

## ESSAY

# Improving team effectiveness using a program evaluation logic model: case study of the largest provincial simulation program in Canada

Alyshah Kaba<sup>1</sup>, Theresa Cronin<sup>1</sup>, Walter Tavares<sup>2</sup>,  
Tanya Horsley<sup>3</sup>, Vincent J Grant<sup>1,✉</sup>, Mirette Dube<sup>1</sup>

<sup>1</sup>eSIM Provincial Simulation Program, Alberta Health Services, Calgary, Alberta, Canada

<sup>2</sup>The Wilson Centre and Post MD Education, Post Graduate Medical Education & Continuing Professional Development, Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada

<sup>3</sup>Research Unit, Royal College of Physicians and Surgeons of Canada, Ottawa, Ontario, Canada

**Corresponding author:** Alyshah Kaba, [alyshah.kaba@albertahealthservices.ca](mailto:alyshah.kaba@albertahealthservices.ca)

<https://ijohs.com/article/doi/10.54531/FQZQ4032>

## ABSTRACT

Historically simulation-based education (SBE) has primarily focused on program development and delivery as a means for improving the effectiveness of team behaviours; however, these programs rarely embed formal evaluations of the programs themselves. Logic models can provide simulation programs with a systematic framework by which organizations and their evaluators can begin to understand complex interprofessional teams and their programs to determine inputs, activities, outputs and outcomes. By leveraging their use, organizational leaders of simulation programs can contribute to both *demonstrating value* and *impact* to healthcare teams, in addition to establishing a growing culture of evaluation at any health system level. This case study describes a complex program evaluation for improving team effectiveness *outputs* and *outcomes* across more than one simulation program, discipline, speciality, department in the largest health authority in Canada and provides considerations for other simulation programs globally to advance the science of program evaluation within the SBE community.

### What this essay adds

- There is a paucity of formal program evaluation studies in simulation-based education that demonstrate its impact in consistently improving team effectiveness outcomes across more than one program, discipline, speciality, department and health system.
- The application program evaluation logic models (i.e. a visual tool such as an 'if-then' representation of a program) is used to shape the development and evaluation strategy of a simulation program life cycle.
- Logic models can provide simulation programs with a framework by which organizations and their evaluators can begin to understand and dissect complex interprofessional teams and their programs evaluation strategy to determine inputs, activities, outputs and outcomes.
- The key is that the educational programs like simulation-based education can help in *demonstrating value* to a healthcare organization if its *impact* can be demonstrated systematically using a logic model.
- By leveraging the use of logic model, organizational leaders of simulation programs can contribute to establishing and growing their culture of evaluation at any health system level.
- There is no prescriptive method for conducting program evaluation in simulation; rather, we suggest use the logic model approach as a road map for the program evaluation process, encouraging the simulation community globally to move beyond asking *whether a program worked*, to establishing *how it worked* and *why it worked* and *what else happened*.

## Introduction

Simulation has emerged as an effective method to practice, reflect on, and improve Interprofessional Collaboration (IPC) and team effectiveness behaviours that can lead to safer patient care, staff safety and higher quality outcomes [1–3]. Historically simulation-based education (SBE) has primarily focused on a program development and delivery model as a means for improving the effectiveness of team behaviours [4–7]; however, these programs rarely embed formal evaluations of the programs themselves [8,9].

There is a paucity of program evaluation studies in SBE that demonstrate its impact in consistently improving team effectiveness outcomes across more than one program, discipline, speciality, department and health system. As a result, simulation programs are left without an established approach or tool to evaluate the scale of their overall impact at the organizational level [10].

Utilizing an approach that supports *program evaluation*, logic models are helpful tools in evaluating the impact of the simulation program in this local context where the environment is complex and has several covariates and where traditional research-based approaches are challenged [11,12]. The application of program evaluation and logic models (i.e. a visual tool such as an 'if-then' representation of a program) is used to shape the development and evaluation strategy of a program [13–15]. Logic models can provide simulation programs with a framework by which organizations and their evaluators can begin to understand and dissect complex interprofessional teams and their programs to determine inputs, activities, outputs and outcomes that demonstrate the value to an organization [16,17]. By leveraging their use, organizational leaders of

simulation programs can contribute to (a) demonstrating their *value* to the organization and (b) establishing and growing their culture of evaluation at any health system level [16,17].

The *goal of this paper* is to describe a case study of a complex program evaluation and logic model for improving team effectiveness *outputs* and *outcomes* across more than one simulation program, discipline, speciality, department in the largest health authority in Canada and provides considerations for other simulation programs globally to tailor these evaluation approaches to their own institutions to further advance the science of program evaluation within the SBE community.

### Background literature: program evaluation and simulation

Over the last 20 years there has been burgeoning of literature from health professional educational programs applying theoretical and outcomes evaluation frameworks such as Kern's and Kirkpatrick's [18,19]; yet, many simulation programs still struggle with only capturing lower levels of outcomes and outputs data (e.g. learners reactions, knowledge and attitudes) [20–23] but are unable to demonstrate behaviour change or system level impacts [24–26].

Further, a scoping review by Batt et al. found that most single studies in health professional education literature examine educational effectiveness at the individual learner level but less commonly explore outputs and outcomes on healthcare teams [27]. This may leave simulation programs as the less desirable applicants to those competing for resource allocation (e.g. funds, human resources, space, etc.). Even with the highest quality SBE

programming, without a clear demonstration of impact to an organization, there is a risk of losing program support and lost opportunity to share evidence of program effectiveness and sustainability [27]. Despite the obvious intuitive link, and the importance of establishing a culture of evaluation within SBE, there are many reasons why this may not occur including lack of time, lack of evaluation expertise on the team, variability in assessment measures, lack of comparators, the number of consistently changing variables in a complex healthcare system, in addition to a misunderstanding that program evaluation is only relevant for well-established simulation programs [16,28–30].

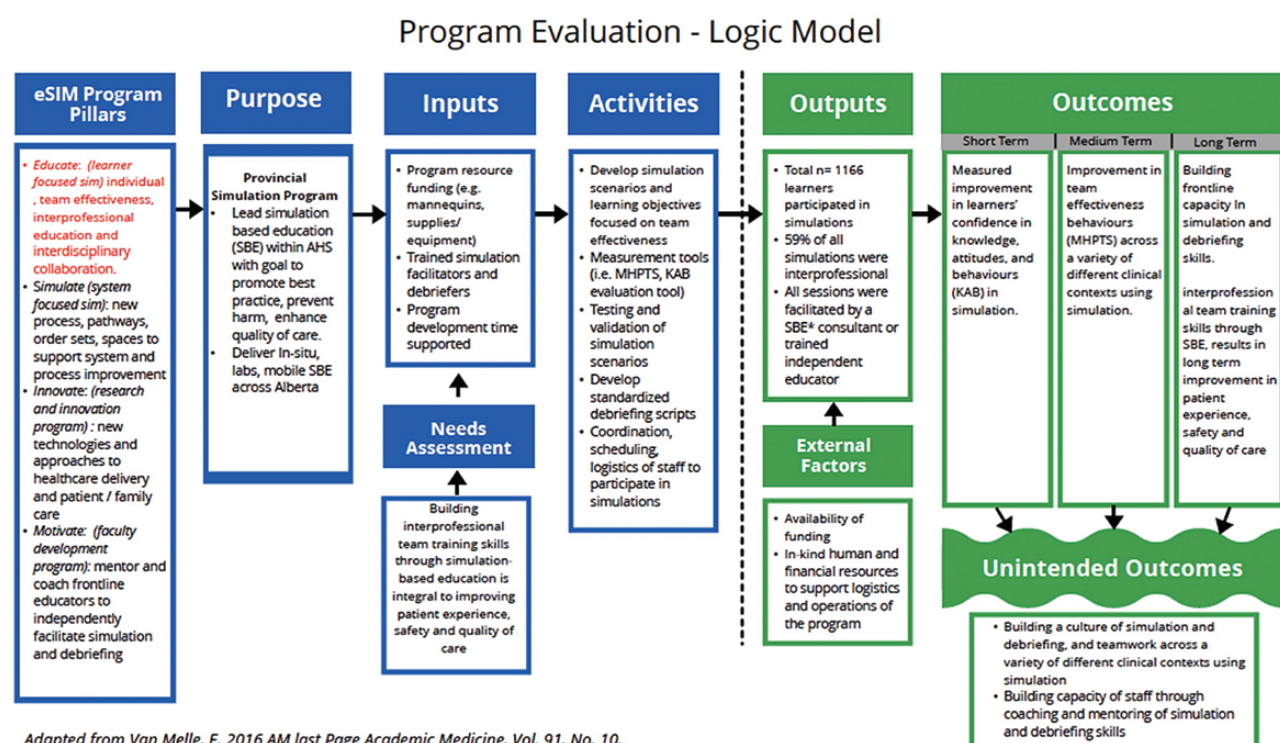
Despite this, it is *never* too early or too late to start evaluating your SBE program [15,31]. Program evaluation provides a systematic approach to measure the *impact* of SBE program outcomes and evaluate a program's implementation. In this way, program evaluation is a more organized approach to examine a program's *outcomes* ('Does it work?') and/or process ('How or why does it work?'), at any stage of the simulation program development cycle [32]. The logic model (Figure 1) supports a systematic method for identifying key questions, as well as collecting, analysing and using information to assess your simulation program outcomes and/or process [15]. There are several key features of a logic model which include the *inputs* that refer to resources that are deemed to be necessary for the simulation education program to have its desired outcome or to achieve its intended purpose [33]. *Activities* capture the critical components of the program – what it is that you are doing (with the inputs) that are allowing a simulation program to achieve its purpose directly or indirectly [33]. An *output* is the tangible product or service that arises as a result of the program activities. (i.e. products or things you can count)

[33]. An *outcome* is a change that occurs as a result of an individual's exposure to the simulation program [33].

Applied to SBE programs, examples of logic model *inputs* include data collection measures, human resources (i.e. faculty facilitators, context experts), simulation space (in lab or *in situ*), participants, etc. *Activities* include simulation scenarios and debriefing, and *outputs* include changes that happen as result of the SBE. These include numbers that you can count (i.e. number of participants, sessions, latent safety threats, etc.) The short-, medium- and long-term outcomes can include overall program evaluation at the macro level, but can also be broken down by specific program goals (i.e. change in knowledge skills, team behaviours) [34]. These elements of the logic model can serve as a road map for the program evaluation process, which allow the simulation community to move beyond asking *whether a program worked*, to establishing *how it worked* and *why it worked* and *what else happened* [32].

There is no prescriptive approach to program evaluation studies for simulation programs [34] but there are evaluation principles and guidelines that can be applied from other disciplines (e.g. realist, developmental and appreciative inquiry, etc.), each having their own epistemological and methodological considerations that underpin the design and unique limitations, which can be integrated in the evaluation of simulation programs [16,31,35–37]. For example, by understanding the *context, mechanisms and outcomes* of SBE interventions, a realist evaluation framework can provide a deeper level of understanding of what types of IPC simulations work for whom in what circumstances [30,38]. Simply participating in this process can engage organizational leaders in productive discussions and debates, generate ideas, support deliberations, identify relationships and provide

Figure 1: Logic model.



Adapted from Van Melle, E. 2016 AM last Page Academic Medicine, Vol. 91, No. 10.

opportunities to review strengths and weaknesses of the simulation program [14,31,34].

In this paper we propose two unique contributions to the literature (a) demonstration of a successful case study whereby we provide evidence of simulation's value in applying a program evaluation approach to improving team effectiveness at the healthcare system level, across multiple programs, professional groups and hospital sites (b) describe a framework for how to go about applying program evaluation using a logical model to share your simulation programs impact.

## Methods

### Case study

The healthcare system in Alberta serves a population of more than 4.3 million and is organized into five geographic regions that are referred to as zones: South, Calgary, Central, Edmonton and North. eSIM (educate, simulate, innovate, motivate) is the provincial simulation program for part of the larger health authority serving a geographic area = 661,848 km<sup>2</sup> offering services to an array of over 15 health professional disciplines, 147 programs, 650 facilities and over 102,700 staff and 8,400 physicians [39]. There are several hundred programs with both clinical (i.e. physicians, nurses, allied health) and non-clinical (i.e. protective services, housekeeping, portering, etc.) team members that engage in simulation-based activities using the services of eSIM across Alberta.

Based on eSIM service delivery model and infrastructure, the program made an early investment of time and effort into identify, understand and engage stakeholders with the aim of enhancing continuous evaluation efforts and areas of focus to support organizational learning specific to team effectiveness. The four key pillars for the eSIM Provincial Simulation Program: 1. Educate (i.e. learner-focused simulation); 2. Simulate (i.e. system-focused simulation); 3. Innovate (i.e. research and innovation) and 4. Motivate (i.e. faculty development program) [39].

Therefore, for the purpose of this case study, the authors focused the context of program evaluation only on eSIM Program Pillar 1 'Educate' which is primarily targets *learner-focused simulations*, individual, team effectiveness, interprofessional education and interdisciplinary collaboration (Figure 1). These pillars were developed in consultation with key simulation champions across a large provincial health authority who were engaged in simulation practices. Specifically, in Alberta, a targeted needs assessment with sites, staff and leadership revealed that IPC, teamwork training and communication was priority for all acute care and inpatient settings and vital to patient safety and quality of care. Simulation was identified as an education resource to support this need by offering an ability for frontline teams to practice, reflect on team effectiveness behaviours for safer patient care.

The program evaluation logic model (Figure 1) describes both the processes and outcomes specific to improving team effectiveness across more than one simulation program, disciplines, specialities, departments.

### Data collection

Two outcome measures were used to measure short-term, medium-term and long-term outcomes for Pillar 1. Both measures were administered to interprofessional frontline teams participating in SBE across Alberta:

1. **Team Effectiveness Evaluation (MHPTS):** The measurement tool was administered twice, after two consecutive simulation sessions, and paired samples *t*-tests were used to analyse the difference between the two scores for the provincial data. The tool is based on the validated Mayo High Performance Teamwork Scale (MHPTS) [40] which was selected as the 8-item teamwork behaviour constructs in the MHPTS (e.g. leadership, situational awareness, communication, etc.) were generalizable provincially across multiple teams, sites, practice areas and sites. The MHPTS consists of a behavioural anchored scale of 0 (never or rarely), 1 (inconsistently), 2 (consistently). This instrument was a previously validated tool with a Cronbach's alpha (0.85) suggesting an excellent internal consistency of a measure (i.e. items represent the construct of teamwork) used to measure participants behavioural change in teamwork after the simulation intervention. In addition, this instrument of was selected because of usability of the tool, limited number of items and the time it took for participants to complete the survey.
2. **Learner Evaluation (KAB):** The Knowledge Attitudes and Behaviour (KAB) instrument was developed by the eSIM Provincial Simulation team and administered to learners before a simulation session and repeated after the session to evaluate any changes in confidence in knowledge, attitudes and team behaviours between the two time points. This evaluation was developed as there was no validated tool for practicing health professionals that measured confidence in all three constructs of teamwork knowledge, attitudes and behaviours in simulation, which was an identified gap in the needs assessment conducted by the eSIM provincial program with key stakeholders (e.g. clinical team leaders, educators), simulation champions, physicians, staff and leadership across the health authority.

## Results

Statistical analysis showed that mean score for every question on the *Team Effectiveness Evaluation (MHPTS)* increased significantly from pre-session (1.58, 0.30) to post-mean score (1.81, 0.29). The number of interprofessional participants (nursing, physician, EMS, allied health) *n* = 284 which represented both acute care and inpatient settings. Nurses were the largest group of participants representing 60.6% of the sample. Most of the sessions were eSIM consultant-supported (97.4%) and simulation sessions took place in either simulation labs or patient cares where healthcare professionals work, across all five zones in Alberta.



All  $p$ -values for team behaviours were highly statistically significant at  $t(283) = 6.32, p \leq 0.000, d = 0.77$  with a medium effect size. This implies that we can consistently produce highly effective interprofessional teams across a variety of different clinical contexts by teaching teamwork behaviours using simulation (Table 1).

Statistical analysis for the *Learner Evaluation (Knowledge, Attitudes, Behaviours – KAB)* showed that the mean score for every question on the Learner Evaluation increased significantly from pre-session (3.78, 0.27) to post (4.29, 0.26) mean score. The number of interprofessional participants (nursing, physician, EMS, allied health)  $n = 882$  which represented both acute care and inpatient settings. Nurses were the largest group of participants representing

66.6% of the sample. Most of the sessions were eSIM consultant-supported (94.7%) and simulation sessions took place in either simulation labs or patient cares where healthcare professionals work, across all five zones in Alberta.

All  $p$ -values for learners' confidence were highly statistically significant at  $t(881) = 7.45, p \leq 0.000, d = 0.94$ . This implies that the simulation sessions were highly effective at improving participants' confidence in their knowledge, attitudes and behaviours. The results of the Learner Evaluation showed that simulation increased participants' confidence (KAB) in their ability to execute procedures and interventions as a team to improve quality patient safety and patient experience (Table 1).

**Table 1:** Statistical analysis for learner and team effectiveness evaluations

Learner Evaluation (KAB)			Team Effectiveness Evaluation (MHPTS)		
I feel confident in my ability to... $n = 882$	Mean, SD, $t$ -static	$p$ -value	Team behaviour $n = 284$	Mean, SD, $t$ -static	$p$ -value
Participate as a team leader or follower	Mean -0.58 SD 0.82 $t$ -statistic -19.98	0.000	1. A leader is clearly recognized by all team members.	Mean -0.24 SD 0.53 $t$ -statistic -7.32	0.000
Delegate and be receptive to direction	Mean -0.46 SD 0.73 $t$ -statistic -17.79	0.000	2. The team member assures maintenance of an appropriate balance between command authority and team member participation.	Mean -0.22 SD 0.57 $t$ -statistic -6.19	0.000
Understand my role and fulfil responsibilities as part of the team	Mean -0.54 SD 0.82 $t$ -statistic -18.42	0.000	3. Each team member demonstrates a clear understanding of his or her role.	Mean 0.16 SD 0.57 $t$ -statistic -4.54	0.000
Recognize a change in clinical status or deteriorating situation	Mean -0.44 SD 0.68 $t$ -statistic -18.04	0.000	4. The team prompts each other to attend to all significant clinical indicators throughout the procedure/intervention.	Mean -0.25 SD 0.54 $t$ -statistic -7.74	0.000
Work collaboratively with patients and families to improve patient experience	Mean -0.31 SD 0.64 $t$ -statistic -13.37	0.000	5. When team members are actively involved with the patient, they verbalize their activities aloud.	Mean 0.21 SD 0.56 $t$ -statistic -5.99	0.000
Communicate effectively by addressing members directly, repeating back and seeking clarity	Mean -0.54 SD 0.76 $t$ -statistic -20.18	0.000	6. Team members repeat or paraphrase instructions and clarifications to indicate that they heard them correctly.	Mean -0.26 SD 0.61 $t$ -statistic -6.69	0.000
Understand when and how to use available equipment	Mean -0.61 SD 0.77 $t$ -statistic -22.32	0.000	7. Team members refer to established protocols and checklists for the procedure/intervention.	Mean -0.13 SD 0.55 $t$ -statistic -3.52	0.000
Refer to established protocols and checklists for the procedure/intervention	Mean -0.57 SD 0.81 $t$ -statistic -19.85	0.000	8. All members of the team are appropriately involved and participate in the activity.	Mean -0.17 SD 0.44 $t$ -statistic -6.26	0.000
Speak up and voice my concerns as appropriate in a clinical event	Mean -0.57 SD 0.77 $t$ -statistic -21.03	0.000			
Know when to seek additional resources and call for help when necessary	Mean -0.50 SD 0.73 $t$ -statistic -19.25	0.000			
<b>Total score</b>	Mean -0.51 SD 0.55 $t$ -statistic -26.50	0.000	<b>Total score</b>	Mean 0.22 SD 0.29 $t$ -statistic -12.00	0.000

## Discussion

The findings from this case study demonstrate building a sustainable and impactful simulation program, regardless of size or breath of programming or services, requires thoughtful consideration for program evaluation activities [16,41]. By evaluating team effectiveness behaviours across different programs, discipline, speciality and department, we provided a program evaluation model that engaged leaders and simulation champions across local sites that could be generalized and adopted broadly across the entire healthcare system, regardless of its size. Everyone was *'rowing in the same direction'*, towards greater team effectiveness' and were part of something larger than improving the teamwork within their individualized simulation programs, which is paramount to safe, quality healthcare and optimal patient outcomes. This *'institutionalizing'* of program evaluation was a strategy for continuous improvement that enabled ongoing engagement, sustainability and organizational learning, as additional programs adopted simulation. Traditionally, simulation programs often overlook opportunities for program evaluation to demonstrate their value or impact to organizations or more specifically, those who are funding the program [12,41]. This is evident in the lack of literature demonstrating program evaluation for simulation programs and their impact at an organization level – across multiple teams, professional groups and even hospital sites [22,24,42,43]. Nonetheless, demonstrating value is the key to successful simulation program delivery, growth and potential future revenue generation to support ongoing resource allocation within an organization [44].

Despite the complexity of variables in our case study such as team cultures, difference in clinical practices across urban and rural sites, and varying acuity and experience levels of staff, which are often overarching complexities that obligates program evaluation studies, there was a statically significant improvement in knowledge attitude and teamwork behaviours across both the Learner Evaluation ( $n = 882$ ) and the Team Effectiveness Evaluation ( $N = 284$ ). Even more critical than these short-term and medium-term outcomes (Figure 1) is the intention to be transparent in sharing these outcomes from simulation program evaluations studies to leadership and key stakeholders within the healthcare organization to demonstrate the *value* of how SBE is critical to improving safer patient care and staff safety [34].

When evaluating simulation education programs specifically, this case study highlights that program developers can use a logic model to organize and articulate program components with the ultimate intent of identifying evaluation questions. Logic models also provide a structure to explore and explain relationships between one or more theoretical models [33]. For example, in our case study we used complexity theory, which emphasizes that interactions are constantly changing and unpredictable [45]. We applied complexity theory to build our provincial simulation program and identify key questions on simulation's role in improving team effectiveness. We used a logic model, to inform our process (inputs, activities/resources) and

the ability to demonstrate impact (outputs and outcomes) where we expected it, and where we did not based on the limitations/external factors that influenced our results. Overall, the logic model provides a systematic approach to study the program evaluation process while also contributing to the evidence.

Given that healthcare education interventions are not singular entities and consist of myriad of components interacting in a complex healthcare system, there were several unintended consequences of this simulation program evaluation. First recognized by Michael Scriven in 1970 as *'emergence'* [46], the unintended consequences of eSIM program evaluation resulted in a creating provincial culture of simulation, debriefing and teamwork across a variety of different clinical contexts, in addition to building capacity of staff through coaching and mentoring of simulation and debriefing skills. This approach to capturing emergent outcomes, recognized *'what else happened'*, as result of the program evaluation, whether these outcomes were intended or not within complex healthcare systems [32].

In summary, a program evaluation and logic models are a helpful tools for any size of simulation program to plan their evaluation strategy looking at the programs purpose, inputs, activities, outputs and outcomes [34]. As our provincial simulation program continues to expand, so will our evaluation strategies within the various pillars and also serve as an engagement strategy that is institutionalized into the daily program evaluation activities of the organization. As this was the first step in establishing program evaluation within one of the eSIM pillars, the outcomes from this program evaluation focused on team effectiveness will inform future program evaluations within each of the individuals pillars (e.g. systems, faculty development, etc.) of the provincial simulation program in Alberta. In sharing our measurement approach and logic model our formulation is not meant to be prescriptive method for conduction program evaluation, rather we use these elements as a road map for the program evaluation process, which allow the simulation community to move beyond asking *whether a program worked*, to establishing *how it worked* and *why it worked* and *what else happened*. It is anticipated that other local, national and international simulation programs will be able to generalize the application of these findings to tailor to their own individual institutions, as we continue to advance the science of program evaluation studies within the simulation community.

## Declarations

## Acknowledgements

This project could not have been accomplished without the leadership support from eSIM Provincial Program, Alberta Health Services.

## Authors' contributions

All authors contributed to manuscript conception and design. Material, preparation, data collection and analysis

were performed by AK TC. The first draft of the manuscript was written by AK and all the authors TC, WT, TH, VG, MD commented on previous versions of the manuscript. All authors TC, WT, TH, VG, MD read and approved the final manuscript.

## Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sector.

## Availability of data and materials

None declared.

## Ethics approval and consent to participate

Ethics approval and consent to participate is not applicable as this was a non-research project. This project followed the successful completions of the “A Project Ethics Community Consensus Initiative ARECCI” screening tool (<https://arecci.albertainnovates.ca/ethics-screening-tool/>). This decision support tool identified the primary purpose of the project as quality improvement/program evaluation and that the project involves minimal risks; therefore, review by the research ethics board was not required.

## Competing interests

MD and (AK) are CEO and consultants for Healthcare Systems Simulation International Inc. which provides simulation education and consulting services. The author(s) TC, WT, TH, VG, declare(s) no conflict of interests.

## References

- Reeves S, Fletcher S, Barr H, et al. A BEME systematic review of the effects of interprofessional education: BEME Guide No. 39. *Medical Teacher*. 2015 May 6;38(7):656–668.
- Palaganas JC, Epps C, Raemer DB. A history of simulation-enhanced interprofessional education. *Journal of Interprofessional Care*. 2014 Mar 1;28(2):110–115.
- Salas E, DiazGranados D, Klein C, et al. Does team training improve team performance? A meta-analysis. *Human Factors*. 2008 Dec 1;50(6):903–933.
- O’Dea A, O’Connor P, Keogh I. A meta-analysis of the effectiveness of crew resource management training in acute care domains. *Postgraduate Medical Journal*. 2014 Nov 12; 90(1070):699–708.
- Grant RE, Goldman J, LeGrow K, MacMillan KM, Soeren M van, Kitto S. A scoping review of interprofessional education within Canadian nursing literature. *Journal of Interprofessional Care*. 2016 Sep 2;30(5):620–626.
- Pannick S, Davis R, Ashrafiyan H, et al. Effects of interdisciplinary team care interventions on general medical wards: a systematic review. *JAMA Internal Medicine*. 2015 Aug 1;175(8):1288–1298.
- Eppich W, Howard V, Vozenilek J, Curran I. Simulation-based team training in healthcare. *Simulation in Healthcare: Journal of the Society for Simulation in Healthcare*. 2011 Aug;6(Suppl):S14–S19.
- Hinde T, Gale T, Anderson I, Roberts M, Sice P. A study to assess the influence of interprofessional point of care simulation training on safety culture in the operating theatre environment of a university teaching hospital. *Journal of Interprofessional Care*. 2016 Mar 3;30(2):251–253.
- Chakraborti C, Boonyasai RT, Wright SM, Kern DE. A systematic review of teamwork training interventions in medical student and resident education. *Journal of General Internal Medicine*. 2008 Jun 1;23(6):846–853.
- Reeves S, van Schaik S. Simulation: a panacea for interprofessional learning? *Journal of Interprofessional Care*. 2012 Apr 23;26(3):167–169.
- Dobson D. Avoiding type III error in program evaluation: results from a field experiment. *Evaluation and Program Planning*. 1980 Jan 1;3(4):269–276.
- Craig P, Dieppe P, Macintyre S, et al. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ*. 2008 Sep 29;337(337):a1655.
- Scheirer MA. Planning evaluation through the program life cycle. *American Journal of Evaluation*. 2012 Jun 1;33(2):263–294.
- Chen H. *Practical program evaluation*. 2nd edition. Los Angeles: Sage Publications. 2014. 464 p.
- Van Melle E. Using a logic model to assist in the planning, implementation, and evaluation of educational programs. *Academic Medicine*. 2016 Jun 21;91(10):1464.
- Frye AW, Hemmer PA. Program evaluation models and related theories: AMEE Guide No. 67. *Medical Teacher*. 2012 May 1;34(5):e288–e299.
- Funnell S, Rogers P. Purposeful program theory: effective use of theories of change and logic models [Internet]. San Francisco, CA: Jossey-Bass. 2011 [cited 2018 May 14]. Available from: <https://www.wiley.com/en-us/Purposeful+Program+Theory%3A+Effective+Use+of+Theories+of+Change+and+Logic+Models-p-9780470478578>
- Kirkpatrick DL, Kirkpatrick JD. *Evaluating training programs: the four levels*. 3rd edition. San Francisco, CA: Berrett-Koehler Publishers. 2006. 379 p.
- Thomas PA, Kern DE, Hughes MT, Tackett SA, Chen BY. *Curriculum development for medical education: a six-step approach*. Baltimore, MD: Johns Hopkins University Press. 2022. 463 p.
- Barr H, Hammick M, Koppel I, Reeves S. Evaluating interprofessional education: two systematic reviews for health and social care. *British Educational Research Journal*. 1999 Sep 1;25(4):533–544.
- Neill MA, Wotton K. High-fidelity simulation debriefing in nursing education: a literature review. *Clinical Simulation in Nursing*. 2011 Sep 1;7(5):e161–e168.
- Fung L, Boet S, Bould MD, et al. Impact of crisis resource management simulation-based training for interprofessional and interdisciplinary teams: a systematic review. *Journal of Interprofessional Care*. 2015 Aug 28;29(5):433–444.
- Gilfoyle E, Koot DA, Annear JC, et al. Improved clinical performance and teamwork of pediatric interprofessional resuscitation teams with a simulation-based educational intervention. *Pediatric Critical Care Medicine: A Journal of the Society of Critical Care Medicine and the World Federation of Pediatric Intensive and Critical Care Societies*. 2017 Feb 1;18(2):e62–e69.
- Cook DA. How much evidence does it take? A cumulative meta-analysis of outcomes of simulation-based education. *Medical Education*. 2014 Aug 1;48(8):750–760.

25. Walsh K, Reeves S, Maloney S. Exploring issues of cost and value in professional and interprofessional education. *Journal of Interprofessional Care*. 2014 Nov 1;28(6):493–494.
26. Peterson ED, Heidarian S, Effinger S, et al. Outcomes of an interprofessional team learning and improvement project aimed at reducing post-surgical delirium in elderly patients admitted with hip fracture. *CE Measure*. 2014 Mar 28;8(1):2–7.
27. Batt AM, Tavares W, Williams B. The development of competency frameworks in healthcare professions: a scoping review. *Advances in Health Sciences Education*. 2020 Oct 1;25(4):913–987.
28. Mertens DM. Research and evaluation in education and psychology: integrating diversity with quantitative, qualitative, and mixed methods. 3rd edition. Los Angeles: Sage Publications, Inc. 2009. 552 p.
29. Madey DL. Some benefits of integrating qualitative and quantitative methods in program evaluation, with illustrations. *Educational Evaluation and Policy Analysis*. 1982 Jun 1;4(2):223–236.
30. Blamey A, Mackenzie M. Theories of change and realistic evaluation: peas in a pod or apples and oranges? *Evaluation*. 2007 Oct 1;13(4):439–455.
31. Weiss CH. Theory-based evaluation: past, present, and future. *New Directions for Evaluation*. 1997 Dec 1;1997(76):41–55.
32. Haji F, Morin MP, Parker K. Rethinking programme evaluation in health professions education: beyond ‘did it work?’ *Medical Education*. 2013 Apr 1;47(4):342–351.
33. McLaughlin J, Jordan G. Using logic models. In: Wholey KJS, Hatry HP, Newcomer KE, editors. *Handbook of practical program evaluation*. 2nd edition. Hoboken, NJ: Jossey-Bass. 2004. p. 492.
34. Kaba A, Van Melle E, Horsley T, Tavares W. Evaluating simulation programs throughout the program development life cycle. In: Chiniara G, editor. *Clinical simulation: education, operations and engineering*. 2nd edition. Academic Press. 2019. p. 890–900.
35. Patton MQ. *Utilization-focused evaluation*. 4th edition. Thousand Oaks: Sage Publications, Inc. 2008. 688 p.
36. Fetterman D, Wandersman A. Empowerment evaluation: yesterday, today, and tomorrow. *American Journal of Evaluation*. 2007 Jun 1;28(2):179–198.
37. Smutylo T. Outcome mapping: a method for tracking behavioural changes in development programs. 2005 Aug [cited 2018 May 14]. Available from: <https://cgspace.cgiar.org/handle/10568/70174>
38. Graham AC, McAleer S. An overview of realist evaluation for simulation-based education. *Advances in Simulation*. 2018 Jul 17;3(1):13.
39. Dube M, Barnes S, Cronin T, et al. Our story: building the largest geographical provincial simulation program in Canada. Accepted article in *Medical Training Magazine* [Internet]. 2019 Jul;(3). Available from: <https://medicalsimulation.training/articles/largest-geographical-provincial-simulation-proram/>
40. Malec JF, Torsher LC, Dunn WF, et al. The mayo high performance teamwork scale: reliability and validity for evaluating key crew resource management skills. *Simulation in Healthcare*. 2007 Spring;2(1):4–10.
41. Stufflebeam D, Coryn C. *Evaluation theory, models, & applications*. San Francisco, CA: Jossey-Bass. 2015.
42. Zwarenstein M, Reeves S. Knowledge translation and interprofessional collaboration: where the rubber of evidence-based care hits the road of teamwork. *Journal of Continuing Education in the Health Professions*. 2006 Dec 1;26(1):46–54.
43. Stocker M, Burmester M, Allen M. Optimisation of simulated team training through the application of learning theories: a debate for a conceptual framework. *BMC Medical Education*. 2014 Apr 3;14(1):69.
44. Zendejas B, Wang AT, Brydges R, Hamstra SJ, Cook DA. Cost: the missing outcome in simulation-based medical education research: a systematic review. *Surgery*. 2013 Feb 1;153(2):160–176.
45. Manson SM. Simplifying complexity: a review of complexity theory. *Geoforum*. 2001 Aug 1;32(3):405–414.
46. Scriven M. A unified theory approach to teacher evaluation. *Studies in Educational Evaluation*. 1995 Jan 1;21(2):111–129.