/provider survey results). For each source, consider an appropriate data collection strategy (e.g., qualitative, quantitative, or mixed methods) and the bases of comparison that can be used to determine success. Bases of comparison include an in-person care program, a comparable virtual care program (internal or external), clinical norms, or clinical guidelines as appropriate.

One of the key attributes of evaluation quality is feasibility. This framework provides an organized approach to deciding on outcomes of interest for different program stages (see WHO guidance [19]), priorities, and durations of evaluation (e.g., a program duration may be too short to assess health outcomes, but sufficient to assess program implementation). To demonstrate framework feasibility and how evaluation questions and measures tie back to specific program objectives, two pediatric endocrinologists applied the framework to an existing pediatric type 1 diabetes virtual care program (see Additional file 3). This allowed us to explore how the evaluation framework supported development of a program-specific evaluation plan. They appreciated the detailed guidance the framework provided, felt it was helpful in guiding a comprehensive and systematic evaluation plan, and that it was easy to use. The breadth of the 19 subdomains was deemed to cover all potential areas of interest for evaluation. This exercise also highlighted the need to consider which sub-domains are most relevant in unique contexts. While the pediatric type 1 diabetes virtual care program example illustrates a comprehensive evaluation plan for

one clinical area, and the framework's utility in developing program-specific evaluation questions and measures that tie back to program objectives, we recognize not all programs will have access to robust data sets for benchmarking or measurement. Furthermore, it is not practical for every evaluation to address all sub-domains of the framework in a single initiative. Thus, we believe it is important to consider all domains when scoping out an evaluation, but also that discrete elements of this framework can be used as required, depending on evaluation objectives, audience, and resources.

Discussion

Herein, we present a comprehensive virtual care evaluation framework that can be used by clinicians, researchers, and program leaders to guide evaluation of a broad range of virtual care programs and services at various stages of development, from planning through implementation and optimization. We describe its development by a multidisciplinary team using consensus methodology and provide instructions for use. Using consensus development panels was an efficient way to obtain results, as developing an evaluation framework was new to many team members due to their clinical and operational (rather than methodological) expertise. Moving forward, team members will be able to apply this knowledge to support development of program-specific virtual care evaluation plans within our institution.

We believe this framework can be used across a variety of virtual care programs, allowing for evidence synthesis and clear comparison of meaningful results. Using a standardized evaluation framework allows for the development of a common set of indicators and ensures program-specific evaluations can be generalized within and across health centres and clinical programs [40]. Ongoing consideration of all domains and sub-domains, even if the focus of a specific evaluation is only on a few, can support program development, implementation and/or optimization and set the stage to streamline a future evaluation (i.e., if a program is implemented with predetermined meaningful measures and data sources, planned or post-hoc analysis of process and outcome measures is simplified and relevant). We encourage using this framework at all stages of program development to guide continuous quality improvement and monitoring of patient and care delivery outcomes ensuring intended benefits are realized and unintended harm avoided. If the primary objective is a holistic evaluation of a virtual care program, we agree with the recommendation of SPROUT to include at least one measure from each domain [20, 41]. However, we feel our framework can also be useful from a research lens where the focus or objective may be within one domain at a time.

Our proposed framework has several strengths. First, it was developed by a multidisciplinary team with patient and family representation, and varied expertise and perspectives in clinical care, health systems, and evaluation. The consensus method used to develop the framework allowed for ongoing discussion about example questions, measures, and data sources in each domain, which led members to take ownership over the material and improved applicability of the framework to multiple contexts. We built upon existing literature to develop a user-friendly, comprehensive framework that considers nuances of pediatric virtual health care delivery. Second, the structure of this framework integrates the opportunity for benchmarking throughout, by requiring users to identify bases of comparison for each selected measure. Finally, application of this framework to an existing virtual care program demonstrated its utility and adaptability. Given the built-in flexibility of the framework, where sub-domains may be tailored to fit individual contexts, we believe it can be used for evidence synthesis across a variety of virtual care programs and support development of evaluation plans appropriate for either research or quality improvement initiatives. This agility is important given the often-overlapping goals of research and quality improvement projects particularly when evaluating health systems impact. Though this framework was developed for the pediatric context by a team with expertise in pediatric care, users have the flexibility to apply it to non-pediatric contexts by modifying or substituting measures as appropriate.

The main limitation to the work presented here is limited testing. Implementation of the framework has not been formally evaluated. Additional testing within an organization that did not develop the framework, or a program that was not represented in our multidisciplinary team, is required to generate validity evidence and further demonstrate its usability, applicability, and versatility.

This framework does not provide the guidance necessary to conduct an economic analysis that would consider the societal value of a particular virtual care program in the broader context of healthcare delivery. However, tools are available to support measuring the value of pediatric telehealth, [8] which we believe are complimentary.

We recognize that some elements of this framework may be difficult to measure robustly at a program or organization level, and the importance of establishing a standardized framework and measures that can be compared across programs, organizations, and institutions. Specifically, risks associated with virtual care, such as wrong, missed, or delayed diagnoses due to the inability to physically examine patients or inability to address emergencies that arise during virtual encounters, may not be easily measured. In an effort to contribute to this

important safety evaluation, two members of our team (EG, SV) are leading a national surveillance project for adverse events related to virtual care [42].

We recognize the dynamic nature of virtual care, where innovative technologies, patient and family preferences and needs, and evidence can change and emerge rapidly. Naturally, our framework will need to be adapted for different settings and updated with advancements in this field. Furthermore, while standardized data definitions will facilitate meaningful comparisons and benchmarking, defining these was beyond the scope of our working group and would require broader stakeholder input.

Conclusion

Our aim in developing a pediatric virtual care evaluation framework is to support development of a standardized approach for evaluating virtual care delivery models to ensure high quality programs and services. Our proposed methodology supports and expands on previously published virtual care evaluation frameworks by providing additional detail, including evaluation sub-domains, and a structure that facilitates practical application.

We believe this framework will be useful in a variety of healthcare contexts. Future research studying the implementation of this framework is required to generate evidence supporting its utility at the program, organization, and system-level. We anticipate a close collaboration with SPROUT and others towards a unified approach for pediatric virtual care evaluation, development of standardized data definitions, benchmarking, knowledge sharing, and ultimately the best health outcomes for our patients.

Abbreviations

AAP	American Academy of Pediatrics
AHRQ	Agency for Healthcare Research and Quality
CHIBEF	Canada Health Infoway Benefits Evaluation Framework
COMET	Core Outcome Measures in Effectiveness Trials
COVID-19	COronalVirus Disease of 2019
KPIs	Key Performance Indicators
NQF	National Quality Forum
QoL	Quality of Life
SARS CoV-2	Severe Acute Respiratory Syndrome Coronavirus 2
SMART	Specific, Measurable, Achievable, Relevant, and Time-bound
SPROUT	Supporting Pediatric Research on Outcomes and Utilization of Telehealth
STEM	SPROUT Telehealth Evaluation and Measurement
WHO	World Health Organization

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12887-023-04229-1>.

Additional file 1. Elements of an Evaluation Framework.

Additional file 2. Virtual Care Evaluation Framework.

Additional file 3. Example Virtual Care Evaluation Plan for Type 1 Diabetes clinic, integrated EHR context.

Acknowledgements

We would like to thank CHEO and CHEO Research Institute Executive Leadership for sponsoring this project and providing the resources necessary to see it through. Jim Bottomley's project management support was much appreciated, particularly at the outset of the project. Further, we are grateful to Dr. John Chuo, Pediatric Neonatologist at the Children's Hospital of Pennsylvania and co-author of the American Academy of Pediatrics (AAP) Section on Telehealth Care's SPROUT (Supporting Pediatric Research on Outcomes and Utilization of Telehealth) telehealth evaluation & measurement (STEM) framework, for enthusiastically reviewing a final draft of this manuscript and providing positive and helpful feedback in advance of submission.

Authors' contributions

CD, EG, and WJK conceptualized the project. CD and EG led the project. SS provided guidance on methods and conduct of consensus development panels. TA, MB, PC, CD, LF, EG, HH, RJ, CK, WJK, JL, CM, NM, KP, SS and CZ participated in the consensus development panels that produced the framework. CD, EG, SS and SV drafted, reviewed, and refined the manuscript. All authors critically revised and approved the final manuscript.

Funding

Shelley Vanderhout was supported by a CIHR Health Systems Impact Postdoctoral Fellowship award. CHEO Foundation, CHEO Research Institute and CHEO provided funding for human resources to support this work. The funding bodies did not have a role in the design of the study, data collection, analysis or interpretation, or in writing the manuscript.

Availability of data and materials

Not applicable.

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

¹Children's Hospital of Eastern Ontario, 401 Smyth Rd, Ottawa, ON K1H 8L1, Canada. ²Children's Hospital of Eastern Ontario Research Institute, 401 Smyth Rd, Ottawa, ON, Canada. ³School of Epidemiology and Public Health, University of Ottawa, 600 Peter Morand Crescent, Room 101, Ottawa, ON, Canada. ⁴Faculty of Medicine, University of Ottawa, 451 Smyth Road, Ottawa, ON, Canada. ⁵Children's Hospital of Eastern Ontario Research Institute Family Leader Program, 401 Smyth Rd, Ottawa, ON, Canada.

Received: 30 June 2022 Accepted: 2 August 2023

Published online: 17 August 2023

References

- Hollander JE, Carr BG. Virtually Perfect? Telemedicine for Covid-19. *N Engl J Med*. 2020;382(18):1679–81. <https://doi.org/10.1056/NEJMp2003539>.
- Canadians' Health Care Experiences During COVID-19 | Canada Health Infoway. <https://www.infoway-inforoute.ca/en/component/edocman/resources/reports/benefits-evaluation/3828-canadians-health-care-experiences-during-covid-19>. Accessed 16 Feb 2022.
- Digital Health Canada - Virtual Care Resources. <https://www.digitalhealthcanadamembers.com/Virtual-Care-Resources>. Accessed 1 Mar 2022.
- Digital Health Canada. Virtual Care in Canada: Lexicon. Published online 2021. <https://digitalhealthcanada.com/wp-content/uploads/2022/06/Virtual-Care-National-Lexicon-v-JAN2521.pdf>.
- Nanda M, Sharma R. A Review of Patient Satisfaction and Experience with Telemedicine: A Virtual Solution During and Beyond COVID-19 Pandemic. *Telemed E-Health*. 2021;27(12):1325–31. <https://doi.org/10.1089/tmj.2020.0570>.

6. Canada Health Infoway. Canadian Digital Health Survey: What Canadians Think. <https://www.infoway-inforoute.ca/en/component/edocman/3856-canadian-digital-health-survey-what-canadians-think/view-document?Itemid=0>. Accessed 6 Jun 2022.
7. Curfman A, McSwain SD, Chuo J, et al. Pediatric Telehealth in the COVID-19 Pandemic Era and Beyond. *Pediatrics*. 2021;148(3):e2020047795. <https://doi.org/10.1542/peds.2020-047795>.
8. Curfman A, McSwain SD, Chuo J, Olson CA, Simpson K. An Economic Framework to Measure Value of Pediatric Telehealth. *Telemed E-Health*. 2021;27(12):1440–2. <https://doi.org/10.1089/tmj.2020.0520>.
9. Shaker MS, Oppenheimer J, Grayson M, et al. COVID-19: Pandemic Contingency Planning for the Allergy and Immunology Clinic. *J Allergy Clin Immunol Pract*. 2020;8(5):1477–1488.e5. <https://doi.org/10.1016/j.jaip.2020.03.012>.
10. Chuo J, Macy ML, Lorch SA. Strategies for Evaluating Telehealth. *Pediatrics*. 2020;146(5):e20201781. <https://doi.org/10.1542/peds.2020-1781>.
11. Shaw J, Jamieson T, Agarwal P, Griffin B, Wong I, Bhatia RS. Virtual care policy recommendations for patient-centred primary care: findings of a consensus policy dialogue using a nominal group technique. *J Telemed Telecare*. 2018;24(9):608–15. <https://doi.org/10.1177/1357633X17730444>.
12. Young JD, Abdel-Massih R, Herchline T, et al. Infectious Diseases Society of America Position Statement on Telehealth and Telemedicine as Applied to the Practice of Infectious Diseases. *Clin Infect Dis Off Publ Infect Dis Soc Am*. 2019;68(9):1437–43. <https://doi.org/10.1093/cid/ciy907>.
13. Canadian Medical Association, Royal College of Physicians and Surgeons of Canada, College of Family Physicians of Canada. VIRTUAL CARE RECOMMENDATIONS FOR SCALING UP VIRTUAL MEDICAL SERVICES. Published online February 2020. <https://www.cma.ca/sites/default/files/pdf/virtual-care/ReportoftheVirtualCareTaskForce.pdf>. Accessed June 2022.
14. Newcomer KE, Hatry HP, Wholey JS. Handbook of Practical Program Evaluation. 2015. <http://site.ebrary.com/id/11090807>. Accessed 16 Feb 2022.
15. Parsons B, Jessup P. The Logic Model Guidebook: Better Strategies for Great Results (Second Ed.): Book Review. *The Foundation Review*. 2012;4(3). <https://doi.org/10.4087/FOUNDATIONREVIEWD-12-00050.1>
16. Alkin MC. Evaluation Essentials: From A to Z. Guilford Press; 2010. <http://public.eblib.com/choice/publicfullrecord.aspx?p=581949>. Accessed 9 Mar 2022.
17. Wang CJ, Liu TT, Car J, Zuckerman B. Design, Adoption, Implementation, Scalability, and Sustainability of Telehealth Programs. *Pediatr Clin North Am*. 2020;67(4):675–82. <https://doi.org/10.1016/j.pcl.2020.04.011>.
18. Ontario Health Quality. Adopting and Integrating Virtual Visits into Care: Draft Clinical Guidance. 2020. https://quorum.hqontario.ca/Portals/0/Users/170/54/10154/Draft%20Clinical%20Guidance_Adopting%20and%20Integrating%20virtual%20visits%20into%20care_V1.pdf?ver=2020-03-13-091936-370. Accessed 6 Jun 2022.
19. WHO | Monitoring and evaluating digital health interventions. WHO. <http://www.who.int/reproductivehealth/publications/mhealth/digital-health-interventions/en/>. Accessed 9 Mar 2022.
20. Chuo J, Olson C, McSwain D, Macy M. Using the STEM Framework to Evaluate Digital Health Interventions in Pediatric Care. Presented at: Pediatric Academic Societies; Denver, CO. 2022. <https://2022.pas-meeting.org/searchGlobal.asp>. Accessed June 2022.
21. NQF: Creating a Framework to Support Measure Development for Telehealth. https://www.qualityforum.org/Publications/2017/08/Creating_a_Framework_to_Support_Measure_Development_for_Telehealth.aspx. Accessed 5 Apr 2022.
22. Totten AM, Womack DM, Eden KB, et al. Telehealth: Mapping the Evidence for Patient Outcomes From Systematic Reviews. Agency for Healthcare Research and Quality (US); 2016. <http://www.ncbi.nlm.nih.gov/books/NBK379320/>. Accessed 5 Apr 2022.
23. Bodenheimer T, Sinsky C. From triple to quadruple aim: care of the patient requires care of the provider. *Ann Fam Med*. 2014;12(6):573–6. <https://doi.org/10.1370/afm.1713>.
24. Sikka R, Morath JM, Leape L. The Quadruple Aim: care, health, cost and meaning in work. *BMJ Qual Saf*. 2015;24(10):608–10. <https://doi.org/10.1136/bmjqs-2015-004160>.
25. Goldbloom E, Buba M, Bhatt M. Innovative virtual care delivery in a Canadian pediatric tertiary care centre. *Paediatr Child Health: Published online; 2022*.
26. Waggoner J, Carline JD, Durning SJ. Is There a Consensus on Consensus Methodology? Descriptions and Recommendations for Future Consensus Research. *Acad Med J Assoc Am Med Coll*. 2016;91(5):663–8. <https://doi.org/10.1097/ACM.0000000000001092>.
27. Halcomb E, Davidson P, Hardaker L. Using the consensus development conference method in healthcare research. *Nurse Res*. 2008;16(1):56–71. <https://doi.org/10.7748/nr2008.10.16.1.56.c6753>.
28. Fink A, Kossecoff J, Chassin M, Brook RH. Consensus methods: characteristics and guidelines for use. *Am J Public Health*. 1984;74(9):979–83. <https://doi.org/10.2105/ajph.74.9.979>.
29. Nair R, Aggarwal R, Khanna D. Methods of formal consensus in classification/diagnostic criteria and guideline development. *Semin Arthritis Rheum*. 2011;41(2):95–105. <https://doi.org/10.1016/j.semarthrit.2010.12.001>.
30. Dodd S, Clarke M, Becker L, Mavergames C, Fish R, Williamson PR. A taxonomy has been developed for outcomes in medical research to help improve knowledge discovery. *J Clin Epidemiol*. 2018;96:84–92. <https://doi.org/10.1016/j.jclinepi.2017.12.020>.
31. Benefits Evaluation Indicators Technical Report | Canada Health Infoway. <https://www.infoway-inforoute.ca/en/component/edocman/resources/reports/450-benefits-evaluation-indicators-technical-report-version-2-0>. Accessed 1 Mar 2022.
32. Hajesmaeel-Gohari S, Bahaadinbeigy K. The most used questionnaires for evaluating telemedicine services. *BMC Med Inform Decis Mak*. 2021;21(1):36. <https://doi.org/10.1186/s12911-021-01407-y>.
33. Danseco E, Kurzawa J, Sundar P, Brown J, Huang C. Evaluating the sector-wide implementation of virtual child and youth mental health services in response to the COVID-19 pandemic: Perspectives from service providers, agency leaders and clients. *Implement Res Pract*. 2021;2:26334895211045690. <https://doi.org/10.1177/26334895211045690>.
34. Ellis LA, Meulenbroeks I, Churrua K, et al. The Application of e-Mental Health in Response to COVID-19: Scoping Review and Bibliometric Analysis. *JMIR Ment Health*. 2021;8(12):e32948. <https://doi.org/10.2196/32948>.
35. Vanderhout SM, Li A, Goldbloom E, Newhook D, Garcia M, Dulude C. Evaluation strategies for understanding experiences with virtual care in Canada: mixed methods study. *J Med Internet Res*. 2023 (in press).
36. Parmanto B, Allen Nelson Lewis J, Graham KM, Bertolet MH. Development of the Telehealth Usability Questionnaire (TUQ). *Int J Telerehabilitation*. 2016;8(1):3–10. <https://doi.org/10.5195/ijt.2016.6196>.
37. Davis FD. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Q*. 1989;13(3):319–40. <https://doi.org/10.2307/249008>.
38. Centers for Disease Control and Prevention. Developing Program Goals and Measurable Objectives. <https://www.cdc.gov/std/program/pupestd/developing%20program%20goals%20and%20objectives.pdf>. Accessed 6 Jun 2022.
39. Agency for Healthcare Research and Quality. What Are the Objectives for Your Health Care Quality Report?. 2018. <https://www.ahrq.gov/talkingquality/plan/objectives.html>. Accessed 6 Jun 2022.
40. Posavac EJ. Program Evaluation: Methods and Case Studies (8th ed.). Routledge; 2010. <https://doi.org/10.4324/9781315664972>.
41. Chuo J. Measuring Telehealth in Clinical Care and Research: The STEM Framework. Presented at: AAP Tele-Research ECHO Program; 2020. https://players.brightcove.net/6056665225001/default_default/index.html?videoId=6216421366001. Accessed 12 May 2022.
42. Adverse events related to virtual care - CPSP. <https://cpsp.cps.ca/surveillance/study-etude>. Accessed 26 Mar 2023.

Publisher's Note

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