


BMJ Open Methodology paper for the General Medicine Inpatient Initiative Medical Education Database (GEMINI MedED): a retrospective cohort study of internal medicine resident case-mix, clinical care and patient outcomes

Andrew CL Lam ¹, Brandon Tang,² Anushka Lalwani,³ Amol A Verma,^{1,3,4} Brian M Wong,^{1,5} Fahad Razak,^{1,3,4} Shiphra Ginsburg^{6,7}

To cite: Lam ACL, Tang B, Lalwani A, *et al.* Methodology paper for the General Medicine Inpatient Initiative Medical Education Database (GEMINI MedED): a retrospective cohort study of internal medicine resident case-mix, clinical care and patient outcomes. *BMJ Open* 2022;**12**:e062264. doi:10.1136/bmjopen-2022-062264

► Prepublication history for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2022-062264>).

ACL and BT contributed equally.

Received 21 February 2022
Accepted 22 August 2022



© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Dr Shiphra Ginsburg;
shiphra.ginsburg@utoronto.ca

ABSTRACT

Introduction Unwarranted variation in patient care among physicians is associated with negative patient outcomes and increased healthcare costs. Care variation likely also exists for resident physicians. Despite the global movement towards outcomes-based and competency-based medical education, current assessment strategies in residency do not routinely incorporate clinical outcomes. The widespread use of electronic health records (EHRs) may enable the implementation of in-training assessments that incorporate clinical care and patient outcomes.

Methods and analysis The General Medicine Inpatient Initiative Medical Education Database (GEMINI MedED) is a retrospective cohort study of senior residents (postgraduate year 2/3) enrolled in the University of Toronto Internal Medicine (IM) programme between 1 April 2010 and 31 December 2020. This study focuses on senior IM residents and patients they admit overnight to four academic hospitals. Senior IM residents are responsible for overseeing all overnight admissions; thus, care processes and outcomes for these clinical encounters can be at least partially attributed to the care they provide. Call schedules from each hospital, which list the date, location and senior resident on-call, will be used to link senior residents to EHR data of patients admitted during their on-call shifts. Patient data will be derived from the GEMINI database, which contains administrative (eg, demographic and disposition) and clinical data (eg, laboratory and radiological investigation results) for patients admitted to IM at the four academic hospitals. Overall, this study will examine three domains of resident practice: (1) case-mix variation across residents, hospitals and academic year, (2) resident-sensitive quality measures (EHR-derived metrics that are partially attributable to resident care) and (3) variations in patient outcomes across residents and factors that contribute to such variation.

Ethics and dissemination GEMINI MedED was approved by the University of Toronto Ethics Board (RIS#39339). Results from this study will be presented in academic conferences and peer-reviewed journals.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- ⇒ Multicentre retrospective cohort study of internal medicine (IM) resident clinical practice.
- ⇒ Clinical practice includes case-mix, clinical care in the form of resident-sensitive quality measures, and patient outcomes.
- ⇒ Study population includes >900 unique residents, data from 1 180 000 unique IM admissions and clinical data derived from electronic records of 4 hospitals across 10 academic years.
- ⇒ Limitations include difficulty attributing care and outcomes entirely to residents and reliance on data from a single residency programme.

INTRODUCTION

Unwarranted variations in patient care have been documented for virtually every major medical condition across multiple disciplines and are associated with negative patient outcomes and increased healthcare costs.^{1–5} For instance, in a Toronto-based study, our team found a 62% relative difference in inpatient mortality between the lowest and highest quartiles of internists in seven academic hospitals; in the same study, significant variations were also found in length-of-stay, use of diagnostic imaging and readmission rates.³ Other examples of physician-level practice variations and patient outcomes have been documented in general surgery, emergency medicine (EM) and primary care.^{2 4 6}

One source of physician-level variation can be traced to residency, as suggested by divergent patient outcomes based on residency location.^{7–11} Clinical outcomes, ranging from surgical complications to healthcare expenditure, varied significantly between physicians



from different residency programmes, with variations persisting years after graduation.⁷⁹ However, most studies examining the link between residency location and patient care only explored outcomes after residents transitioned into independent practice. There are few studies of care quality and patient outcomes linked to physicians during residency.^{10 12} When it comes to measuring resident practice, assessments emphasise educational outcomes (eg, preceptor ratings and test scores) rather than clinical outcomes and do not incorporate the most important part of medical practice—the patient.¹³ Measuring resident clinical exposure, care quality and patient outcomes in real time may identify unwarranted variation in practice and provide a valuable source of feedback to inform ongoing learning activities and improve educational outcomes.

Measuring exposure and assessing competency in residency

With increased calls to incorporate clinical activities directly into resident assessments, long gone are residency programmes that use standardised test scores or clinical time as the sole means of assessment.¹⁴ But despite the movement towards workplace-based assessments (WBAs) focused on clinical activities, there still remain deficiencies within residency programmes as it pertains to tracking resident case-mix, clinical care and patient outcomes.

In terms of case mix, studies linking resident clinical exposure have found a moderate association between clinical volumes and diversity with in-training examination scores.^{15 16} However, outside of research studies, most programmes lack systematic and accurate methods of tracking case mix.^{17 18} Past attempts using resident-driven case logs, manual extraction from health records and prescription profiles were labour intensive and could not comprehensively capture case mix.^{17 19} The most popular method, resident-driven case logs, suffer from substantial recall and selection bias as well as coding error rates ranging from 19% to 47%.^{20–22} Although it would be erroneous to attribute the quality of residency education entirely to clinical exposure, case mix is a vital measure that may guide residents to seek out additional clinical experiences.

In terms of resident assessment, there is a global shift in residency programmes towards competency-based medical education (CBME).^{23–25} A guiding principle behind this shift is to increase the number of clinical data points available to both residents and educators to quantify performance. For instance, instead of a single summative end-of-rotation evaluation, competency-based programmes opt for multiple formative/summative WBAs throughout each rotation. These preceptor-driven WBAs are an invaluable source of actionable, specific and timely feedback for residents, and are more accurate markers of on-the-job performance for educators.²⁶ However, WBAs, like most subjective evaluations, have their limitations, particularly in terms of objectivity and comprehensiveness.^{27–29} In terms of objectivity, preceptor-driven

assessments often suffer from variable frames of reference, differing rating scales and confounding by external factors.³⁰ In terms of comprehensiveness, physicians make anywhere from 7 to 16 clinical decisions per patient encounter, making it unreasonable for preceptor-driven feedback to be provided on all of these decisions.³¹ Taken together, these limitations raise questions on whether an objective clinical data source is required for resident assessment, in addition to preceptor-driven WBAs.

Bridging educational and clinical outcomes

A promising approach to bridge the gap between educational and clinical outcomes, and supplement WBAs in CBME, are resident-sensitive quality measures (RSQMs).^{32 33} RSQMs are defined as measures that are: (1) directly meaningful to patient care in the clinical environment and (2) largely attributable to actions of an individual resident. Unlike WBAs, which may be subjectively resident driven or preceptor driven, RSQMs reflect clinical care provided by residents and are often extracted directly from health records. Previously published RSQMs were developed through consensus methods and emphasised clinical actions that may be attributed to resident care. RSQMs developed and implemented in paediatric EM demonstrated significant variation in the quality of asthma, bronchiolitis and head injury care provided between residents.^{33–35} These studies demonstrated associations between RSQM scores and traditional entrustment decisions, suggesting RSQMs could be employed in summative resident assessments.^{35 36} The same studies suggested RSQMs may also informally highlight to residents the elements of patient care that directly improve care quality, signalling the role RSQMs may also play in formative feedback and self-reflection.

Leveraging electronic clinical data in medical education

Despite the initial promise of these RSQM studies, one major hurdle was their reliance on large quantities of clinical data, often requiring manual chart review that limits scalability. The time is ripe to leverage large data sets in medical education; the infrastructure to build such databases is widely available as 96% of American hospitals already use electronic health records (EHRs).^{37 38} Furthermore, natural language processing technologies have shown promise in automating data extraction of clinical notes, ranging from radiology reports to discharge summaries.³⁹ In the EHRs of academic health centres, patients are assigned to staff physicians such that data-driven metrics are already being generated to support audit and feedback, physician report card initiatives, and broader quality improvement efforts.^{17 40 41} However, the transient nature of learners in our health system creates unique challenges when trying to link patient data to resident care, especially in the inpatient setting. As such, there are no known examples where large administrative databases have linked patient-level quality-of-care and clinical outcomes to individual residents in the acute care environment.

General Medicine Inpatient Initiative Medical Education Database

The General Medicine Inpatient Initiative Medical Education Database (GEMINI MedED) links patient care processes and outcomes to internal medicine (IM) residents that have spent clinical time at any of four major urban Toronto hospitals with the aim of studying resident clinical practice. We define clinical practice by three domains that can be measured using routinely collected data: case-mix, clinical care (in the form of RSQMs) and patient outcomes. GEMINI MedED will leverage data from the GEMINI dataset, which encompasses EHR data from 1 180 000 unique IM admissions, including admissions to University of Toronto (UofT)-affiliated hospitals.⁴² GEMINI MedED will study the IM resident clinical practice in a retrospective study that will span four hospital sites and 10 academic years.

IM is an ideal ecosystem to examine learner exposure and competence given its relevance to a large number of trainees (IM programmes have the most residency positions in the USA and second most in Canada of all specialties), diversity of case mix (~40% of hospital admissions are IM patients) and the senior IM resident's relative independence with respect to clinical decision-making during the overnight on-call period.^{43–45} This study will focus on the senior IM resident on overnight IM call shifts. At all UofT-affiliated hospitals, one senior resident is responsible for triaging and supervising all overnight admissions to the IM service; there are no in-hospital IM staff physicians overnight and the majority of admitted cases are reviewed by the IM staff physician the next morning. As such, clinical care and patient outcomes overnight are at least partially attributable to the care of a single senior IM resident. GEMINI MedED is a proof-of-concept study, tying a large clinical informatic database with an IM residency programme that spans multiple academic hospitals. Once GEMINI MedED is complete, we aim to demonstrate that potentially useful data in the form of measuring case-mix exposure, resident clinical care (in the form of RSQMs) and patient outcomes may be extracted. GEMINI MedED has relevance to three

stakeholder groups: residents, educators/researchers and patients. For residents, this database may be a source of objective and comprehensive feedback based on patient care and outcomes partly attributable to an individual resident's practice. This feedback may supplement the current preceptor-driven WBAs with minimal increase in resident and preceptor workloads. For educators and researchers, in addition to residents, this system can track resident clinical exposure and practice patterns across training sites and academic years without reliance on resident self-report or manual chart extraction. This can help plan future rotations and curriculum design. Finally, for patients and the health system, the goal is to ensure residents are providing high-quality care both during and after residency. Given both the trust and financial investment the public places into training residents, this study aims to bridge the gap between educational and clinical outcomes, in line with medical education's social accountability to train physicians who can provide better and safer care.

METHODS AND ANALYSIS

Study design and setting

GEMINI MedED is a retrospective cohort study of residents in the UofT IM programme and the patients cared for at four of its affiliated hospitals. The retrospective analysis will include data from 1 April 2010 to 30 December 2020. Data collection and analysis will begin on 1 January 2023, and is anticipated to conclude by 31 December 2023. The UofT IM programme enrolls approximately 70 new residents each academic year, with approximately 225 enrolled at any one time. IM residents spend the majority of their training rotating through several UofT-affiliated teaching hospitals including Mount Sinai Hospital, Sunnybrook Health Sciences Centre, Toronto General Hospital and Toronto Western Hospital; see [table 1](#) for hospital details. Although each hospital is affiliated with UofT, the daily operations of each hospital's IM department are independent of one another.

Table 1 Metrics for academic hospitals involved in the General Medicine Inpatient Initiative Medical Education Database for the 2020 fiscal year (1 April 2020 to 3 March 2021, inclusive)⁶³

	Mount Sinai Hospital	Sunnybrook Health Sciences Centre	University Health Network*
Number of emergency department visits	42 865	49 304	84 707
Number of acute care admissions	25 277	31 112	32 562
Number of acute care beds	291	623	909
Mean admission length (days)	4.6	6.9	8.3
Hospital occupancy rate	79%	82%	79%
Electronic Health Record System (Manufacturer)	PowerChart (Cerner)	SunnyCare	EPR (QuadraMed)

*The University Health Network is comprised of the Toronto General Hospital and the Toronto Western Hospital; both hospitals have internal medicine departments that admit inpatients.
EPR, Electronic Patient Record.

Each hospital has an EM department that refers patients to the IM service for possible admission. EM physicians are generally scheduled for 8 hour shifts with a variable number of EM physicians scheduled at any given time based on the predicted emergency department volume patterns. Conversely, the IM service at each hospital is generally comprised of four core IM teams, each staffed by a single IM physician who is arbitrarily assigned a senior IM resident, junior residents and medical students. Scheduling for the EM and the IM departments is performed entirely independent of one another. During the day, each core IM team cares for admitted inpatients with a separate day team (staffed by an internist) responsible for EM admissions. Overnight, a single senior IM resident is responsible for overseeing all patients referred to IM as they are admitted to the four IM teams. All IM admissions performed by junior residents (defined as residents in postgraduate year one) and medical students, including the associated investigations, admission and treatment orders, are reviewed by the senior IM resident. Additionally, for admissions performed by medical students, senior IM residents must electronically cosign all orders entered by medical students, who do not have ordering privileges. There is one other major training site affiliated with the UofT IM programme (St. Michael's Hospital), for which we did not have historical resident call schedules and therefore will not be included. Together, these hospitals serve the city of Toronto (population 2 960 000).⁴⁶

GENERAL MEDICINE INPATIENT INITIATIVE

GEMINI MedED is an extension of the GEMINI retrospective cohort study (formerly known as the General Medicine Inpatient Initiative) and will derive clinical data from GEMINI. GEMINI is a previously established research and quality improvement database that collects data from every patient admitted to the IM department at multiple Ontario hospitals, including the four academic hospitals listed above. The development of the GEMINI dataset has been previously described.⁴² In brief, the GEMINI study collects clinical data from EHRs and other hospital clinical repositories, then links these data to administrative data as reported by hospitals to the Canadian Institute of Health Information (CIHI) Discharge Abstract Database and National Ambulatory Care Reporting System. Clinical data includes laboratory and radiology investigations, medication and treatment orders, dietary orders and vitals/clinical monitoring. CIHI administrative data include patient demographics, most responsible discharge diagnoses, comorbidities, discharge dispositions, in-hospital interventions and resource use/cost associated with each admission. Based on these raw data, the GEMINI study also derives other variables including the aggregate comorbidity level (eg, Charlson Comorbidity Index) and acuity of admission (eg, Laboratory-based Acute Physiology Score).^{47 48} Data is currently available between 1 April 2010 and 31 December 2020, is updated two-to-three times per year and has been used

in studies ranging from resource utilisation to tracking coronavirus disease of 2019 patient outcomes.^{45 49}

Participants

This study will include both residents and patients. Senior IM residents (defined as residents in postgraduate years 2 or 3) enrolled in the UofT IM programme between 1 April 2010 and 31 December 2020 are eligible for inclusion; this timeframe was selected to match the data availability of the GEMINI database. Approximately 900 IM residents will be included in the database, but the exact number will be calculated once data collection is complete. Patients admitted during an overnight call shift to IM from the EM department at the four academic teaching hospitals between 1 April 2010 and 31 December 2020 are eligible for inclusion; this data will be derived from the GEMINI database. Patients admitted overnight to the intensive care unit (ICU), family medicine-staffed hospitalist teams or any team not under the purview of the senior IM resident will not be included to this study. Furthermore, patients who are already admitted and then transferred to an IM team will not be included. Overall, GEMINI MedED allows linkage of individual patients to the senior IM resident who supervised their care overnight, thus facilitating attribution of care and outcomes to individual residents.

Data collection and extraction

All senior IM resident data will be derived from IM overnight call schedules. Call schedules will be retrospectively collected from the postgraduate medical education offices at all four participating hospitals. Only the latest version of each call schedule will be used, ensuring any last-minute substitutions (including absences due to illness and vacations) are accounted for. The hospital site, date of call shift and name of the senior resident leading the overnight team will be manually extracted from each call schedule. Overall, 5% of the data will be randomly selected for independent parallel extraction to ensure accuracy; 5% was an arbitrarily selected value used previously in the literature.⁵⁰

Data linkage

As senior IM residents are scheduled for call shifts multiple times throughout the year and across different hospital sites, residents will be linked to all their overnight shifts by name. Every attempt will be made to link shifts to residents by first and last name. However, inferences may be necessary if the call schedule does not provide enough information to confidently link a resident to a call shift; the most common scenario being that only the last name of the resident was listed on the call schedule. In that event, that shift will be assigned to the resident working in the same academic year and hospital with the same last name. If two residents in the same academic year and hospital have the same last name, then the shift will be assigned to the resident who most recently worked a shift. In a preliminary survey of the data, less than 1% of

our resident population have the same last names and are assigned to the same hospital, in the same academic year. After residents are linked to their overnight shifts, they will be deidentified and given a unique study ID.

We will assign all patients admitted during the overnight period to the senior IM resident on-call, at that hospital site, on that date through the GEMINI database. All patients admitted between 18:00 and 8:00 the following day will be assigned to the senior IM resident on-call that night. Although overnight call shifts typically start at 17:00, we chose an 18:00 start time for assigning patients to minimise the possibility of inappropriately assigning patients admitted by the daytime team to the overnight on-call resident. Weekday, weekend and holiday admissions will be treated with the same data linkage procedure (ie, only overnight admissions will be included in this study). The time of admission will be used as the start point to distinguish orders placed by the IM team from those placed by the EM team. 8:00 will be the endpoint to distinguish orders placed by the overnight IM team from those placed by the attending physician and the day team.

Patient and public involvement

It was not appropriate or possible to involve patients or the public in the design, or conduct, or reporting, or dissemination plans of our research.

Study aims

Aim 1: Describe case mix and volumes of overnight IM admissions and measure variation across residents, hospitals and over time.

Aim 2: Describe a proof-of-concept approach to measuring unwarranted variation in senior IM resident clinical practice using RSQMs.

Aim 3: Describe variation in patient outcomes between senior IM residents and link variation in patient outcomes to resident care using a proof-of-concept approach.

Study outcomes and statistical analyses

Aim 1: case mix

GEMINI MedED will examine the typical case mix encountered by IM residents overnight and how case mix varies among residents, hospitals and academic years. Case-mix data includes patient demographics, patient volume, breadth of presentations, acuity and complexity.^{47 48} See [table 2](#) for proposed case-mix variables. For case-mix measures that tend to evolve over the course of admissions (ie, acuity), a time limiter will be placed on each measure. For instance, ICU transfers may be limited to transfers occurring within 24 hours of admission. This data will be summarised using descriptive statistics and differences between residents, hospital sites and over time will be summarised using measures of variance.

Aim 2: resident sensitive quality measures

Goals of analysis

We aim to demonstrate a proof-of-concept approach to measuring variation in IM resident clinical practice using RSQMs. This study will explore unanswered questions

Table 2 Factors, outcome variables and their definitions to be included in the case-mix analysis

Variable	Definition
Factors	
Resident	Each unique resident
Site	Hospital site of call shift
Time	Academic year (each academic year is between 1 July and 30 June of the next year, not level resident level of training)
Outcome variables	
Demographic outcomes	
Patient age	Age of patient on admission
Patient sex	Biological sex of patient
Long-term residence	If the patient resided in a long-term care residence prior to admission
Volume Outcomes	
Volume	Number of admissions per call shift
Breadth outcomes	
Primary discharge diagnosis	Most responsible diagnosis assigned to the admission on discharge
Complexity outcomes	
Comorbidities	Pre-existing diagnoses associated with the patient that are not the primary discharge diagnosis
Charlson Comorbidity Index (CCI)	CCI derived from patient age and comorbidities at the time of discharge
Hospital Frailty Risk Score (HFRS)	HFRS derived from admission diagnoses and comorbidities at the time of discharge
Readmission	Patients with an admission to an acute care hospital in the past 30 days
Acuity outcomes	
Laboratory Acute Physiology Score (LAPS)	LAPS derived from the earliest set of laboratory values available during admission
In-hospital death	Patient death during the hospital admission (divided into overall admission, within 48 hours and within 7 days)
Intensive care transfer	Patients requiring a subsequent intensive care unit admission after an initial admission to the internal medicine department (divided into overall admission, within 48 hours and within 7 days)

regarding how RSQMs can be operationalised for learner assessment/feedback:

- Identification of resident-attributable care. Given clinical outcomes often reflect the work of teams rather than individual residents, how can measures



be developed that are mostly attributable to resident care?

- ▶ Identifying warranted versus unwarranted variation. How can measures be developed that identify truly unwarranted variation in resident care?
- ▶ Feasibility of data-driven metrics. How can measures be feasibly developed in the context of EHR data limitations?
- ▶ Applicability to medical education. How will this data be useful at different levels of medical education as a measure of clinical and educational quality?

RSQM classification

We will categorise RSQMs into three main groups based on available clinical evidence: guideline-concordant RSQMs, discretionary RSQMs, and guideline-discordant RSQMs. We define this novel nomenclature as follows: guideline-concordant RSQMs are measures that should be performed nearly 100% of the time based on clinical evidence (eg, following local antimicrobial guidelines when ordering antibiotics for suspected pneumonia); discretionary RSQMs have equivocal evidence and their performance is context-dependent; and guideline-discordant RSQMs should rarely be performed based on clinical evidence (eg, generalised ordering of benzodiazepines for sleep in patients over the age of 65).

This classification scheme focuses on the identification of warranted versus unwarranted variation, in keeping with Wennberg's framework for care variation (effective care, preference-sensitive care and supply-sensitive care).⁵¹ Wennberg's framework helps predict whether identified variation is truly unwarranted, or the result of equivocal evidence and/or healthcare supply. For example, if residents were consistently underperforming guideline-concordant RSQMs, this

would likely represent unwarranted variation. Moreover, overperformance of guideline-discordant RSQMs, which represent potentially harmful practices, would also represent unwarranted variation. Discretionary RSQMs reflect preference-sensitive and supply-sensitive care, where there is more than one accepted approach to care, and variation in these metrics may not necessarily represent unwarranted variation. Instead, variation in these measures likely implies wide-ranging clinical practices and offers potential utility from an educational lens, such as promoting standardisation where necessary. Thus, our study aims to improve our understanding of which RSQMs are the most appropriate metrics for resident assessment. We anticipate that guideline-concordant and guideline-discordant RSQMs (ie, metrics which should be performed or avoided nearly 100% of the time) are likely the most appropriate for resident assessment

RSQM selection

An initial list of RSQMs will be drafted by drawing on four sources: (1) previously proposed RSQMs in IM,^{33–35} (2) previously proposed disease-specific quality measures,^{52–54} (3) Canadian and international practice guidelines^{55–57} and (4) existing physician practice feedback reports^{3 40} (see table 3). We plan to include both disease-specific and general measures. Disease-specific measures reflect care for a particular diagnosis, such as prescribing guideline-directed antibiotic therapy for pneumonia. We anticipate including metrics for three to five common IM diagnoses, such as pneumonia, congestive heart failure (CHF) and chronic obstructive pulmonary disease (figure 1). In turn, general measures are not specific to a particular diagnosis (eg, prescribing venous thromboembolism prophylaxis) (see table 3).

Table 3 Examples of resident-sensitive quality measures (RSQMs) for pneumoniaspecific and general clinical care

Metric	Numerator	Denominator	RSQM classification*
Pneumonia-specific RSQMs			
% Guideline-recommended antibiotic therapy	Total number of patients with guideline-recommended antibiotic treatment on admission	Number of admissions	Guideline-concordant
% CT of the chest	Total number of patients with CT of the chest order on admission	Number of admissions	Discretionary
% Empiric anaerobic coverage in aspiration pneumonia	Total number of patients with empiric anaerobic antibiotics orders on admission	Number of admissions	Guideline-discordant
General RSQMs			
% VTE prophylaxis order set completed	Total number of patients with VTEp order set implemented (including those with stated reason for not ordering VTEp) on admission	Number of admissions	Guideline-concordant
% Broad-spectrum antibiotic use	Total number of patients with broad spectrum antibiotic orders on admission	Number of admissions	Discretionary
% Benzodiazepine use	Total number of patients over 65 years old with benzodiazepine orders on admission	Number of admissions	Guideline-discordant

*Guideline-concordant RSQMs are measures that should be performed most of the time based on clinical evidence. Discretionary RSQMs have equivocal evidence and whether they should be performed is context dependent. Guideline-discordant RSQMs are measures that should be rarely performed based on clinical evidence.

RSQMs, resident-sensitive quality measures; VTE, venous thromboembolism.

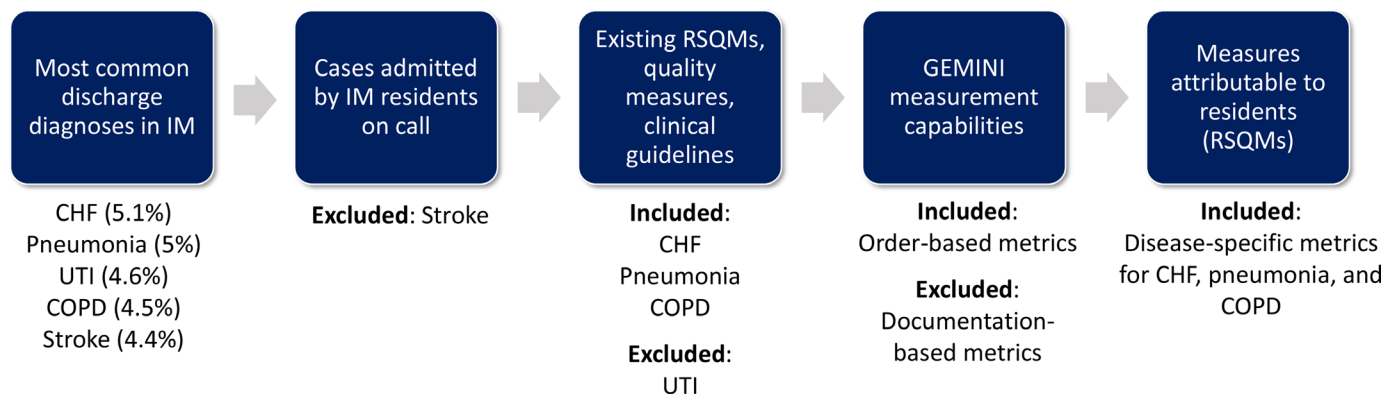


Figure 1 Selection process for developing disease-specific RSQMs for internal medicine. The three diseases selected for this study were congestive heart failure, pneumonia and chronic obstructive pulmonary disease. These three diseases were selected based on their commonality in internal medicine. A literature review will be conducted to determine RSQMs for these three diseases. Four major sources will be the primary contributor to RSQMs: (1) previously proposed RSQMs in internal medicine, (2) previously proposed disease-specific quality measures, (3) Canadian and international practice guidelines and (4) existing physician practice feedback reports. Urinary tract infections were excluded as the management largely depends on the local resistance patterns of microbes; thus, practice guidelines often suggest following local antimicrobial sensitivities. RSQMs were also selected based on the GEMINI dataset's measurement capabilities. This includes clinical data (laboratory and radiology investigations, medication and treatment orders, dietary orders and vitals/clinical monitoring) and administrative data (patient demographics, most responsible discharge diagnoses, comorbidities, discharge dispositions, in-hospital interventions and resource use/cost). Based on these raw data, the GEMINI study also derives other variables including the aggregate comorbidity level (eg, Charlson Score) and acuity of admission (eg, Lab-based Acute Physiology Score). This selection process will be repeated for general RSQMs. CHF, congestive heart failure; COPD, chronic obstructive pulmonary disease; GEMINI, General Medicine Inpatient Initiative; IM, internal medicine; RSQM, resident-sensitive quality measure; UTI, urinary tract infection.

The proposed RSQMs will then be reviewed by project team members using the modified Delphi method, with specific prompting to quantitatively rate proposed measures using our core guiding principles and provide qualitative comments.^{58 59} This process will also be used to categorise each measure as a guideline-concordant, discretionary or guideline-discordant RSQM through consensus methods. In keeping with prior Delphi approaches to develop RSQMs, we will conduct three rounds with a priori inclusion and exclusion thresholds based on the final number of participants. In each subsequent round, participants will receive a spreadsheet that includes measures that have not yet met the threshold for inclusion or exclusion, their own previous rating, the distribution of the group's ratings and anonymised comments.³⁵ Of note, we are not aiming to specifically validate these RSQMs or propose they are widely generalisable for other residency programmes. Rather, we aim to explore their feasibility in measuring resident-level care variation in our local context as a proof of concept.

Finally, one important consideration is that this study will be conducted in a staged approach, meaning that the RSQM analysis will begin after the case-mix analysis is completed. The feasibility of disease-specific RSQMs relies on sufficient case volumes to reduce the effect of potential confounding variables (eg, particular hospital sites or care being driven by a recurring subset of EM physicians). Hence, knowledge generated from the case-mix analysis may prompt a pivot in our analytical approach to RSQMs,

such as a shift toward general measures if resident case volumes for specific diagnoses are globally low.

Analysis plan

Our analysis will focus on identifying patterns of variation in RSQMs. In keeping with prior studies, RSQMs will be operationalised into a binary outcome (eg, ordered or not ordered) and performance rates of individual RSQMs will be summarised with descriptive statistics and measures of variance. For each RSQM, the variation will be determined between residents, hospital sites and over time. We anticipate that guideline-concordant RSQMs will have a high frequency of performance approaching 100%, while guideline-discordant RSQMs will have a low frequency of performance approaching 0%. Significant deviation from the anticipated frequency would likely represent unwarranted variation. We will perform a separate analysis for discretionary RSQMs, which are only indicated in particular clinical circumstances (eg, ordering CT of the chest in patients with pneumonia) and therefore the desired frequency of performance is not clear. These measures tend to reflect preference-sensitive and supply-sensitive care, and we anticipate greater variation in these metrics.⁵¹ Again, this variation will be characterised between residents, sites, and over time. Both uniformity and variance in RSQMs may offer valuable insight. For example, within-hospital and within-year variance in RSQM performance may suggest resident-level variation, while between-hospital or between-year variance may suggest systems-level variation.³⁴ We anticipate wide

variation in completion of measures defined as preference-sensitive care, given their reliance on provider/patient decision-making.

We recognise that some measures of effective care may be performed by EM physicians or other providers prior to the senior resident's involvement. For some RSQMs, distinguishing the physician who entered the order may be important in attributing patient care to residents. For these measures, the time the order was placed relative to admission can be used to distinguish between orders placed by the EM physician, overnight IM team and daytime IM team. Please see the Methods and Analysis section, subsection Data Linkage, for additional details. However, for other RSQMs, distinguishing the physician who entered the order is not as important in attributing patient care to residents. It is the ultimate responsibility of the senior IM resident to ensure a care process is completed, irrespective of which physician enacts these care processes. For example, we anticipate that chest x-ray (CXR) ordering in pneumonia should approach 100% as it is a clear measure of effective care (ie, almost all patients admitted overnight with pneumonia should have had a CXR ordered, whether by the resident or another provider). In this case, regardless of whether the referring EM physician ordered the test, it is the senior IM resident's responsibility to ensure this standard of care is met for their patients. For these metrics, it is the deviation from the desired outcome, or the variation seen in these metrics, that may be attributable to resident care (while the absolute frequency is driven by both EM and IM physicians).

Aim 3: patient outcomes

There will be two analyses of patient outcomes. The first analysis will measure variation in patient outcomes between residents. Traditionally, patient outcomes are attributed entirely to staff physicians, but emerging evidence suggests that resident practices also influence care processes for patients in academic hospitals.⁶⁰ In this study, patient outcomes more proximal to the admission date may be more readily attributable to care provided by the on-call senior resident. For instance, the proportion of ICU transfers within 24 hours of admission is temporally close to admission.³ Conversely, outcomes distal to admission, for instance 30-day readmission rates, are less attributable and more likely to be influenced by other providers and external factors. Thus, a time limit may be placed on patient outcomes (eg, within 24 or 48 hours of admission) to ensure that clinical outcomes are at least partially attributable to the admitting senior IM resident, and less likely to be confounded by care provided by the day teams.

Similar to RSQMs, outcomes may be further stratified as disease-specific or general (figure 2). GEMINI tracks five general outcomes, which are ICU utilisation rates, in-hospital mortality, 30-day readmission rates, use of advanced imaging modalities and length of stay. These same outcomes may be used as disease-specific outcomes

if the patient population is restricted to only include those with a particular diagnosis. In addition, disease-specific outcomes may include unique measures only relevant to one disease: for instance, number of days requiring oxygen therapy in pneumonia. As this is a staged study, the exact disease-specific outcomes will be decided based on data availability, results from the case-mix analysis and attributability/relevance to senior residents. All outcomes will be summarised with descriptive statistics and variations between residents, hospital sites, and over time will be summarised using measures of variance.

The second analysis will be an exploratory analysis aimed at linking resident care quality with resulting patient outcomes. One would hypothesise that residents with positive educational outcomes would also have better patient outcomes. But in the literature, there is scarce evidence linking strong academic performance in CBME with improvements in patient outcomes.⁶¹ Thus, the aim will be to validate RSQM composite scores as a marker for overall resident care quality and, in turn, link better care with improved patient outcomes. A multivariate regression model may be used to determine whether RSQM composite scores are associated with improved patient outcomes. The exact RSQMs included in the composite score and the patient outcomes included in the regression model will largely depend on the results of previous analyses. For a measure to be selected for inclusion to the composite score, in addition to fitting the definition of an RSQM (ie, directly meaningful to patient care and attributable to a single resident), there needs to be sufficient variation between residents to enable the composite score to distinguish variation in resident performance. Of note, discretionary RSQMs are not evidence based or based only on weak evidence. As a result, any correlation between discretionary RSQMs and improvements/detriments to patient outcomes may reflect over-detection or confounding rather than a true association. To mitigate this risk, only guideline-concordant and guideline-discordant RSQMs (which are both rooted in established clinical evidence) will be included as RSQMs for aim 3. Finally, availability of data in GEMINI will also guide selection. Of note, the validation of individual RSQMs on improvements/detriments in patient outcomes is beyond the scope of this study.

LIMITATIONS AND STRENGTHS

There are four major limitations to this study. First is the difficulty attributing patient outcomes and, to a lesser extent, RSQMs entirely to senior residents. External factors such as hospital policies, time trends and input from other members of the healthcare team all influence these measures. For instance, care processes may be enacted by the EM physician even before the resident becomes involved in the patient's care. We will attempt to control these external factors by examining senior residents in a clinical environment where they have substantial influence over patient care (overnight call shifts)

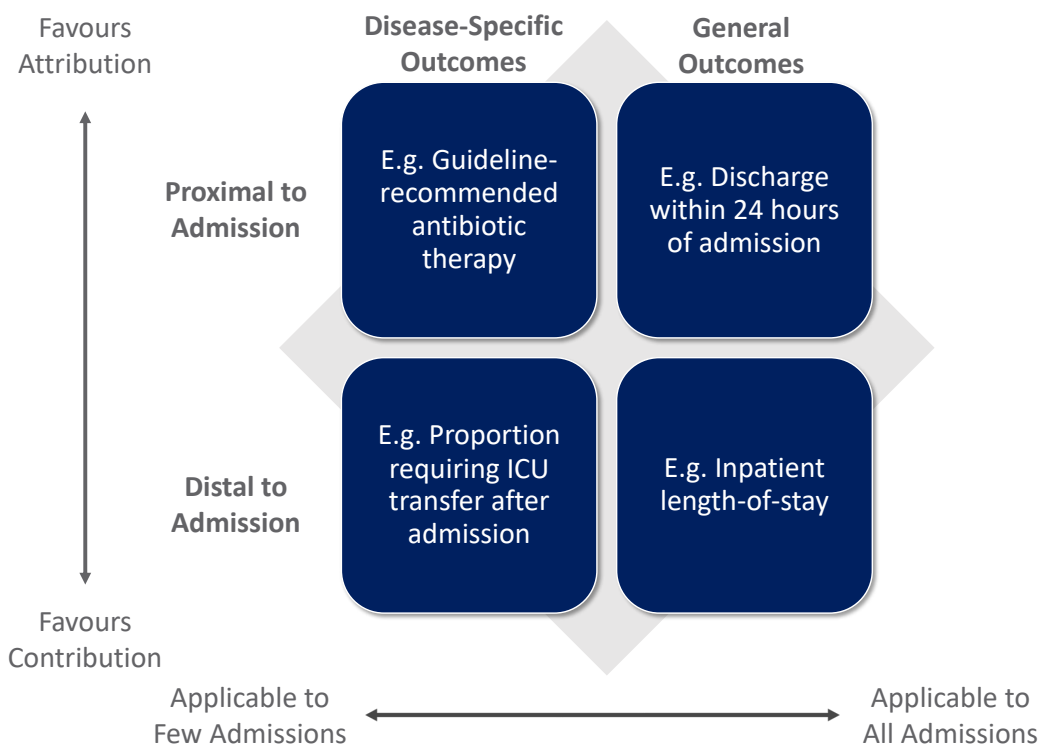


Figure 2 Visual representation of attribution and applicability spectrum of patient outcomes for senior internal residents on overnight internal medicine call. With respect to the attributability, outcomes more proximal to the time of admission are more attributable to care provided by residents overnight. In contrast, outcomes more distal to the time of admission may only be partially influenced by the care provided by the senior resident overnight. With respect to the applicability, outcomes focused on one disease process may only apply to a few patients within our study cohort. In contrast, general outcomes may apply to most or all patients in our cohort. ICU, intensive care unit.

and by collecting data across multiple hospitals and time settings to enable variation between residents to emerge. Of note, residents are given consults by numerous EM physicians and arbitrarily assigned to multiple staff inter-nists throughout their residency. Senior IM residents complete overnight shifts at all four academic hospitals included to this study over the course of their postgraduate years 2 and 3, whereas most EM and IM staff primarily practice out of a single hospital. Additionally, IM residents work overnight shifts throughout the academic year in a ‘1 in 4’ system; this means residents are on-call every 4th night in hospital. Staffing and scheduling IM shifts are done entirely independently from that of the EM team, the latter of which do not follow a ‘1 in 4’ system. Given a large enough sample, variations in clinical care and patient outcomes may partially reflect resident practice rather than entirely reflect staff influence. However, as ‘resident-staff’ pairings are not tracked at our hospitals, this remains a potential limitation of this study. Second is this study will be carried out at a single residency programme. Although multiple hospitals are included, all hospitals and residents are affiliated with the UofT IM programme. Other residency training programmes have staffing policies, geographic and/or programme design differences that may limit the direct generalisability of this study’s design and results. Regardless, our study may

serve as a proof of concept, demonstrating the feasibility of linking residents to important clinical data and patient outcomes.

Third, this study relies on an existing clinical database (GEMINI) for patient data, which may not capture certain clinical and patient outcomes. For instance, the ordering physician for investigations and treatment is not captured by the database. However, at all four participating institutions orders (especially for therapy) can only be placed by the IM team after the patient has been formally admitted. As such, the timing of orders relative to the timing of admission may be a suitable surrogate for the author of the order. Finally, this study is a proof-of-concept study with an overarching goal of implementing Big Data and clinical informatics into medical education. Although this study is focused on resident clinical care through RSQMs, the study of whether this data improves resident performance, the utility of data for resident assessment and promotion, and the validation of individual clinical actions/measures with patient outcomes is beyond the scope of this study.

There are several strengths to this study. First, patient data for this study is derived from the GEMINI database, which is based on objective clinical data derived from EHRs that has been previously validated for accuracy.⁶² Unlike trainee-reported case logs or preceptor-driven



evaluations, both patient and resident data are derived from empiric sources with a lower risk of reporting bias and error. Second, this study derives data from multiple hospitals and academic years. This provides a comprehensive view of the IM residents' clinical practice. Although call schedules may not be available for certain timepoints over the past decade, the vast majority of residents in the programme will be represented within the final database.

Residency education remains the cornerstone of developing physicians capable of delivering high-quality and safe care. The introduction of CBME has emphasised the importance of WBAs but large deficiencies still remain in terms of tracking case mix and using objective clinical data in resident assessment. GEMINI MedED will bring together large clinical datasets with new methods of resident assessment to further our understanding of how best to incorporate clinical measures and outcomes into medical education. Ultimately, this may provide residents with more in-depth feedback on their clinical training, educators with more analytics of their residency programme, and patients with the assurance they are receiving the highest quality care.

Author affiliations

¹Department of Medicine, University of Toronto Faculty of Medicine, Toronto, Ontario, Canada

²Department of Medicine, Division of General Internal Medicine, University of Toronto Faculty of Medicine, Toronto, Ontario, Canada

³Li Ka Shing Knowledge Institute, Unity Health Toronto, Toronto, Ontario, Canada

⁴Division of General Internal Medicine, Unity Health Toronto, Toronto, Ontario, Canada

⁵Division of General Internal Medicine, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada

⁶Department of Medicine, Division of Respiriology, University of Toronto Faculty of Medicine, Toronto, Ontario, Canada

⁷Division of Respiriology, Sinai Health System, Toronto, Ontario, Canada

Twitter Brandon Tang @DrBrandonTang, Anushka Lalwani @AnushkaKLalwani, Amol A Verma @AmolAVerma, Brian M Wong @Brian_M_Wong, Fahad Razak @DrFahadRazak and Shiphra Ginsburg @sginsburg1

Acknowledgements The authors thank Dr Daniel Schumacher and Dr Ben Kinnear for their invaluable insight and feedback on this project. The authors thank Natasha Campbell, Heather Smith-St. Kitts and James LeBlanc for their invaluable contribution collecting and organising study data. The authors thank Serina Cheung for her invaluable help with copy editing and formatting the final manuscript.

Contributors ACLL, BT, BW, FR and SG were involved in the planning, conceptualisation and design of this study. ACLL, BT, AL, AV, FR and SG were involved in data acquisition. ACLL, BT, AV, BW, FR and SG were involved in data interpretation and analysis. AV, BW, FR and SG were involved in supervision and project administration. All authors were involved in the writing and approval of the final manuscript. ACLL and BT contributed equally to this manuscript. FR and SG contributed equally to this manuscript.

Funding This work was supported by the Ontario Medical Student Association Medical Student Education Research Grant (2020–2021). No grant number is available for this grant.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which

permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Andrew CL Lam <http://orcid.org/0000-0001-6923-2721>

REFERENCES

- Qian F, Osler TM, Eaton MP, *et al.* Variation of blood transfusion in patients undergoing major noncardiac surgery. *Ann Surg* 2013;257:266–78.
- Shashar S, Ellen M, Codish S, *et al.* Medical practice variation among primary care physicians: 1 decade, 14 health services, and 3,238,498 Patient-Years. *Ann Fam Med* 2021;19:30–7.
- Verma AA, Guo Y, Jung HY, *et al.* Physician-level variation in clinical outcomes and resource use in inpatient general internal medicine: an observational study. *BMJ Qual Saf* 2021;30:123–32.
- Hartford EA, Klein EJ, Migita R, *et al.* Improving patient outcomes by addressing provider variation in emergency department asthma care. *Pediatr Qual Saf* 2021;6:e372.
- Corallo AN, Croxford R, Goodman DC, *et al.* A systematic review of medical practice variation in OECD countries. *Health Policy* 2014;114:5–14.
- Birkmeyer JD, Finks JF, O'Reilly A, *et al.* Surgical skill and complication rates after bariatric surgery. *N Engl J Med* 2013;369:1434–42.
- Asch DA, Nicholson S, Srinivas S, *et al.* Evaluating obstetrical residency programs using patient outcomes. *Obstet Gynecol Surv* 2010;65:152–3.
- Bansal N, Simmons KD, Epstein AJ, *et al.* Using patient outcomes to evaluate general surgery residency program performance. *JAMA Surg* 2016;151:111–9.
- Chen C, Petterson S, Phillips R. Spending patterns in region of residency training and subsequent expenditures for care provided by practicing physicians for Medicare beneficiaries 2021;20037:2385–93.
- Phillips RL, Holmboe ES, Bazemore AW, *et al.* Purposeful imprinting in graduate medical education: opportunities for partnership. *Fam Med* 2021;53:574–7.
- Sirovich BE, Lipner RS, Johnston M, *et al.* The association between residency training and internists' ability to practice conservatively. *JAMA Intern Med* 2014;174:1640–8.
- Prystowsky JB, Bordage G. An outcomes research perspective on medical education: the predominance of trainee assessment and satisfaction. *Med Educ* 2001;35:331–6.
- Haan CK, Edwards FH, Poole B, *et al.* A model to begin to use clinical outcomes in medical education. *Academic Medicine* 2008;83:574–80.
- Fitzgibbons JP, Bordley DR, Berkowitz LR. Redesigning residency education in internal medicine. *A position paper from the Association of Program Directors in Internal Medicine.* *Ann Intern Med* 2006;144:920–6.
- Iglar K, Murdoch S, Meaney C. Does clinical exposure matter? *Can Fam Physician* 2018;64:42–8.
- McCoy CP, Stenerson MB, Halvorsen AJ, *et al.* Association of volume of patient encounters with residents' in-training examination performance. *J Gen Intern Med* 2013;28:1035–41.
- Rule A, Chiang MF, Hribar MR. Using electronic health record audit logs to study clinical activity: a systematic review of aims, measures, and methods. *J Am Med Inform Assoc* 2020;27:480–90.
- Mattana J, Kerpen H, Lee C, *et al.* Quantifying internal medicine resident clinical experience using resident-selected primary diagnosis codes. *J Hosp Med* 2011;6:395–400.
- Bachur RG, Nagler J. Use of an automated electronic case log to assess fellowship training: tracking the pediatric emergency medicine experience. *Pediatr Emerg Care* 2008;24:75–82.
- Balla F, Garwe T, Motghare P, *et al.* Evaluating coding accuracy in general surgery residents' accreditation Council for graduate medical education procedural case logs. *J Surg Educ* 2016;73:e59–63.
- Dermoddy SM, Gao W, McGinn JD, *et al.* Case-Logging practices in otolaryngology residency training: national survey of residents and program directors. *Otolaryngol Head Neck Surg* 2017;156:1072–7.
- McPheeters MJ, Talcott RD, Hubbard ME, *et al.* Assessing the accuracy of neurological surgery resident case logs at a single institution. *Surg Neurol Int* 2017;8:206.
- Whitcomb ME. Transforming medical education: is competency-based medical education the right approach? *Acad Med* 2016;91:618–20.

- 24 Lockyer J, Carraccio C, Chan M-K, *et al*. Core principles of assessment in competency-based medical education. *Med Teach* 2017;39:609–16.
- 25 Thoma B, Caretta-Weyer H, Schumacher DJ. The deliberately developmental organization: a conceptual framework for CBME. *Med Teach* 2021.
- 26 Gofton W, Dudek N, Barton G. Workplace-Based Assessments Implementation Guide: formative tips for medical teaching practice. In: *R coll physicians Surg Canada*. 1st ed, 2017: 1–12.
- 27 Ten Cate O, Olle TC. Competency-Based postgraduate medical education: past, present and future. *GMS J Med Educ* 2017;34:Doc69 <http://www.ncbi.nlm.nih.gov/pubmed/29226237%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=>
- 28 Massie J, Ali JM. Workplace-Based assessment: a review of user perceptions and strategies to address the identified shortcomings. *Adv Health Sci Educ Theory Pract* 2016;21:455–73.
- 29 Miller A, Archer J. Impact of workplace based assessment on doctors' education and performance: a systematic review. *BMJ* 2010;341:c5064.
- 30 Kogan JR, Conforti L, Bernabeo E, *et al*. Opening the black box of clinical skills assessment via observation: a conceptual model. *Med Educ* 2011;45:1048–60.
- 31 Ofstad EH, Frich JC, Schei E, *et al*. Clinical decisions presented to patients in hospital encounters: a cross-sectional study using a novel taxonomy. *BMJ Open* 2018;8:e018042–9.
- 32 Schumacher DJ, Holmboe ES, van der Vleuten C, *et al*. Developing resident-sensitive quality measures: a model from pediatric emergency medicine. *Acad Med* 2018;93:1071–8.
- 33 Schumacher DJ, Martini A, Holmboe E, *et al*. Initial implementation of resident-sensitive quality measures in the pediatric emergency department: a wide range of performance. *Acad Med* 2020;95:1248–55.
- 34 Schumacher DJ, Wu DTY, Meganathan K, *et al*. A feasibility study to attribute patients to primary interns on inpatient ward teams using electronic health record data. *Acad Med* 2019;94:1376–83.
- 35 Kinnear B, Kelleher M, Sall D, *et al*. Development of Resident-Sensitive quality measures for inpatient general internal medicine. *J Gen Intern Med* 2021;36:1271–8.
- 36 Schumacher DJ, Holmboe E, Carraccio C, *et al*. Resident-sensitive quality measures in the pediatric emergency department: exploring relationships with supervisor entrustment and patient acuity and complexity. *Acad Med* 2020;95:1256–64.
- 37 Office of the National Coordinator for Health Information Technology. *Non-federal*. Washington DC: Acute Care Hospital Electronic Health Record Adoption, 2015. <https://www.healthit.gov/data/quickstats/non-federal-acute-care-hospital-electronic-health-record-adoption>
- 38 Weinstein DF. Optimizing GME by measuring its outcomes. *N Engl J Med* 2017;377:2007–9.
- 39 Iroju OG, Olaleke JO. A systematic review of natural language processing in healthcare. *IJITCS* 2015;7:44–50.
- 40 Verma AA, Razak F. Lessons for hospital care from the first wave of COVID-19 in Ontario, Canada. *Hosp Pract* 2021;49:229–31.
- 41 Van den Bulck S, De Burghraeve T, Raat W, *et al*. The effect of automated audit and feedback on data completeness in the electronic health record of the general physician: protocol for a cluster randomized controlled trial. *Trials* 2021;22:1–9.
- 42 Verma AA, Guo Y, Kwan JL, *et al*. Patient characteristics, resource use and outcomes associated with general internal medicine hospital care: the general medicine inpatient initiative (gemini) retrospective cohort study. *CMAJ Open* 2017;5:E842–9.
- 43 R-1 Main Residency Match - first iteration Quota Overview by Discipline. *Can. Resid. Matching Serv* 2021;2021 <https://www.carms.ca/match/r-1-main-residency-match/program-descriptions-archive-first-iteration/>
- 44 Top 10 specialties with the most medical residency positions | American Medical association, 2019. *Am. Med. Assoc*. Available: <https://www.ama-assn.org/residents-students/specialty-profiles/top-10-specialties-most-medical-residency-positions> [Accessed 17 Sep 2021].
- 45 Verma AA, Guo Y, Kwan JL, *et al*. Prevalence and costs of discharge diagnoses in inpatient general internal medicine: a multi-center cross-sectional study. *J Gen Intern Med* 2018;33:1899–904.
- 46 *Toronto C of Toronto at a glance*.2021
- 47 Sundararajan V, Henderson T, Perry C, *et al*. New ICD-10 version of the Charlson comorbidity index predicted in-hospital mortality. *J Clin Epidemiol* 2004;57:1288–94.
- 48 Escobar GJ, Greene JD, Scheirer P. Risk-Adjusting Hospital inpatient mortality using databases. *Med Care* 2008;46:232–9 <http://www.ncbi.nlm.nih.gov/pubmed/18388836>
- 49 Verma AA, Hora T, Jung HY, *et al*. Characteristics and outcomes of hospital admissions for COVID-19 and influenza in the Toronto area. *CMAJ* 2021;193:E410–8.
- 50 Liddy C, Wiens M, Hogg W. Methods to achieve high interrater reliability in data collection from primary care medical records. *Ann Fam Med* 2011;9:57–62.
- 51 Wennberg JE. Unwarranted variations in healthcare delivery: implications for academic medical centres. *BMJ* 2002;325:961–4.
- 52 Goldstein R, Kaplan A. Chronic obstructive pulmonary disease care in the community for adults, 2018. Available: <https://www.hqontario.ca/Evidence-to-Improve-Care/Quality-Standards/View-all-Quality-Standards/Chronic-Obstructive-Pulmonary-Disease/Resources>
- 53 Rhew DC. Quality indicators for the management of pneumonia in vulnerable elders. *Ann Intern Med* 2001;135:736–43.
- 54 Bonow RO, Ganiats TG, Beam CT, *et al*. ACCF/AHA/AMA-PCPI 2011 performance measures for adults with heart failure: a report of the American College of Cardiology Foundation/American heart association Task force on performance measures and the American medical Association-Physician Consortium for performance improvement. *Circulation* 2012;125:2382–401.
- 55 Wedzicha JA, Miravittles M, Hurst JR, *et al*. Management of COPD exacerbations: a European Respiratory Society/ American Thoracic Society guideline. *Eur Respir J* 2017;49. doi:10.1183/13993003.00791-2016. [Epub ahead of print: 15 Mar 2017].
- 56 Mandell LA, Marrie TJ, Grossman RF. Canadian guidelines for the initial management of community-acquired pneumonia: an evidence-based update by the Canadian infectious diseases Society and the Canadian thoracic Society. *Can J Cardiol* 2000;31:383–421.
- 57 Ezekowitz JA, O'Meara E, McDonald MA. Comprehensive update of the Canadian cardiovascular Society guidelines for the management of heart failure. *Can J Cardiol* 2017;2017:1342–433.
- 58 Fitch K, Berstein S, Aguilar M. The RAND/UCLA Appropriateness Method User's Manual. *RAND Corp St Monica CA* 2001.
- 59 Dalkey N, Olaf H. An experimental application of the Delphi method to the use of experts. *Science M* 2016;9:458–67 <http://www.jstor.org/stable/2627117>
- 60 Lau BD, Streiff MB, Pronovost PJ, *et al*. Attending physician performance measure scores and resident physicians' ordering practices. *JAMA Surg* 2015;150:813–4.
- 61 Brydges R, Boyd VA, Tavares W, *et al*. Assumptions about competency-based medical education and the state of the underlying evidence: a critical narrative review. *Acad Med* 2021;96:296–306.
- 62 Verma AA, Pasricha SV, Jung HY, *et al*. Assessing the quality of clinical and administrative data extracted from hospitals: the general medicine inpatient initiative (gemini) experience. *J Am Med Inform Assoc* 2021;28:578–87.
- 63 Explore your health system. *Can. Inst. Heal. Inf* 2020 <https://yourhealthsystem.cihi.ca/hsp/indepth?lang=en#/>