BMJ Open

Validation of a classification to identify emergency department visits suitable for subacute and virtual care models: a randomised single-blinded agreement study protocol

Ryan P Strum ,[1] Shawn Mondoux,[2,3] Fabrice Mowbray ,[1] Andrew Worster,[1,2] Lauren E Griffith ,[1,4] Walter Tavares,[5] Paul Miller,[6,7] Erich Hanel,[6] Komal Aryal,[1] Ravi Sivakumaran,[8] Andrew P Costa [1,9]

To cite: Strum RP, Mondoux S, Mowbray F, *et al*. Validation of a classification to identify emergency department visits suitable for subacute and virtual care models: a randomised single-blinded agreement study protocol. *BMJ Open* 2022;**12**:e068488. doi:10.1136/bmjopen-2022-068488

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2022-068488).

Received 19 September 2022
Accepted 05 December 2022

For numbered affiliations see end of article.

Correspondence to
Mr Ryan P Strum;
strumr@mcmaster.ca

ABSTRACT

Introduction Redirecting suitable patients from the emergency department (ED) to alternative subacute settings may assist in reducing ED overcrowding while delivering equivalent care. The Emergency Department Avoidance Classification (EDAC) was constructed to retrospectively classify ED visits that may have been suitable for safe management in a subacute or virtual clinical setting. The EDAC has established face and content validity but has not been tested against a reference standard as a criterion.

Objectives Our primary objective is to examine the agreement between the EDAC and ED physician judgements in retrospectively identifying ED visits suitable for subacute care management. Our secondary objective is to assess the validity of ED physicians' judgement as a criterion standard. Our tertiary objective is to examine how the ED physician's perception of a virtual ED care alternative correlates with the EDAC.

Methods and analysis A randomised single-centre, single-blinded agreement study. We will randomly select ED charts between 1 January and 31 December 2019 from an academic hospital in Hamilton, Canada. ED charts will be randomly assigned to participating ED physicians who will evaluate if this ED visit could have been managed appropriately and safely in a subacute and/or virtual model of care. Each chart will be reviewed by two physicians independently. We compute our needed sample size to be 79 charts. We will use kappa statistics to measure inter-rater agreement. A repeated measures regression model of physician ratings will provide variance estimates that we will use to assess the intraclass correlation of ED physician ratings and the EDAC.

Ethics and dissemination This study has been approved by the Hamilton Integrated Research Ethics Board (2022-14625). If validated, the EDAC may provide an ED-based classification to identify potentially avoidable ED visits, monitor ED visit trends, and proactively delineate those best suited for subacute or virtual care models.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ To the best of our knowledge, this is the first study to validate an epidemiological classification against physician determination as a reference standard.
⇒ Emergency department physicians will be used as a criterion reference in absence of a 'gold standard' for validation and examined for inter-rater reliability.
⇒ Single-centre study at an academic hospital.
⇒ Validation of this classification could permit epidemiologists to accurately identify retrospective emergency department visits that were more suitable for subacute or virtual care models.
⇒ Physicians are not blinded to personal health identifying information within the patient record, which may incorporate some implicit bias.

INTRODUCTION

Ontario emergency departments (EDs) are challenged with providing timely medical care despite steady increases in utilisation and overcrowding.[1–3] EDs are commonly the first point of contact to engage with the healthcare system, independent of need or ability to seek non-emergency alternatives.[4 5] Non-emergent visits constitute the majority of all ED encounters, and play an important role in determining measures of performance and quality of care (ie, time to physician assessment, patient satisfaction, overall department workload).[3 6–9] In Ontario, ED utilisation by patients with non-emergent conditions has doubled population growth (13.4% vs 6.2%) in the past decade.[6 10] With Ontario's continued population growth, the demand for ED healthcare may continue to increase, further challenging departments to manage already overburdened workloads.[11–13] Challenges of ED overcrowding extend beyond Canada; international research reports similar increases in patient volumes and longer admission times.[14–16]

The Emergency Department Avoidance Classification (EDAC) was constructed using a multistage, multicentred, consensus process of leading emergency and primary physicians in Ontario, Canada.[3 17] The EDAC aims to retrospectively identify ED visits that could have been appropriately managed in a subacute clinical setting.[3 17] This classification addresses gaps in previously developed models by uniquely identifying informative patient features (beyond acuity or diagnostic category alone) and including criterion validity examination, a core component for generalisability absent from numerous ED classifications.[18–20] Specifically, the EDAC (1) identifies patients who could have sought care in a subacute centre, (2) determines which subacute setting could be appropriate (urgent care and/or general practice), (3) has high specificity to avoid including patients who require ED care, and (4) has established face and construct validity through a consensus process.[10] Limitations of the EDAC's utility persist without understanding its agreement with a reference standard, and constitute a challenge in the absence of a gold-standard for comparison.[21 22] If the EDAC can be validated against a criterion standard, such as ED physician evaluations of potentially avoidable emergency visits, the EDAC could be used to support proactive decision-making about health resource allocation.

Our primary objective is to examine agreement between the EDAC and ED physicians in retrospectively identifying ED visits that could have been redirected to subacute primary care. Our secondary objective is to assess the validity of ED physicians as a criterion standard by examining inter-rater agreement among ED physicians. Our tertiary objective is to examine how the ED physician's perception of a virtual ED care alternative correlates with EDAC.

METHODS/DESIGN

Study design and setting

We will conduct a single-centre, single-blinded, randomised agreement study. We will recruit ED physicians from a single academic hospital to review randomly selected electronic ED patient charts. ED charts will be categorised into three study blocks based on the EDAC criteria. Physicians will rate whether the ED visit could have been safely managed in a subacute care setting and/or via a virtual care visit, while blinded to the study block the ED chart belongs. ED charts will be identified as potentially avoidable in this study using the inclusion criteria of the EDAC, and agreement between ED physician ratings and the EDAC will be analysed. This study will commence in September 2022.

RECRUITMENT OF PHYSICIAN RATERS

Physician eligibility and recruitment

Inclusion in our study will require ED physicians to meet the following criteria: (1) currently clinically practising and (2) holding a staff emergency physician position at the academic hospital. Eligible study physicians will be recruited by the study's principal investigator (PI). An information letter and consent form will be provided, and all will be given the opportunity to review and ask questions prior to enrolling. Upon acceptance, each physician will sign and return a study consent form. A participant demographic questionnaire will be distributed to all physicians to report aggregate characteristic information in the final manuscript (ie, sex, years of practice, primary practice setting, college designation). We estimate a minimum of 10 physicians should be recruited to participate in this study. Participating physicians will receive financial compensation equivalent to their hourly compensation for completing the study's tasks.

ED CHARTS

ED patient chart eligibility

Patient charts will be eligible for inclusion in the study if:

1. All patient fields that specify the EDAC are inputted (patient age, triage Canadian Triage and Acuity Scale (CTAS), physician main intervention, specialist consult completed, ED visit outcome).
2. Patients did not leave against medical advice.
3. The visit occurred between 1 January 2019 and 31 December 2019. This time frame represents the most recent 12-month period prior to the COVID-19 pandemic when ED utilisation changed.[23]

Chart selection and randomisation

We will provide the academic hospital's Health Information Management Department with the criteria needed to specify an EDAC visit, shown in online supplemental appendix 1. When eligible ED charts have been identified, study relevant data elements of the charts will be extracted by the academic hospital and provided to the study investigators using an encrypted file. Data from ED charts will include the medical report number (MRN), mode of arrival, month and time of visit, sex, main diagnosis, previously attended the ED within 30 days and all criteria of the EDAC (age, triage acuity, specialist consult conducted in the ED, ED visit disposition, main physician intervention). Due to personal health identification legislation, only these features of the ED charts can be extracted, not the entire ED chart.

We will group all eligible ED charts into one of three study blocks. Assignment to a study block will be dependent on an ED chart's alignment with the EDAC, categorised as: (1) all EDAC criteria met, (2) all EDAC criteria met except with an urgent triage acuity (CTAS 3) and (3) not all EDAC criteria met. Block 2 will be used to assess a plausible middle level of the EDAC, where some ED visits could have been suitable for subacute care models but are not recognised by the EDAC. We hypothesise that classifying all urgently triaged patients as ineligible for subacute care models may limit the range of the classification to assess ED visits that are likely to be suitable

for subacute models. Given the EDAC was constructed using a conservative and highly specific approach to identify patients retrospectively, a middle-level classification concerning EDAC visits but with an urgent acuity is plausible and warrants investigation.

We will randomly select an equal number of ED charts from each block; these will be used as the study charts. We will assign all ED charts used in the study a unique study ID number; the study key matching MRNs with their corresponding study ID will be securely stored with only the PI.

Sample size estimates

We estimate this study will require a minimum of 79 ED charts to be reviewed by two independent physicians to draw meaningful conclusions using a 95% CI, though we aim to complete more than this minimum to increase our sample size and statistical power.

Based on 80% (±5%) physician agreement for subacute care suitability and chance agreement estimated at 25%, we expect a Cohen's kappa statistic of 0.8. This kappa constitutes a very high level of agreement to infer study conclusions that are beyond the probability of chance, though a kappa of 0.6–0.8 will be acceptable, which indicates substantial agreement.[24] We estimate a minimum sample size of 79 ED charts are required given a minimum acceptable kappa of 0.6, our anticipated kappa of 0.8, proportion of the outcome is 0.5 (binary outcome, yes/no of redirection suitability), a 0.05 alpha (two tailed), using a power of 70%.[25] We aspire to complete 126 ED chart reviews given feasibility, which would increase our study's power to an optimal 80%.

STUDY PROCESS

Data collection and handling

We will provide all participating physicians with a password-protected Excel file containing all assigned ED chart MRNs to evaluate in the academic hospital electronic ED database using secure electronic communication. Physicians will be requested to review the ED chart and complete a questionnaire within the Excel file related to this specific ED chart. Following completion, the Excel file will be returned to the PI using secure electronic communication.

We will randomly distribute ED charts evenly among the participating physicians. Each individual ED chart must be rated in duplicate by independent physicians, and no two physicians can be paired to rate an ED chart more than once in each block. Given each physician has an equal possibility of being paired with another participant to rate one ED chart, but cannot occur twice, we estimate each physician will rate 13 ED charts. We will encourage physicians to complete more ED charts than their minimum, and make every effort to recruit more than 12 physicians. We estimate a single ED chart may take 2–5 min to complete, culminating in 25–60 min of study contributions from each participant.

Outcome measures

Table 1 shows the data collection questionnaire for the outcome measures. First, physicians will be requested to judge whether an ED visit could have been appropriately and safely managed in a subacute and/or virtual care model. Physicians will be permitted to select multiple care settings. Physicians will be given a 5-point Likert scale to rate their confidence in their decision, ranging from not confident (1) to very confident (5).[26] Descriptions, definitions, staffing, diagnostic imaging and care services (ie, laboratory, pharmaceutical) will be provided for each centre to align understanding among physicians prior to ratings. Second, physicians will be requested to judge which of the subacute or virtual care models is the best care centre that could have managed this ED visit. A second 5-point Likert scale will be used to rate the confidence in their selection. All study questions were reviewed independently by three ED physicians (none eligible for study inclusion); no interpretation issues were identified.

Blinding

Physicians will be blinded to the knowledge of the EDAC criteria or components, and the study block to which

Table 1 Study questions to be completed by each physician for each ED chart review

Study question	Possible answers	Confidence score*
From a retrospective position, which of the following care settings **could** have appropriately and safely managed this ED visit? (*Select all that apply*)	► Urgent care centre ► Family medicine centre (with their family physician) ► Family medicine walk-in care centre ► Virtual care with an emergency physician ► Virtual care with a family physician ► Only the ED	1. Not confident 2. Slightly confident 3. Moderately confident 4. Very confident 5. Extremely confident
From a retrospective position, which of the following is the **best** care setting to appropriately and safely manage this ED visit? (*One selection only*)	► Urgent care centre ► Family medicine centre (with their family physician) ► Family medicine walk-in care centre ► Virtual care with an emergency physician ► Virtual care with a family physician ► ED	1. Not confident 2. Slightly confident 3. Moderately confident 4. Very confident 5. Extremely confident

*A confidence score will be requested for each care setting selection.
ED, emergency department.

an ED chart belongs. The ED chart format, information and presentation will not be altered in any way for the study. ED charts will be shown in the format the physicians expect when completing the charts at the time of a patient's ED visit.

Data handling

All Excel questionnaires and related study documents will be stored and only accessible to the PI. After all questionnaires have been completed and returned to the PI, all questionnaires will be combined for analytical purposes. At this time, all MRNs will be removed and replaced with their study ID numbers to minimise any risk of MRN reidentification. All study files returned to the PI will be stored in a locked computer as an encrypted file, only accessible to the PI until deidentification occurs (MRNs substituted with study IDs). Following the completion of the study, the MRN cross-walk will be permanently deleted.

Study steps

Figure 1 shows a summary of all study steps to assemble the study data used to compute the analysis.

Statistical plan

Participant demographic and ED chart descriptive statistics will be reported using general measures of central tendency and frequency.

Physician ratings will be tested as a criterion standard by computing the inter-rater quadratically weighted kappa agreement of all ED charts. Criterion will be established if a kappa of 0.6 is achieved on ED charts where the EDAC criteria were met. To examine the agreement of physician ratings with the EDAC, we will use a repeated measures regression model to calculate a variance estimate. This variance estimate will inform the final computation of an intraclass correlation of the EDAC in each block.

To examine if a plausible middle level could be incorporated into the EDAC, physician ratings of the EDAC visits that have an urgent triage acuity (block 2) will be compared with the EDAC (block 1) using repeated measures regression model and intraclass correlation analysis, and compared with the first intraclass correlation model.

ED physician perceptions of virtual care ED alternatives will be measured for inter-rater kappa agreement of all ED charts. A kappa regression will be computed with physician answers and their confidence scores.

Patient and public involvement

Potential implications of this study's findings were discussed with Ontario ED clinicians and epidemiologists to gauge their satisfaction with this study's methodology and outcomes. This protocol received input from EDs across Ontario, where frontline ED clinicians were asked focused questions on their perspectives of alternative healthcare centres for patients who require primary care. Additionally, clinicians were asked to speculate on patients' perceptions of alternative healthcare centres and whether they would have been more appropriate than the ED. All input helped to modify the study design.

ETHICS AND DISSEMINATION

Ethics approval

This study protocol has been approved by the Hamilton Integrated Research Ethics Board (review reference 2022-14625-GRA).

Risk to participating physicians

No known risks to physicians are anticipated as a result of study participation. Participating physicians will be asked to review and rate ED patient charts, a task that is within their scope of clinical practice. Psychological distress is unlikely, though withdrawal is permitted at any time for any reason (ie, burden of time). All physicians will be informed of their rights and can terminate their participation without any consequence, with their data withdrawn upon request.

Confidentiality

All participating physicians in this study will have their anonymity maintained by the researchers. All documents will be stored securely and are only assessable by the investigators.

Results dissemination

The results of this study will be made public through peer-reviewed publication, study registries, conference publications and thesis manuscripts. Communication will be sent to relevant stakeholders with the study's results for distribution in reports and newsletters.

DISCUSSION

Assessing criterion validity for the EDAC against an independent standard could establish EDAC as an accurate identifier of retrospective ED visits potentially manageable in non-ED centres.[27] Additionally, with appropriate validation, this algorithm could be used to proactively delineate patients better suited for subacute clinical settings. Leveraging physician determination as a criterion is an appropriate measure to test experimental validity in this circumstance, as there is no gold-standard classification to establish appropriateness or has not been experimentally validated.[28] Given validation testing tends to perform better on data used to construct a classification or scale, examining EDAC with a panel of physicians who are unaware of the EDAC's components can provide a more unbiased estimate of EDAC's true accuracy and performance.[22 29 30] This also mimics clinical practice, where emergency physicians are not given this information during clinical decision-making. If the performance of EDAC is validated against a criterion standard using complete ED records, this could be a strong indicator the EDAC is reproducible and generalisable to new settings as a criterion to describe avoidable ED visits.[21]

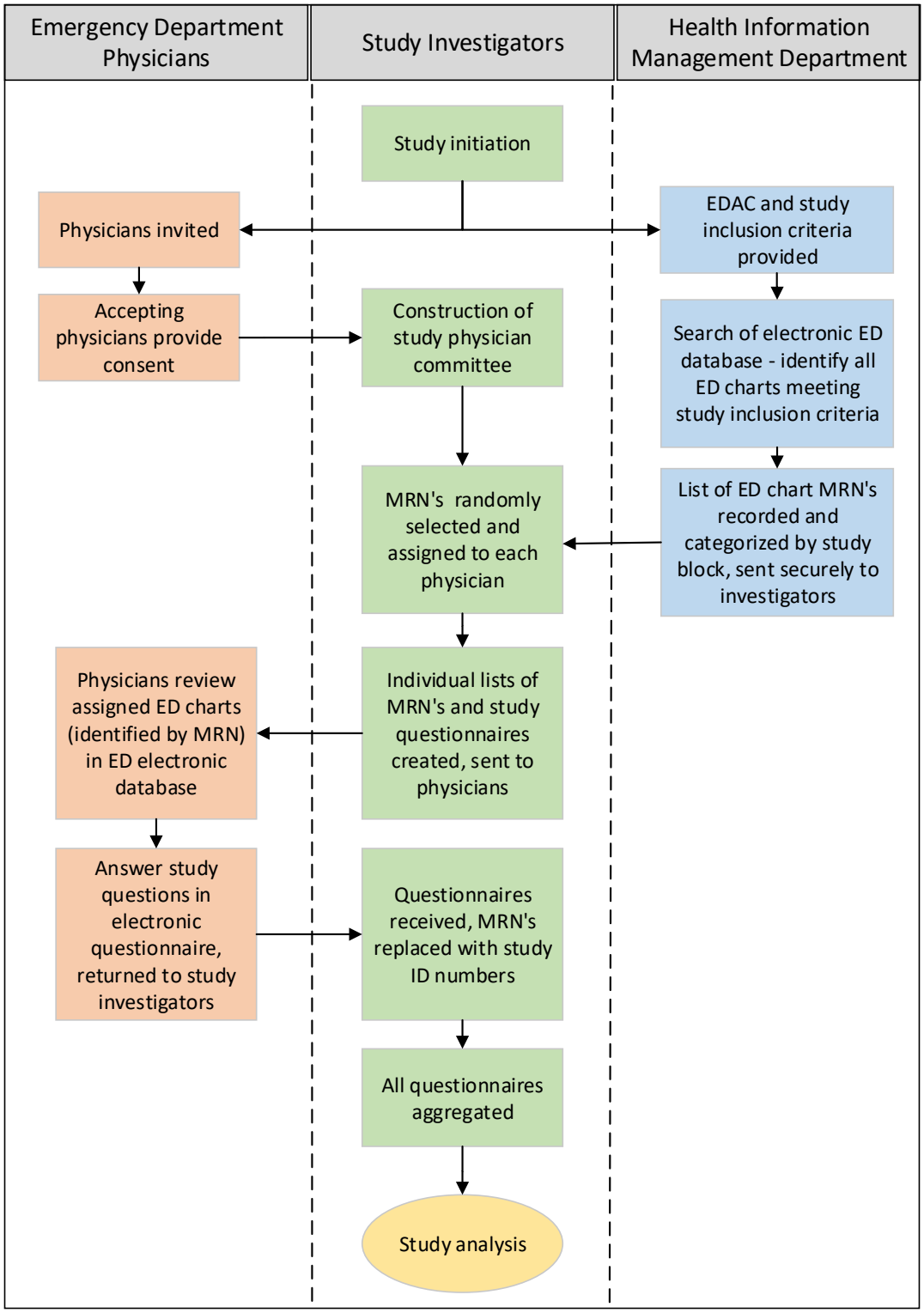

Figure 1 Steps to collect study data from recruited academic hospital physicians. ED, emergency department; EDAC, Emergency Department Avoidance Classification; MRNs, medical report numbers.

If validated, the EDAC could be established as an authentic benchmark to inform numerous epidemiological fields, including prehospital, paramedicine, emergency, virtual and primary care. In paramedicine, the EDAC could support prospective research in alternative destination models of care and aim to improve the safety of paramedic clinical judgements regarding transport decision-making. In emergency medicine, the EDAC could be used to identify patient cohorts for further research to understand the rationale and healthcare-seeking behaviours of patients who sought scarce ED resources when non-ED care alternatives were as appropriate.

Additionally, the EDAC could be leveraged to develop a prospective tool for ED triage to direct care resources, and compared with standard practices in future research. The EDAC could inform virtual care applications to classify which ED visits may have been avoidable with support of a virtual care visit with a physician and support virtual care as a model for further research. Epidemiologically, the EDAC could help to provide an ED-based resource to explain ED appropriateness and monitor trends in ED resources over time. Lastly, an output of this study could construct an ordinal scale of the EDAC, instead of a binary classification, which could aid in describing likely avoidable ED visits that presently not recognised in the EDAC criteria.

To our knowledge, this is the first study to examine an ED avoidance classification for criterion validity. This research study is methodologically rigorous and sufficiently powered to evaluate the experimental validity of EDAC. This research has the potential to contribute evidence to support new care models of care development that can identify patients who sought ED care, when subacute alternatives were likely more appropriate.

Several outcomes are plausible from this study that are informative of validation of ED physicians as a criterion and the EDAC as a validated classification. If ED physicians' perceptions of ED charts are in agreement, ED physicians can be established as an appropriate criterion standard. If a criterion is established and demonstrates agreement with the EDAC, the EDAC can demonstrate criterion validity. In circumstances where ED physician perceptions do not agree on ratings of ED charts, an appropriate criterion standard cannot be created. In the absence of establishing a criterion standard but the perceptions are not significantly different from between rater agreement with the EDAC, the EDAC could be determined to be no different than ED physician perspectives, thus cannot be validated.

This study conveys a risk the EDAC may not be validated by an external group of physicians for retrospective epidemiological purposes. However, we contend the high internal validity process used to construct the EDAC has assembled favourably conservative inclusion criteria that carry the highest potential for validation. We anticipate the EDAC to establish criterion validity in this study.

Study timeline

We expect to conduct this study between September and November 2022, with probable publication by January 2023. The academic hospital's ED physicians will be recruited beginning in September 2022 and may occur throughout the study period. Assignment of ED charts to participants will occur from October to November 2022, with data collection occurring upon completion of each ED chart review. Analyses of results will occur at the conclusion of the study, anticipated for December 2022.

Author affiliations
[1]Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada
[2]Medicine, Division of Emergency Medicine, McMaster University, Hamilton, Ontario, Canada
[3]Institute for Health Policy, Management and Evaluation, University of Toronto, Toronto, Ontario, Canada
[4]McMaster Institute for Research and Aging, McMaster University, Hamilton, Ontario, Canada
[5]The Wilson Centre, University of Toronto, Toronto, Ontario, Canada
[6]Medicine, McMaster University, Hamilton, Ontario, Canada
[7]Centre for Paramedic Education and Research, Hamilton Health Sciences, Hamilton, Ontario, Canada
[8]Health Information Management, Saint Joseph's Healthcare Hamilton, Hamilton, Ontario, Canada
[9]Institute for Clinical Evaluative Sciences, Toronto, Ontario, Canada

Twitter Ryan P Strum @Ryan_Strum, Fabrice Mowbray @modeRNepi and Lauren E Griffith @LaurenGriff1

Contributors RPS, SM and APC led the conceptualisation of the study objective and methodology. RPS designed the study, drafted and revised the manuscript. RPS and LEG computed the estimated sample size and statistical plan. SM, FM, AW, LEG, WT, PM, EH, KA, RS and APC made critical revisions to the design, methodology and manuscript, and agreed to be accountable for all aspects of this manuscript.

Funding This work was substantially supported by Schlegel Chair in Clinical Epidemiology in Ageing at McMaster University (award/grant number is not applicable).

Competing interests None declared.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iDs
Ryan P Strum http://orcid.org/0000-0003-1902-4734
Fabrice Mowbray http://orcid.org/0000-0002-8844-5850
Lauren E Griffith http://orcid.org/0000-0002-2794-9692
Andrew P Costa http://orcid.org/0000-0001-9212-5641

REFERENCES

1 Affleck A, Parks P, Drummond A, *et al*. Emergency department overcrowding and access block. *CJEM* 2013;15:359–70.
2 Bond K, Ospina MB, Blitz S, *et al*. Frequency, determinants and impact of overcrowding in emergency departments in Canada: a national survey. *Healthc Q* 2007;10:32–40.
3 Strum RP, Tavares W, Worster A, *et al*. Development of the PriCARE classification for potentially preventable emergency department visits by ambulance: a RAND/UCLA modified Delphi study protocol. *BMJ Open* 2021;11:e045351.
4 Martin D, Miller AP, Quesnel-Vallée A, *et al*. Canada's universal health-care system: achieving its potential. *Lancet* 2018;391:1718–35.

5 National ambulatory care reporting system metadata (NACRS) | CIHI. Available: https://www.cihi.ca/en/national-ambulatory-care-reporting-system-metadata-nacrs [Accessed 14 Nov 2020].
6 Strum RP, Mowbray FI, Worster A, *et al*. Examining the association between paramedic transport to the emergency department and hospital admission: a population-based cohort study. *BMC Emerg Med* 2021;21:1–9.
7 Forster AJ, Stiell I, Wells G, *et al*. The effect of hospital occupancy on emergency department length of stay and patient disposition. *Acad Emerg Med* 2003;10:127–33.
8 Derlet RW, Richards JR. Ten solutions for emergency department crowding. *West J Emerg Med* 2008;9:24–7.
9 Vieth TL, Rhodes KV. The effect of crowding on access and quality in an academic ED. *Am J Emerg Med* 2006;24:787–94.
10 Strum RP, Tavares W, Worster A, *et al*. Identifying patient characteristics associated with potentially redirectable paramedic transported emergency department visits in Ontario, Canada: a population-based cohort study. *BMJ Open* 2021;11:e054625.
11 Health Quality Ontario. Under pressure: emergency department performance in Ontario, 2016. Available: http://www.hqontario.ca/Portals/0/documents/system-performance/under-pressure-report-en.pdf
12 Ministry of Finance. Ontario population projections. government of Ontario, 2020. Available: https://www.ontario.ca/page/ontario-population-projections [Accessed 3 Apr 2022].
13 Strum RP, Drennan IR, Mowbray FI. Increased demand for paramedic transports to the emergency department in Ontario, Canada: a population-level descriptive study from 2010 to 2019. *Can J Emerg Med* 2022.
14 Iacobucci G. Overcrowding and long delays in A&E caused over 4000 deaths last year in England, analysis shows. *BMJ* 2021;375:n2835.
15 Panovska-Griffiths J, Ross J, Elkhodair S, *et al*. Exploring overcrowding trends in an inner City emergence department in the UK before and during COVID-19 epidemic. *BMC Emerg Med* 2021;21:43.
16 Janke AT, Melnick ER, Venkatesh AK. Hospital occupancy and emergency department boarding during the COVID-19 pandemic. *JAMA Netw Open* 2022;5:e2233964.
17 Strum RP, Tavares W, Worster A, *et al*. Emergency department interventions that could be conducted in subacute care settings for patients with nonemergent conditions transported by paramedics: a modified Delphi study. *CMAJ Open* 2022;10:E1–7.
18 Purdy S, Griffin T, Salisbury C, *et al*. Ambulatory care sensitive conditions: terminology and disease coding need to be more specific to aid policy makers and clinicians. *Public Health* 2009;123:169–73.
19 Nagree Y, Camarda VJ, Fatovich DM, *et al*. Quantifying the proportion of general practice and low-acuity patients in the emergency department. *Med J Aust* 2013;198:612–5.
20 Gruneir A, Bell CM, Bronskill SE, *et al*. Frequency and pattern of emergency department visits by long-term care residents--a population-based study. *J Am Geriatr Soc* 2010;58:510–7.
21 Ramspek CL, Jager KJ, Dekker FW, *et al*. External validation of prognostic models: what, why, how, when and where? *Clin Kidney J* 2021;14:49–58.
22 Bleeker SE, Moll HA, Steyerberg EW, *et al*. External validation is necessary in prediction research: a clinical example. *J Clin Epidemiol* 2003;56:826–32.
23 Ferron R, Agarwal G, Cooper R, *et al*. The effect of COVID-19 on emergency medical service call volumes and patient acuity: a cross-sectional study in Niagara, Ontario. *BMC Emerg Med* 2021;21:39.
24 McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med* 2012;22:276–82.
25 Donner A, Eliasziw M. A goodness-of-fit approach to inference procedures for the kappa statistic: confidence interval construction, significance-testing and sample size estimation. *Stat Med* 1992;11:1511–9.
26 Sullivan GM, Artino AR. Analyzing and interpreting data from likert-type scales. *J Grad Med Educ* 2013;5:541–2.
27 Bellamy N. 2 - Principles of clinical outcome assessment. In: Hochberg MC, Silman AJ, Smolen JS, *et al*, eds. *Rheumatology*. Sixth Edition. Mosby, 2015: 9–19.
28 Geffen LN, Kelly G, Morris JN, *et al*. "Establishing the criterion validity of the interRAI Check-Up Self-Report instrument". *BMC Geriatr* 2020;20:260.
29 Harrell FE, Lee KL, Mark DB. *Multivariable Prognostic Models: Issues in Developing Models, Evaluating Assumptions and Adequacy, and Measuring and Reducing Errors*. Vol 15. Statistical Methods. Wiley Library. Available: https://doi.org/10.1002/(SICI)1097-0258(19960229)15:4<361::AID-SIM168>3.0.CO;2-4 [Accessed 29 Dec 2021].
30 Moons KG, van Es GA, Michel BC, *et al*. Redundancy of single diagnostic test evaluation. *Epidemiology* 1999;10:276–81.