Measuring Overcrowding in Emergency Departments: A Call for Standardization

Supporting Informed Decisions
Until April 2006, the Canadian Agency for Drugs and Technologies in Health (CADTH) was known as the Canadian Coordinating Office for Health Technology Assessment (CCOHTA).


This report and the French version entitled Mesurer le surpeuplement de l’urgence : la nécessité d’uniformiser are available on CADTH’s web site.

This is the first in a series of four CADTH reports on emergency department (ED) overcrowding in Canada. The series looks at measures of ED overcrowding, and examines databases and information systems to monitor the issue. It also examines the frequency, determinants, and impacts of overcrowding. Finally, the series explores interventions used to reduce ED overcrowding and reviews which interventions are successful. An overview report on the series is available.

Production of this report is made possible by financial contributions from Health Canada and the governments of Alberta, British Columbia, Manitoba, New Brunswick, Newfoundland and Labrador, Northwest Territories, Nova Scotia, Nunavut, Ontario, Prince Edward Island, Saskatchewan, and Yukon. The Canadian Agency for Drugs and Technologies in Health takes sole responsibility for the final form and content of this report. The views expressed herein do not necessarily represent the views of Health Canada or any provincial or territorial government.

Reproduction of this document for non-commercial purposes is permitted provided appropriate credit is given to CADTH. CADTH is funded by Canadian federal, provincial, and territorial governments.

Legal Deposit - 2006
National Library of Canada

PUBLICATIONS MAIL AGREEMENT NO. 40026386
RETURN UNDELIVERABLE CANADIAN ADDRESSES TO
CANADIAN AGENCY FOR DRUGS AND TECHNOLOGIES IN HEALTH
600-865 CARLING AVENUE
OTTAWA ON K1S 5S8
Measuring Overcrowding in Emergency Departments: A Call for Standardization

Maria B. Ospina BSc MSc¹
Kenneth Bond BEd MA¹
Michael Schull MD MSc FRCPC²
Grant Innes MD CCFP FRCPC³
Sandra Blitz MSc⁴
Carol Friesen MA MLIS¹
Brian H. Rowe MD MSc CCFP (EM) FCCP⁵

May 2006

¹ University of Alberta Evidence-based Practice Center, University of Alberta, Edmonton AB
² Division of Emergency Medicine, University of Toronto and Institute of Clinical Evaluative Sciences, Toronto ON
³ Department of Emergency Medicine, Providence Health Care, St. Paul’s Hospital, Vancouver BC
⁴ Department of Emergency Medicine, University of Alberta, Edmonton AB
⁵ Department of Emergency Medicine, University of Alberta and Evidence-based Practice Center, University of Alberta, Edmonton AB
Reviewers

These individuals kindly provided comments on this report.

External Reviewers

Professor Matthew Cooke, PhD FRCS(Ed)  Ran Goldman, MD  
FFAEM DipIMC  Associate Professor  
Professor of Emergency Medicine  Department of Pediatrics  
University of Warwick  Division of Emergency Medicine and Clinical  
Coventry, UK  Pharmacology and Toxicology  

University of Toronto  
Toronto ON

CADTH’s Scientific Advisory Panel Reviewers

Jeff Scott, MBChB MHSc MHSA FRCPC  Charles Wright, MD MSc FRCS (C,E,Ed)  
Provincial Medical Officer of Health  Scientific Officer, Canadian Health Services  
Province of Nova Scotia  Research Foundation  
Halifax NS  Toronto ON

This report is a review of existing public literature, studies, materials and other information and documentation (collectively the “source documentation”) which are available to CADTH. The accuracy of the contents of the source documentation on which this report is based is not warranted, assured or represented in any way by CADTH, and CADTH does not assume responsibility for the quality, propriety, inaccuracies or reasonableness of any statements, information or conclusions contained in the source documentation.

CADTH takes sole responsibility for the final form and content of this report. The statements and conclusions in this report are those of CADTH and not of its Panel members or reviewers.

Authorship

Maria B. Ospina contributed to the conception and design of the review, Delphi study logistics, acquisition of data, and data analysis and interpretation; co-wrote the report; prepared the report for publication; and approved the final version of the manuscript.

Kenneth Bond contributed to study design, Delphi study logistics, acquisition of data, and data analysis and interpretation; co-wrote the report; prepared the report for publication; and approved the final version of the manuscript.

Michael Schull contributed to the conception of the study design, and acquisition of data; revised the report for intellectual content; and approved the final version of the manuscript.

Grant Innes contributed to the conception of the study design, and acquisition of data; revised the report for intellectual content; and approved the final version of the manuscript.

Sandra Blitz contributed to the data synthesis plan and data processing, and statistical analysis and interpretation; participated in drafting and revising the report; and approved the final version of the manuscript.
Carol Friesen designed and executed literature searches, and wrote the associated appendix on literature searching.

Brian H. Rowe is the guarantor of the study. He contributed to the conception and design of the review, acquisition of data, and design of data analysis and interpretation; co-wrote the report; revised it critically for intellectual content; and approved the final version of the manuscript.

**Acknowledgements**

The authors are grateful to the following individuals who served as members of the Technical Expert Panel for this report. Their expertise, contributions, and thoughtful comments were invaluable for the development and completion of this project:

- Marc Afilalo, MD, MCFP(EM), FACEP, FRCPC, Emergency Department, Sir Mortimer B Davis - Jewish General Hospital, and Faculty of Medicine, McGill University, Montréal QC
- Michael Bullard, BSc, MD, FRCPC, FACEP, CFPC-EM, Department of Emergency Medicine, University of Alberta, Edmonton AB
- Sam Campbell, MB, BCh, CCFPEM, Department of Emergency Medicine, Dalhousie University, Halifax NS
- Terry Klassen, MD, MSc, FRCPC, Department of Pediatrics and Evidence-based Practice Center, University of Alberta, Edmonton AB
- Carla Policicchio, BScN, RN, National Emergency Nurses Affiliation, Edmonton AB
- Michael Schull, MD, MSc, FRCPC, Division of Emergency Medicine, University of Toronto, and Institute for Clinical Evaluative Sciences, Toronto ON
- Doug Sinclair, MD, FRCPC, Department of Emergency Medicine, Dalhousie University, Halifax NS
- Mukarram Ali Zaidi, MBBS, MSc, Canadian Council on Health Services Accreditation, Ottawa ON

We are grateful to all the emergency physicians who contributed to the development of the pilot questionnaire used for the Delphi survey, and to those nurses and physicians who participated in the Delphi survey.

The authors also thank Carol Spooner, Harris Lari, and Nicola Hutton for their assistance selecting studies for the review. Liza Bialy helped design and manage the electronic forms for data extraction. Thanks to Allison Jaskiw, Kenneth Moreau, John Russell, and Jennifer Au for their contributions to this project.

**Conflicts of Interest**

Maria B. Ospina, Kenneth Bond, Sandra Blitz, and Carol Friesen disclosed no conflicts of interest. Brian Rowe, Michael Schull, and Grant Innes are authors of several abstracts or manuscripts considered for inclusion in this report, but they were not involved in the search and selection of the included studies, and did not comment specifically on the content of their articles.
Measuring Overcrowding in Emergency Departments:  
A Call for Standardization

Issues and Methods
Although there is a growing concern about ED overcrowding, there is currently no consistent standard for measuring this phenomenon. It is also unclear what measures are important to Canadian administrators, ED providers, and researchers. A systematic review of published literature and a modified Delphi study (a consensus method) with 38 Canadian ED experts were conducted to examine these areas.

Implications for Decision Making
- **The percentage of the ED occupied by inpatients is perceived to be the most important measure of ED overcrowding.** This refers to the percentage of patients in the ED who have been admitted, but have not been transferred to a hospital ward because of a lack of bed availability. The use of consistent measures such as this may help with intra- and inter-institutional benchmarking, as well as in designing and implementing interventions to reduce overcrowding in Canadian EDs.
- **Other important measures are perceived to be total ED patients, total time in the ED, percentage of time that ED is at or above capacity, and overall bed occupancy.** The factors that are most important in academic, urban EDs may not be identical to those in suburban EDs.

The importance of some measures may vary according to local criteria, and even change over time. For instance, the extent of ambulance diversion may be a useful measure in a large, inner-city institution, but of no value to a regional hospital that is the only choice for ambulance personnel.
- **The least important measure is perceived to relate to staffing.** The number of ED nurses, attending emergency physicians, and staffed acute-care beds were not considered to be important measures of ED overcrowding.
- **The consensus opinion regarding the most important measures for reporting ED overcrowding did not always corroborate with the results of the systematic review.** From 169 studies examined in the systematic review, few of the measures for reporting ED overcrowding identified by the experts were found.
- **Consideration should be given to identifying the most meaningful and consistent measures.** This would ensure that data collected could be used to support decision-making about the seriousness of overcrowding and the development of standardized approaches to addressing the problem.

EXECUTIVE SUMMARY

The Issue

Emergency department (ED) overcrowding is a widely used term referring to a situation where the demand for ED services has exceeded capacity. There is no uniform definition of what overcrowding is or how it can best be measured. The many attributes of ED overcrowding, and the absence of well-defined methods for developing measures have frustrated agreement on appropriate indicators. As a result, attempts to study the causes, characteristics, and effects of ED overcrowding, and to develop effective solutions have been hindered.

Although there is a growing interest in measuring ED overcrowding, little is known about the type of measures that have been reported in the biomedical literature to document this problem. It is unclear what measures are important to researchers, ED providers, and administrators across Canada. This study represents an effort to systematically identify the valid and relevant indicators of ED overcrowding using an evidence-based approach (systematic review) and consensus methodology (Delphi technique).

This is the first in a series of four CADTH reports, which together provide a comprehensive assessment of ED overcrowding in Canada.

Objectives

The objective of the systematic review was to identify and characterize the measures and indicators that have been used to document ED overcrowding in the scientific literature. The objective of the Delphi study was to identify the level of consensus among a group of Canadian ED experts on the importance and relevance of measures to document ED overcrowding.

Methods

A systematic review and modified Delphi technique were used. Relevant published literature was obtained by searching multiple databases using a defined strategy. Studies had to report data for measures and indicators used to document events related to ED overcrowding. All types of study designs were considered. A qualitative synthesis of the results was performed.

A two-round modified Delphi study was conducted involving 38 Canadian experts in ED-related issues (ED physicians, nurses, and administrators). Participants were asked to rate a set of measures for documenting ED overcrowding, and to rank them according to their importance.

Results

From 486 potentially relevant studies, 169 studies were selected; 735 measures documenting ED overcrowding were identified (median number per study=3; interquartile range=2, 5). Operational definitions of ED overcrowding were infrequently reported (31%). Most measures focused on delays in the process of ED care (39.7%), overall volume of patients in the ED (11.6%), volume of patients waiting to receive care at different stages (8.8%), or the proportion of patients being seen at different stages while in the ED (7.6%). Measures such as ED access block (7%), ambulance diversion (7%), number of patients who left without being seen (5.5%), and ED length of stay (4.2%) were less commonly reported. When the measures were analyzed according to the input-throughput-output model of ED service pressures, throughput measures were the ones most commonly used (67.8%),
followed by input and output measures (19.5% and 11.8% respectively). System measures were reported less often (0.8%).

For the Delphi study, 32 participants (84% response rate) completed the questionnaire in the first round, and 33 participants (87% response rate) completed the second round. The most important measure identified was the percentage of the ED occupied by in-patients. The other top five measures, in order of importance, were total ED patients, total time in the ED, percentage of time ED is at or above capacity, and overall bed occupancy. The top 10 measures reported accord with measures that researchers in English-speaking countries outside North America have considered to be important for documenting ED overcrowding.

**Conclusions**

The evidence reveals the limitations of research on measuring overcrowding in EDs. Many measures and indicators have been used, and there is little agreement on the development of standardized definitions and measures that take into account regional variations and differences between EDs. The multitude of measures and their variable definitions highlight the complexity of this issue. The inconsistent use of definitions, indicators, and measures of ED overcrowding has created a contradictory research base.

This report describes a set of indicators that may be used to evaluate the problem of ED overcrowding, and thereby make it more consistently measurable across settings. Through the combination of a literature review and a Delphi study, 10 clinically important indicators were identified. The measures developed using consensus techniques have face validity, and those based on the rigorous collection of evidence possess content validity. These indicators should be tested for acceptability, feasibility, reliability, sensitivity to change, and validity to optimize their effectiveness for documenting ED overcrowding across Canada. Without a greater knowledge of the measurement properties of ED overcrowding indicators, study results will remain difficult to interpret, and consequently of limited value to policy makers, clinicians, and patients.
GLOSSARY

Access block: situation in which patients in the emergency department requiring in-patient care cannot gain access to appropriate hospital beds within a reasonable time frame

Ambulance diversion: rerouting of an ambulance(s) from the intended receiving facility to an alternative receiving facility, because of a temporary lack of critical resources in the emergency department of the intended receiving facility

Emergency department gridlock: simultaneous ambulance diversion at multiple emergency departments

Boarding of patients: situation in which patients in the emergency department requiring in-patient care cannot gain access to appropriate hospital beds

Critical care bypass: situation in which the hospital cannot admit even one more critically ill patient without compromising the care of patients already in the department; the emergency department is essentially closed to patients coming by ambulance

Delphi study: collaborative technique for building consensus, involving an iterative group process in which a central source forwards surveys or questionnaires to isolated, anonymous (to each other) participants, whose responses are summarized and recirculated to the participants in multiple rounds for further modification, producing a final group response

Input measures: measures related to the number of patients seeking ED care

Likert scale: ordinal rating scale that measures a respondent’s agreement with a clear statement

Output measures: measures of factors related to those processes that move patients out of the emergency department to other care areas or to discharge

Redirect consideration status: situation where no ambulances are accepted except those with critically ill patients

Throughput measures: measures of the efficiency and capacity of the emergency department system to admit and treat patients requiring emergency care
# ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAEP</td>
<td>Canadian Association of Emergency Physicians</td>
</tr>
<tr>
<td>CI</td>
<td>confidence interval</td>
</tr>
<tr>
<td>CTAS</td>
<td>Canadian Triage and Acuity Scale</td>
</tr>
<tr>
<td>ED</td>
<td>emergency department</td>
</tr>
<tr>
<td>EIP</td>
<td>emergency department occupied by in-patients</td>
</tr>
<tr>
<td>EP</td>
<td>emergency practitioner</td>
</tr>
<tr>
<td>ICU</td>
<td>intensive care unit</td>
</tr>
<tr>
<td>IQR</td>
<td>interquartile range</td>
</tr>
<tr>
<td>LWBS</td>
<td>left or leaving without being seen</td>
</tr>
<tr>
<td>LAMA</td>
<td>left against medical advice</td>
</tr>
<tr>
<td>NENA</td>
<td>National Emergency Nurses Affiliation</td>
</tr>
<tr>
<td>SD</td>
<td>standard deviation</td>
</tr>
<tr>
<td>TEP</td>
<td>technical expert panel</td>
</tr>
</tbody>
</table>
# TABLE OF CONTENTS

**EXECUTIVE SUMMARY** .................................................................................................................................................. iv

**GLOSSARY** ........................................................................................................................................................................ vi

**ABBREVIATIONS** ............................................................................................................................................................. vii

1 **INTRODUCTION** ............................................................................................................................................................. 1
   1.1 Background ................................................................................................................................................................... 1

2 **THE ISSUE** ......................................................................................................................................................................... 1

3 **OBJECTIVES** ...................................................................................................................................................................... 2

4 **SYSTEMATIC REVIEW** ....................................................................................................................................................... 3
   4.1 Methods ........................................................................................................................................................................ 3
      4.1.1 Literature search strategy ................................................................................................................................. 3
      4.1.2 Selection criteria and method .......................................................................................................................... 4
      4.1.3 Data extraction strategy ...................................................................................................................................... 4
      4.1.4 Strategy for quality assessment ...................................................................................................................... 5
      4.1.5 Data analysis methods ....................................................................................................................................... 5
   4.2 Results ............................................................................................................................................................................. 5
      4.2.1 Quantity of research available ........................................................................................................................ 5
      4.2.2 Study characteristics ....................................................................................................................................... 7
      4.2.3 Data analyses and synthesis .......................................................................................................................... 7
   4.3 Discussion ....................................................................................................................................................................... 8

5 **DELPHI STUDY** ............................................................................................................................................................ 10
   5.1 Methods .......................................................................................................................................................................... 10
      5.1.1 Study design ......................................................................................................................................................... 10
      5.1.2 Study participants ............................................................................................................................................... 10
      5.1.3 Development of questionnaires ...................................................................................................................... 10
      5.1.4 Study procedures ............................................................................................................................................... 11
      5.1.5 Data analysis ....................................................................................................................................................... 13
      5.1.6 Ethics .................................................................................................................................................................... 13
   5.2 Results ............................................................................................................................................................................. 13
      5.2.1 Round 1 ............................................................................................................................................................... 13
      5.2.2 Round 2 ............................................................................................................................................................... 13
   5.3 Discussion ....................................................................................................................................................................... 14

6 **DISCUSSION** ................................................................................................................................................................. 17
   6.1 Summary of Results ....................................................................................................................................................... 17
      6.1.1 Systematic review .............................................................................................................................................. 17
      6.1.2 Delphi study ......................................................................................................................................................... 18
   6.2 Study Limitations ............................................................................................................................................................ 18
   6.3 Generalizability of Findings ........................................................................................................................................ 20
   6.4 Health Services Impact ................................................................................................................................................. 20
   6.5 Knowledge Gaps ............................................................................................................................................................ 20
APPENDICES - available from CADTH's web site www.cadth.ca

APPENDIX 1: Literature Search Strategies
APPENDIX 2: Electronic Databases Searched
APPENDIX 3: Form for Inclusion and Exclusion of Studies
APPENDIX 4: Data Extraction Form
APPENDIX 5: Excluded Articles and Reason for Exclusion
APPENDIX 6: Articles Awaiting Assessment
APPENDIX 7: Characteristics of Included Studies in Systematic Review of Measures to Document Overcrowding in ED
APPENDIX 8: Classification of Studies according to Variables under Study
APPENDIX 9: First-round Delphi Questionnaire
APPENDIX 10: Second-round Delphi Questionnaire
1 INTRODUCTION

1.1 Background

Overcrowding in the emergency department (ED) has been difficult to define scientifically, even though it is a growing concern in Canada and other countries. Emergency health care providers and hospital administrators may have a sense of when an ED is becoming overcrowded, but there is no consensus on a definition. A systematic review of definitions of ED overcrowding found that they vary in content and focus, with 43% of the studies stating one explicitly. The authors concluded that although one definition may be unsuitable for all situations, a more consistent approach that focuses on standardized criteria, and uses operational definitions and measures of events occurring in the ED would distinguish among the causes, characteristics, and outcomes of overcrowding.

Several publications have tried to develop standard definitions of ED overcrowding, and a list of proxy measures or indicators. These have not been widely accepted, because of the multiple attributes of overcrowding, and the absence of well defined methods to develop appropriate indicators.

Some authors have called for the use of measures and indicators of ED overcrowding that are valid, reliable, and sensitive to changes through time. Indicators are defined here as measurable items referring to an attribute related to system structures, processes, or outcomes of ED overcrowding. Indicators can also be objective surrogate markers that represent overcrowding.

Isolated indicators do not provide a comprehensive assessment of ED overcrowding; it is necessary to presuppose a conceptual model to better understand how measures relate to one another. An input-throughput-output conceptual model is one way of grouping the ED overcrowding measures (Figure 1). This model is based on engineering principles, queuing theory, and compartmental models of patient flow, allowing most factors related to overcrowding to be put into one of three interdependent categories: input, throughput, and output. Input refers to reproducible measures of the number of patients seeking ED care. Throughput refers to factors related to ED efficiency, workload, and capacity. Finally, output includes measurement of the efficiency and capacity of the in-patient system to admit patients requiring hospital care, and of the ambulatory care system to provide timely care after discharge.

This model allows one to classify the measures that have been used to document ED overcrowding in the medical literature. It was chosen as the framework for this systematic review, because it focuses on the measurement of factors related to overcrowding from the ED’s perspective. The model also highlights the multidimensional measurement problem, and helps to clarify those components that may be alleviated by changes to ED processes.

2 THE ISSUE

Emergency department (ED) overcrowding is a growing concern in many countries, and a frequent topic in the biomedical literature. “ED overcrowding” is a term that is used by academic institutions, professional bodies, hospital administrations, and funding organizations to identify a situation where
the demand for ED services has exceeded capacity. The precise meaning of ED overcrowding varies with the context in which it is used. Many researchers\textsuperscript{1,7,8,10} have noted the lack of a uniform definition, and the difficulty that this has created in studying the causes, characteristics, and effects of ED overcrowding, and in developing effective solutions.\textsuperscript{11}

Although there is increasing interest in measuring the factors associated with ED overcrowding, little is known about the types of measures that have been reported in the biomedical literature to document this problem. It is unclear what measures are important to researchers, ED providers, and administrators across Canada. As a result, the systematic identification of measures is the first step in understanding which ones contribute the most to overcrowding. Identifying these measures by consensus among experts in the field would ensure that these meaningful measures would help develop standardized approaches, and allow for comparisons among EDs of different capacities. The involvement of ED experts in identifying specific indicators to document overcrowding may increase the face validity and content validity of the measures identified in the literature review.

3 OBJECTIVES

The objective of the systematic review was to identify measures and indicators that have been used to document ED overcrowding in the scientific literature, and to identify the contexts in which the term has been used. The secondary objectives were to describe the characteristics of these measures in terms of the input-throughput-output model, and to report on the purposes of their use in the context of research on ED overcrowding.
The objective of the Delphi study was to identify the level of consensus among a group of ED experts across Canada on the importance and relevance of a set of measures to document ED overcrowding. A secondary objective of the study was to develop an inventory of potential and clinically relevant common measures of ED overcrowding, based on expert opinion.

4 SYSTEMATIC REVIEW

4.1 Methods

A protocol for the study was written a priori, and followed throughout the review process.

4.1.1 Literature search strategy

a) Searches of electronic databases

A comprehensive search was undertaken by the research librarian (CF), who with the review team, identified the relevant electronic databases, and developed the search strategies based on the core search terms for “overcrowding” and “emergency departments.” Searches were conducted in 24 electronic databases with the use of appropriate subject headings, and the extensive use of keywords, which was crucial for this poorly indexed topic. The search strategies were designed for comprehensiveness using indexing terms and free-text searching terms after consulting the Technical Expert Panel (TEP), a group of individuals from across Canada with extensive expertise in ED administration, clinical care, nursing, and research; and review team members. Databases that focus on grey literature, such as SIGLE, GrayLit Network, Dissertation Abstracts, and the NLM Gateway were included in the search process. Google and other Internet search engines (Dogpile and Copernic Meta) were used to identify web-based information. The original searches were performed in November 2004, with two updates using additional terms conducted one month later. The search strategies and results are included in Appendix 1, and the list of databases searched appears in Appendix 2. No restrictions on publication status, language, or year of publication were applied in the searches.

b) Manual and grey literature searches

The electronic database searches were supplemented by manual searches, and other searches for grey literature by review team members. Abstracts presented to the scientific meetings of the Canadian Association of Emergency Physicians (CAEP) from 2000 to 2004, and to the Society of Academic Emergency Medicine (SAEM) annual scientific meeting from 1999 to 2004 were screened for relevance.

Reference lists and bibliographies of relevant papers and books were hand-searched for additional citations. Efforts were made to obtain information about unpublished studies through consultations with the TEP. To limit publication bias, dissertations (beyond those retrieved from the electronic database searches) were identified by searching grey literature. These dissertations were used to find related literature.
4.1.2 Selection criteria and method

a) Selection criteria

Topic
Studies were required to refer to ED overcrowding as the main objective or at least as a prominent feature. A formal statement that included the terms “crowding,” “overcrowding,” or other synonyms (e.g., increased or increasing patient volumes, increasing number of visits, increasing ED census, ED congestion, or high demand of ED services) was required in the title, introduction, or methods sections.

Study design
Randomized controlled trials, quasi-randomized trials, before-and-after studies, cohort studies, case-control studies, cross-sectional studies, qualitative research or Delphi studies, and computer simulation studies using measures related to ED overcrowding were considered.

Outcomes
Studies were required to report data for measures or indicators used to document events related to ED overcrowding.

Articles that did not meet all the inclusion criteria were excluded. Only original research was eligible. Systematic reviews, health technology assessment reports, review articles, editorials, opinion letters to the editor, commentaries, and case studies were also excluded. If a study resulted in more than one report (i.e., conference abstract and published manuscript), only the main publication was considered for inclusion.

b) Selection method

Because of the large number of records retrieved from the computerized database, the selection of studies involved a three-stage process. Two reviewers (HL, KB) independently pre-screened study titles to exclude irrelevant citations. Four reviewers (CS, HL, KB, and MBO) independently inspected the titles, subtitles, abstracts, and keywords to select potentially relevant references. Each citation in stage 2 was considered by at least two reviewers. The full texts of the provisionally included articles from the second stage were retrieved. Six reviewers (CF, CS, HL, KTB, MBO, and NH) independently decided to include or exclude studies using a standard form based on the selection criteria (Appendix 3). Each citation in stage 3 was considered by at least two reviewers.

The level of agreement among reviewers at all stages of the selection process was evaluated using kappa (κ) statistics. A κ score in the range of 0.0 to 0.40 was considered to be poor agreement, 0.41 to 0.60 was moderate agreement, and 0.61 to 0.80 represented substantial agreement. Disagreements about inclusion or exclusion of studies were initially resolved by consensus, and when this was impossible, through arbitration by a third reviewer (BHR).

4.1.3 Data extraction strategy

One reviewer (MBO) extracted data using a structured electronic form (Appendix 4). A profile of each study was developed by the extraction of first author, year of publication, country, study design, number of study centres, definition of ED overcrowding, measures or indicators of ED overcrowding, and definitions of the measures. A second reviewer (KB) examined a 10% random
sample of the studies to check the accuracy of the abstracted information. Disagreements were resolved by consensus.

4.1.4 Strategy for quality assessment

It was decided a priori not to conduct a methodological quality assessment of the studies. As the goal of this review was to provide a comprehensive overview of the measures and indicators that have been used to document ED overcrowding in the biomedical literature, a quality assessment of the included studies was considered to be irrelevant.

4.1.5 Data analysis methods

After collection, the overcrowding measures were classified according to the input-throughput-output conceptual model of ED overcrowding.\textsuperscript{8} The measures were categorized according to their purpose (i.e., discriminative or descriptive, predictive, or evaluative). Discriminative measures distinguish individuals or groups using an underlying variable. For example, the number of episodes of ambulance diversion can be used to describe the frequency of overcrowding among EDs. Predictive measures classify individuals or groups into predefined categories, either concurrently or prospectively, to determine whether they predict a particular event. For example, the number of patients seen in the ED can be considered as a predictive factor associated with increased waiting times in the ED. Evaluative measures quantify the magnitude of longitudinal change in an individual or group in the dimension of interest as a result of an intervention. For example, a measure of waiting times can be used to investigate the effect of a fast track scheme implemented in the ED.

The constant comparative method\textsuperscript{14} was used to categorize the measures according to their content across the studies. This is an iterative process of analyzing qualitative data (i.e., text) involving generative and descriptive coding. Using this method, units of text were labelled, compared, and grouped until no new categories emerged. Generative coding involves the development of categories emerging from the data without making prior assumptions about the nature of the categories. Descriptive coding links measures under study to concrete concepts such as actions, events, properties, settings, and processes.

Two reviewers (MBO, KB) coded the measures, and then compared results. Disagreements were resolved by consensus. The results were considered to be unsuitable for formal meta-analysis, because of the multiple outcome measures that were reported in the studies. As a result, a qualitative synthesis of the data including descriptive statistics on the characteristics of ED overcrowding measures was undertaken.

4.2 Results

4.2.1 Quantity of research available

The computerized search identified 15,357 citations. After removing duplicate and irrelevant references, 1,259 potentially relevant studies remained. A second screening of titles and abstracts yielded 632 studies (487 and 145 references from electronic searches and the grey literature respectively) to be retrieved for further review. After application of the selection criteria, 452 studies were excluded, and 169 were included in the review (Figure 2).
There was 78% agreement on included studies [weighted κ: 0.63; (95% confidence interval (CI): 0.58, 0.68); n=619 studies]. Among the excluded studies, 214 were not primary research, 194 were unrelated to ED overcrowding, 20 were conference abstracts subsequently published as full-text manuscripts, 21 did not provide measures of ED overcrowding, and three studies did not meet the study design criteria for inclusion in the review (Appendix 5). Eleven studies are awaiting assessment: two articles are pending translation into English, and nine are pending retrieval (Appendix 6). Because of delays in access to and translation of articles published in languages other than English, the authors could not make a decision about their relevance in this review.
4.2.2 Study characteristics

The 169 eligible studies reported data on 735 measures to document ED overcrowding. The number of measures reported per study varied from 1 to 28 [median: 3; interquartile range (IQR): 2, 5]. The studies were published between 1973 and 2004.

Most of the studies were conducted in the US (79), Canada (36), Australia (20), the UK (12), and Spain (nine). A few studies were performed in other countries, such as Hong Kong, Sweden, Taiwan (two studies each), and Greece, Israel, New Zealand, Pakistan, Singapore, Switzerland, and Turkey (one study each). There were 124 single-centre studies, and 45 that involved >1 centre. Among the studies, 52 used a before-and-after design, 80 used a prospective (41) or retrospective (39) design, and 18 used a cross-sectional design. Other designs included the use of controlled trials (seven), computer simulation models (eight), and quantitative or Delphi studies (two). A randomized controlled trial design was used in two studies. A description of the characteristics of the included studies and their associated definitions and measures appears in Appendices 7 and 8.

4.2.3 Data analyses and synthesis

Few authors defined ED overcrowding. Of 169 studies, 52 (31%) provided a clear definition of the problem.

a) Category of ED overcrowding measures

Each measure was placed in one of three interdependent categories from the input-throughput-output model. Throughput variables were the most commonly used measures (499, 67.8%), followed by input (143, 19.5%), and output (87, 11.8%). Six measures (0.8%) were used to document aspects related to overall ED overcrowding, and did not fit into the model; these measures were classified as system measures.

b) Purpose of measures

The measures reported in the studies served a variety of purposes. More than half (397, 54%) were used to describe the frequency of an underlying dimension related to overcrowding, or to discriminate between people with different levels of an attribute. Over a quarter (204, 27.8%) were used to document the magnitude of longitudinal change in individuals or groups as an effect of an intervention to reduce or control ED overcrowding. Finally, 134 (18.2%) measures were used to classify individuals, concurrently or prospectively, to assess the likelihood of a particular event related to ED overcrowding.

c) Content of measures

After an iterative qualitative analysis of the measures, 11 content categories emerged from the data. ED times was a category describing a heterogeneous group of 292 measures (39.7%) reported in the studies:

- overall times (e.g., total times in the ED)
- times needed for various steps in ED care (e.g., time from registration to consultation with an ED doctor)
- waiting times for a particular action or event in the ED (e.g., time waiting to be seen).

Uniform time intervals were not used, and given the heterogeneity of the metrics, secondary categories could not be created.
Other categories included the overall volume of ED patients (85 measures, 11.6%), and variables such as the total number of ED visits within a given interval, ED census, number of patients registered, and number of ambulances that arrived at a certain time.

Another identified measure was patient waiting volumes (65, 8.8%). It referred to those measures that described the number of patients in the queue at different stages of receiving care (e.g., proportion of patients waiting before being seen by a doctor, maximum queue of patients waiting for examination, and proportion of patients waiting <60 minutes to see a doctor).

The category of ED administration included measures (59, 8.0%) that described aspects related to ED capacity and resources (e.g., ED workload, ratio of ED beds to ED patients), staff-related issues (e.g., providers’ ratings of ED demand or crowding, ratio of physicians and nurses to waiting room patients), ED management decisions (e.g., temporary ED closures, forced openings), and measures related to quality improvement (e.g., number of times the total number of patients in the ED equalled or exceeded the bed capacity, and proportion of patients waiting less than their triage time threshold).

The category of ED patient volumes (56, 7.6%) referred to those measures describing the proportion of patients seen at different stages in their care (e.g., proportion of patients seen within a threshold time, rate of stretcher occupancy, and number of patients under evaluation).

Measures categorized as access block (52, 7%) referred to a situation where patients in the ED requiring in-patient care could not gain access to appropriate hospital beds within a reasonable time frame (e.g., number of admitted patients boarded in the ED, delay from ready for departure to departure >4 hours, and percentage of ED admissions waiting >4 hours).

There were 52 measures (7%) categorized as ambulance diversion measures. These were time or volume measures related to the decision to redirect incoming ambulance traffic when an ED had reached a level of crowding that triggered diversion (e.g., number of hours on diversion, percentage of time per month spent on redirect consideration status or critical care bypass, gridlock time, percentage of EDs that were diverting ambulances at patient registration).

Another category of 40 measures (5.5%) contained variables related to the number of patients who left the ED without being seen (LWBS). A subgroup (three measures, 0.4%) comprised those patients who were seen but left the ED against medical advice (LAMA). The category of ED length of stay (LOS) (31, 4.2%) grouped those measures related to the length of time from patient triage or registration to the time of patient discharge.

### 4.3 Discussion

This systematic review aimed to identify the measures and indicators of ED overcrowding that have been used in the scientific literature. From 169 studies, 735 measures documenting ED overcrowding were identified. This shows that it is common for authors to use a heterogeneous set of variables to investigate events related to overcrowding. More than half of the measures (397 measures, 54.1%) were used to describe the frequency of an underlying dimension of ED overcrowding, or to discriminate between people with different levels of a particular attribute related to ED overcrowding. Examples of the latter would be patients with different levels of acuity, or physicians experiencing varying amounts of stress.
The different measures and the variation in definitions for similar measures highlight the complexity of this topic. The measures provided information about specific events related to overcrowding, but do not provide a clear definition of the phenomena.

Although some measures are conceptually similar and may be highly correlated, they are not necessarily synonymous. The studies were inconsistent in selecting variables to measure overcrowding for many reasons. Inconsistencies may originate from confusion between factors defining overcrowding, possible causes, and effects. Variation in the ways to measure overcrowding in ED research may preclude generalizations. Another source of variation may be the fact that the most obvious aspects of overcrowding change over time with the evolution of the acute care system, volume of demand, and resource allocation, so that the most salient aspects of overcrowding at a particular time do not necessarily remain so.

The measures used may depend on where the studies were performed. For example, ambulance diversion may be a useful measure of ED overcrowding in a large inner-city institution where multiple alternatives exist, yet of no value to a regional hospital that is the only choice for ambulance personnel. Measures that are considered to be crucial in one ED may not be important in others.15

While there are many theoretical models to describe ED overcrowding, the input-throughput-output model was used, because it focuses on measures of overcrowding from the ED’s perspective, and allows for a comprehensive analysis. When the measures were analyzed according to this model, the throughput measures were the ones that were most commonly used to document ED overcrowding (499, 67.8%), followed by input (143, 19.5%), and output (87, 11.8%). This distribution may reflect the growing interest in gaining a detailed understanding of internal, controllable factors related to overcrowding, instead of focusing on external, less controllable factors. The emphasis on measuring throughput variables to define overcrowding may not be equally important for patients, staff, administrators, and policy makers; how much emphasis should be placed on each group of measures, whether internal to the hospital or external, remains unclear.

Of the studies, 31% provided explicit definitions of ED overcrowding. This is lower than the 43% reported in a previous systematic review.1 The discrepancy may be related to variations in the eligibility criteria for studies (e.g., inclusion of synonyms of ED overcrowding), and a difference in the denominator used for calculating these estimates. Hwang et al.1 included 91 studies, but based their calculations on 53 studies classified as primarily about ED crowding or overcrowding. If the total number of studies (i.e., 91) were the denominator, the proportion of studies providing explicit definitions of ED overcrowding would be 25%, a proportion closer to that reported here.

This systematic review is one step in the development of comprehensive indicators to help evaluate and resolve the problem of ED overcrowding. The review satisfied the criteria for a rigorous systematic review. A clear research question was posed a priori, and a comprehensive search of the literature was performed. The study selection process was robust. As authors of similar reviews1 have noted, the eligibility criteria for studies related to ED overcrowding may be criticized because defining overcrowding is a problem, regardless of other issues and attempts to control for bias in the selection of studies.

The evidence provided here highlights the limitations of research on the measurement of ED overcrowding. Many measures and indicators have been used, and there has been little agreement on how to develop standardized definitions and measures that account for regional variations and differences among individual EDs. The inconsistent use of definitions, indicators, and measures of...
ED overcrowding has created a contradictory research base. More investigations are required to provide a greater understanding of measurement in this area. Without a greater knowledge of the operating characteristics and properties of measures used for ED overcrowding, results will remain difficult to interpret, and consequently, of limited value to policy makers, clinicians, and patients.

These findings point to significant gaps in the measurement of ED overcrowding in the scientific literature, and suggest future directions for study. Researchers need to identify measures that are feasible, relevant, reliable, valid, and sensitive to change over time. There is an urgent need for rigorous assessment of the measures being used in ED overcrowding research. Such an assessment will optimize their effectiveness, applicability, and generalized adoption by clinicians, administrators, policy makers, and researchers in the ED overcrowding field.

5 DELPHI STUDY

5.1 Methods

5.1.1 Study design

A two-round modified Delphi study was conducted from February to April 2005. The study followed a protocol that was developed a priori, and shared with the members of the TEP.

The Delphi technique is a research tool developed to deal with complex problems that involve a measure of uncertainty that cannot be overcome by statistical methods or deliberation. Its goal is to obtain the most reliable consensus among a group of experts on a particular topic under evaluation. The technique involves recruiting a group of experts to participate in an iterative process of answering sequential questionnaires, receiving feedback regarding group responses, and revising their opinions in light of this feedback. The four distinguishing features of the Delphi technique in this study were: anonymity, iteration (the procedure involves at least one round), controlled feedback (the results of each round are analyzed separately and responses fed back to members of the Delphi panel), and statistical group response (expression of the degree of consensus of the group). The Delphi technique has been used to develop or to identify indicators and measures in a variety of biomedical areas (Figure 3).

5.1.2 Study participants

Participants were individuals considered to be experts in an aspect of ED operations. For this study, an expert was any Canadian ED administrator, director, physician, or nurse with a known or stated interest in ED overcrowding. A sample of 38 Canadian ED experts was recruited based on recommendations from the TEP.

5.1.3 Development of questionnaires

The selection of the measures to be included in the Delphi study was independent of the systematic review process. An initial list of potential measures to document ED overcrowding was generated from a preliminary list of key articles relevant to the topic that were available before the systematic review was formally conducted. An indicator or measure of overcrowding was defined as something
that changes as a result of changes in ED overcrowding. For example, waiting time, a widely used measure of ED overcrowding, will increase as the ED becomes less able to meet the demands placed on it by an increasing influx of patients.

The list of potentially relevant measures of ED overcrowding with their operational definitions were refined through an iterative review process with the TEP, until a set of 36 measures was reached. These served as the basis for the first-round questionnaire. The measures were classified into patient volumes (six), ED times (13), diversion status (two), ED staffing (five), ED administration (eight) and staff perceptions of ED overcrowding (two).

The first-round questionnaire (Appendix 9) consisted of 36 measures. Participants were asked to rate the importance of each measure for documenting ED overcrowding on a seven-point Likert scale (1=unimportant in all cases, and 7=important in all cases), and to indicate whether the measures were used in their EDs. Demographic and institutional data were also collected.

A pilot survey testing the length of time to complete the questionnaire and its comprehensiveness was sent to a convenient sample of 11 emergency physicians. Pilot respondents were also asked to include any measures that they considered important but were excluded. Of 11 participants, eight completed the pilot survey, and no substantive changes were made. The responses of the pilot participants were not part of the Delphi study.

A second-round questionnaire (Appendix 10) was developed after data from the first round were analyzed. The second round consisted of 10 measures selected from the first-round survey, based on the mean group rating for each measure (mean of ≥6). As the purpose of the Delphi study was to reach a consensus, a measure of what that would be had to be considered. A priori values chosen for the mean score and standard deviation were used as consensus indicators. The mean score is an accurate reflection of the level of agreement, while the standard deviation measures the degree of consensus. In the presence of skewed data, the median and interquartile range (IQR) would be used to select the measures for round 2.18 Time constraints meant that only two rounds could be completed.

Participants in round 2 were provided with the respondents’ first-round ratings, and the mean group rating for each measure. This offered the respondents an opportunity to amend their ratings in light of the opinion expressed by the group. Participants were asked to rank from one to 10, the 10 measures according to their relative importance in documenting ED overcrowding (1=the most important measure, 10=the least important measure).

5.1.4 Study procedures

The TEP identified 38 potential participants in seven provinces in Canada, and one Canadian expert employed in the US. Potential participants were individually contacted by e-mail. They received a personalized letter describing the Delphi process, and the expectations regarding their participation. The survey questionnaire was sent electronically. Non-respondents received three reminders that were sent at two-week intervals. The method used to deliver the second-round survey was the same as that for the first. The second-round questionnaire was sent to all individuals regardless of their participation in the first round.
Measuring Overcrowding in Emergency Departments: A Call for Standardization

Figure 3: Delphi study

- Definition of problem
- Selection of participants (N=38)
  - First round of Delphi (N=32)
  - Second round of Delphi (N=33)
  - Results analyzed for degree of consensus
  - Questionnaire development (round 1)
    - TEP selected participants
    - Pilot survey for feasibility
    - Rating of importance of measures to document ED overcrowding
    - Data analysis for agreement
    - Questionnaire development (round 2)
      - Relative ranking for top 10 measures
      - Data analysis for agreement

Report findings
5.1.5 Data analysis

Data from the electronic questionnaires were captured in a database, and exported to a statistical package (SAS version 6, SAS Institute, Carey NC) for analysis. Proportions and percentages with 95% CI were reported for dichotomous and categorical data. Data from Likert-type scales were treated as continuous data, and reported as means with standard deviations (SDs), or when appropriate, medians with IQR.

5.1.6 Ethics

This study was reviewed and approved by the University of Alberta Health Research Ethics Board. Completion and return of the survey were considered consent to participate. Data from participants were coded, and anonymity was ensured when reporting the results for the group.

5.2 Results

Of the 38 participants who agreed to participate, 32 (84% response rate) completed the first-round questionnaire, and 33 (87% response rate) completed the second. There were 21 males and 10 females (n=31). Not all participants answered each question. Respondents ranged in age from 35 to 54 years with a mean age of 44 (95% CI: 42.5; 46.0). Of the respondents (n=29), 16 were emergency physicians, seven were director or physicians, four were nurses or researchers, and two were administrators. Many (29) reported affiliation with a university or university-affiliated hospital. The annual ED census of their institutions ranged from 25,000 patients to 210,000 patients (median: 60,000; IQR: 50,000, 70,000).

5.2.1 Round 1

The median value of the group ratings was chosen as an indicator of the level of agreement on the top 10 measures to document ED overcrowding (Table 1). Four measures had a median value of 7 (percentage of ED occupied by in-patients, total ED patients, total time in the ED, and percentage of time the ED was at or above its capacity), one had a median of 6.5 (overall bed occupancy), and five had a median of 6 [time from bed request to bed assignment, time from triage to emergency practitioner (EP), MD satisfaction, time from bed ready to transfer to ward, and number of staffed acute-care beds].

5.2.2 Round 2

Of the 33 respondents, 16 ranked the percentage of ED occupied by in-patients as the most important measure to document ED overcrowding from the top 10 measures selected from round 1 (Figure 4). In terms of least importance, 12 of the 31 ranked the number of staffed acute-care beds as the least important indicator of ED overcrowding.

Of the 10 measures, seven were throughput, two were output, and one was system-related. The Delphi group did not select any input measures as important indicators of ED overcrowding. The participants in the Delphi survey were unaware of the use of the input-throughput-output model in the analysis of the results.
5.3 Discussion

The objective of this study was to develop an inventory of potential and clinically relevant measures of ED overcrowding based on expert opinion. By means of a consensus rating process involving a group of Canadian experts in ED-related issues, this study has identified and ranked 10 measures that are considered to be important for the reporting of ED overcrowding in Canada. The most important measure identified was the percentage of the ED occupied by in-patients.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Mean</th>
<th>SD</th>
<th>Median</th>
<th>IQR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total ED patients</td>
<td>6.35</td>
<td>0.75</td>
<td>7</td>
<td>6,7</td>
</tr>
<tr>
<td>Percentage of ED occupied by in-patients</td>
<td>6.53</td>
<td>0.80</td>
<td>7</td>
<td>6,7</td>
</tr>
<tr>
<td>Overall bed occupancy</td>
<td>6.19</td>
<td>0.93</td>
<td>6.5</td>
<td>5,7</td>
</tr>
<tr>
<td>Total time in ED</td>
<td>6.16</td>
<td>1.04</td>
<td>7</td>
<td>5,7</td>
</tr>
<tr>
<td>Percentage of time ED at or above stated capacity</td>
<td>6.16</td>
<td>1.08</td>
<td>7</td>
<td>5,7</td>
</tr>
<tr>
<td>Time from bed request to bed assignment</td>
<td>6.06</td>
<td>1.08</td>
<td>6</td>
<td>5,7</td>
</tr>
<tr>
<td>Number of staffed acute-care beds</td>
<td>5.53</td>
<td>1.57</td>
<td>6</td>
<td>5,7</td>
</tr>
<tr>
<td>Time from triage to EP</td>
<td>5.84</td>
<td>1.08</td>
<td>6</td>
<td>5,7</td>
</tr>
<tr>
<td>Time from bed ready to transfer to ward</td>
<td>5.53</td>
<td>1.72</td>
<td>6</td>
<td>5,7</td>
</tr>
<tr>
<td>MD satisfaction</td>
<td>5.84</td>
<td>1.22</td>
<td>6</td>
<td>5,7</td>
</tr>
<tr>
<td>Time from triage to placement</td>
<td>5.75</td>
<td>1.08</td>
<td>5.5</td>
<td>5,7</td>
</tr>
<tr>
<td>Nurse satisfaction</td>
<td>5.75</td>
<td>1.24</td>
<td>5.5</td>
<td>5,7</td>
</tr>
<tr>
<td>Admission proportions</td>
<td>5.28</td>
<td>1.46</td>
<td>5</td>
<td>5,6,7</td>
</tr>
<tr>
<td>Time from waiting room to patient care area in ED</td>
<td>5.34</td>
<td>1.54</td>
<td>5</td>
<td>4,7</td>
</tr>
<tr>
<td>Patients in waiting room</td>
<td>5.38</td>
<td>1.52</td>
<td>5</td>
<td>5.7</td>
</tr>
<tr>
<td>Longest time in ED for admitted patient since admission</td>
<td>5.19</td>
<td>1.53</td>
<td>5</td>
<td>4.5,6.5</td>
</tr>
<tr>
<td>Patients in triage</td>
<td>5.03</td>
<td>1.86</td>
<td>5</td>
<td>3,7</td>
</tr>
<tr>
<td>Time from triage to bedside nurse</td>
<td>5.47</td>
<td>1.08</td>
<td>5</td>
<td>5,6</td>
</tr>
<tr>
<td>LWBS</td>
<td>5.16</td>
<td>1.51</td>
<td>5</td>
<td>5,6</td>
</tr>
<tr>
<td>Time from consult to disposition decision</td>
<td>5.22</td>
<td>1.43</td>
<td>5</td>
<td>4.5,7</td>
</tr>
<tr>
<td>CTAS categories 2 and 3</td>
<td>5.19</td>
<td>1.47</td>
<td>5</td>
<td>4,7</td>
</tr>
<tr>
<td>Number of hours out of last 24 on diversion</td>
<td>5.17</td>
<td>1.58</td>
<td>5</td>
<td>5,6</td>
</tr>
<tr>
<td>Average and range of patients per hour seen by EP</td>
<td>4.94</td>
<td>1.46</td>
<td>5</td>
<td>4,6</td>
</tr>
<tr>
<td>Percent of time on diversion</td>
<td>5.03</td>
<td>1.71</td>
<td>5</td>
<td>5,7</td>
</tr>
<tr>
<td>Total ED volume</td>
<td>4.75</td>
<td>2.02</td>
<td>5</td>
<td>3,7</td>
</tr>
<tr>
<td>Bed ratio</td>
<td>5.31</td>
<td>1.00</td>
<td>5</td>
<td>5,6</td>
</tr>
<tr>
<td>Time from EP assessment to disposition</td>
<td>5.00</td>
<td>1.27</td>
<td>5</td>
<td>4,6</td>
</tr>
<tr>
<td>Time from laboratory order to laboratory result returned</td>
<td>5.16</td>
<td>1.32</td>
<td>5</td>
<td>5,6</td>
</tr>
<tr>
<td>Longest time in ED since registration</td>
<td>4.66</td>
<td>1.52</td>
<td>5</td>
<td>3,6</td>
</tr>
<tr>
<td>Number of ED nurses</td>
<td>4.22</td>
<td>1.95</td>
<td>4</td>
<td>3,6</td>
</tr>
<tr>
<td>Time from diagnostic imaging order to receipt of result in patient care area</td>
<td>5.03</td>
<td>1.28</td>
<td>5</td>
<td>4.5,5.5</td>
</tr>
<tr>
<td>Number of attending ED physicians</td>
<td>4.28</td>
<td>1.85</td>
<td>4</td>
<td>3,5.5</td>
</tr>
<tr>
<td>Provider ratio</td>
<td>4.81</td>
<td>1.31</td>
<td>5</td>
<td>4.5</td>
</tr>
<tr>
<td>Hours of physician coverage</td>
<td>4.31</td>
<td>1.60</td>
<td>5</td>
<td>3,5</td>
</tr>
<tr>
<td>CTAS categories 4 and 5</td>
<td>4.13</td>
<td>1.68</td>
<td>4</td>
<td>3,5</td>
</tr>
<tr>
<td>LAMA</td>
<td>3.88</td>
<td>1.74</td>
<td>4</td>
<td>3,5</td>
</tr>
</tbody>
</table>
This factor, which has been called access block, is an area of emerging research activity. Access block is a term applied to a situation in which patients requiring emergency hospitalization stay in the ED for extended periods before being transferred to an in-patient bed. This has been described as emergency in-patients (EIP) or “being boarded in the ED.”

Total ED patients, which must be interpreted with caution, is related to the individual ED under consideration. For example, with 50 patients in a 30-bed ED, there is overcrowding, whereas 50 patients in a 75-bed ED may not be overcrowding.

**Figure 4: Results of round 2 of Delphi study**

<table>
<thead>
<tr>
<th>Assigned Rank</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of ED occupied by in-patients</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
</tr>
<tr>
<td>Total ED patients</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
</tr>
<tr>
<td>Total time in ED</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
</tr>
<tr>
<td>Percentage of time ED at or above capacity</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
</tr>
<tr>
<td>Overall bed occupancy</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
</tr>
<tr>
<td>Time from bed request to bed assignment</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
</tr>
<tr>
<td>Time from triage to EP</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
</tr>
<tr>
<td>MD satisfaction</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
</tr>
<tr>
<td>Time from bed ready to transfer to ward</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
</tr>
<tr>
<td>Number of staffed acute-care beds</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
</tr>
</tbody>
</table>

Rank ordering of measures selected by Delphi group to document ED overcrowding was 1) percentage of ED occupied by in-patients (output); 2) total ED patients (throughput); 3) total time in ED (throughput); 4) percentage of time when ED is at or above its capacity (throughput); 5) overall bed occupancy (throughput); 6) time from bed request to bed assignment (throughput); 7) time from triage to EP (throughput); 8) MD satisfaction (system); 9) time from bed ready to transfer to ward (output); 10) number of staffed acute-care beds (throughput).

The third most important measure, total time in the ED, corresponds to findings of the systematic review. A variety of ED times were the most common reported ED overcrowding measure identified in the review. For example, of the 735 measures identified, 292 (40%) involved a component of time in the ED. These measures are categorized as throughput issues in the model used in this study. Some
might say the fourth and fifth ranked measures — overall bed occupancy and percentage of time ED is at or above capacity — represent another form of access block.

Most of the measures selected by the Delphi group addressed factors affecting the ED’s efficiency in coping with the incoming flow of patients (throughput measures), followed by measures related to factors addressing the ability of the in-patient and ambulatory care systems to provide care after ED discharge (output measures).

This study is part of a trend to develop reliable measures for reporting ED overcrowding. The top 10 measures identified show a similarity to measures developed in other studies. The 36 measures that formed the first-round questionnaire were similar to the 38 measures of ED overcrowding identified by Solberg et al. using a group consensus method with 74 experts. Those authors reported “ED throughput time” and “ED boarding time” as the most important throughput and output measures respectively. These are similar in their operational definition to “total time in the ED,” and “time from bed request to bed assignment,” the fourth and sixth ranked measures reported in this study.

A qualitative Canadian study by Estey et al. identified two factors as important to defining and characterizing ED overcrowding: inefficient access to ED beds because of slow throughput of patients, and staff shortages. Schull et al. assembled an expert panel of clinical and administrative personnel in pre-hospital, ED, and hospital settings in Canada to develop a standard definition, and a list of determinants for ED overcrowding. Ambulance diversion was selected as an appropriate operational definition. ED-related factors considered as potentially important determinants included the number of admitted patients held in the ED, intermittent surges in numbers of newly arriving ambulances and ambulatory patients, ED physician staffing (physician-hours per day), ED physician characteristics, ED nurse staffing, ED nurse profiles, availability of ED social work and geriatric teams, response times to ED consultation requests, the enforcement of ED consultation timeliness policies, ED design (e.g., number of stretchers and cardiac monitors, size of department), and availability of radiological imaging off-hours.

The top 10 measures reported here also concur with measures that researchers in English-speaking countries outside North America have considered to be important for documenting ED overcrowding. Access block; the percentage of all patients admitted to, transferred to, or dying in the ED where their total ED time is >8 hours; and daily total patient care time (mean daily occupancy with patients being treated) are the key measures of ED overcrowding in Australia. In Britain, the National Health Service set as a target for the end of 2004 that the “total time in the ED” (the third most highly rated measure in this study) for 98% of ED patients not exceed four hours.

The importance of some measures may vary according to local criteria. For example, “total ED volume” or the “number of ED nurses” are not measures of ED overcrowding; they become measures when placed in the context of a hospital’s capacity. Other measures, such as “percentage of ED occupied by in-patients,” are standardized for sites and do not require a context to make them meaningful. For this reason, some authors have considered the former measures to be surrogate markers of overcrowding and not true measures.

The results of this Delphi study suggest that experts consider physician satisfaction as an indicator of ED overcrowding: whatever form overcrowding takes (long waiting times, access block, or patients in the hallway), it will lead to decreased physician satisfaction. Weiss et al. have provided evidence that the perceptions of ED physicians and nurses, and their feeling of being rushed correlate with...
seven “objective” measures of ED overcrowding: numbers of patients in the waiting room, at triage and at registration, the number of full rooms, hallway patients, patients awaiting beds, and the total number of patients registered.

The low rating assigned to staffing measures agrees with the results of a study by Schull et al. that found that the numbers of nurse-hours is not a significant determinant of overcrowding.

With the exception of ED physician satisfaction, all of the second-round measures provide a quantitative measure of overcrowding. As Bernstein et al. have noted, the development of a quantitative measure of ED activity can help put the concept of overcrowding in context with the overall activity levels of the department. This in turn will provide a more balanced perspective regarding the relationships between overcrowding, activity levels and adverse outcomes for patients.

6 DISCUSSION

6.1 Summary of Results

This report describes comprehensive methods that combine systematic and qualitative approaches to identify and define valid indicators of ED overcrowding in Canada. Through the combination of a detailed literature review and consensus technique among experts, relevant and clinically important indicators were identified.

6.1.1 Systematic review

The systematic review aimed to identify measures and indicators of ED overcrowding that have been used in the scientific literature. From 169 studies, 735 measures documenting ED overcrowding were identified. ED overcrowding is a term that is commonly used but inconsistently defined across studies; 31% of studies provided operational definitions. There is little agreement in the research literature on how to best investigate ED overcrowding, resulting in an assortment of measures and indicators. Overall, the measures provided information about specific events related to ED overcrowding, but alone do not provide a complete picture. Measures and indicators of ED overcrowding reported in the scientific literature have focused on delay represented as intervals throughout ED care (39.7%). Multiple definitions of ED time intervals are often used. Other measures addressed the volume of patients in the ED, as the overall volume of patients attended in the ED (11.6%), the volumes of patients waiting to receive care at different stages while in the ED (8.8%), or the proportion of patients being seen at different stages while in the ED (7.6%). Measures such as ED access block (7.0%), ambulance diversion (7.0%), number of patients who LWBS (5.5%), and ED length of stay (4.2%) were also used in the studies, although less so.

When the measures were analyzed according to the input-throughput-output model, throughput measures were the ones most commonly used to document ED overcrowding (67.8%), followed by input (19.5%), output (11.8%), and system (0.8%) measures. More than half of the measures (54%) were used to describe the frequency of an underlying dimension of ED overcrowding, or to discriminate between people with different levels of a particular attribute related to ED overcrowding. Of all the measures, 27.8% were used to measure the magnitude of longitudinal change on a dimension of interest as an effect of an intervention to reduce or control ED
overcrowding, and 18.2% were used to predict the incidence of a particular event related to ED overcrowding.

6.1.2 Delphi study

The objective of this study was to develop an inventory of potential and clinically relevant measures of ED overcrowding based on expert opinion. The most important measure identified was the percentage of the ED occupied by in-patients, a factor that has been called access block or EIP in the scientific literature. This measure is categorized as an output issue in the model used in this report, and reflects the role that factors external to the ED (e.g., hospital-related factors) play in explaining the problem of ED overcrowding. Research on access block remains under-represented in the scientific literature; only 7.0% of the measures identified in the systematic review documented access-block related events.

The other measures in the top five were total ED patients, total time in the ED, percentage of time ED is at or above capacity, and overall bed occupancy. Some less heavily weighted measures were time from bed request to bed assignment, time from triage to EP, physician satisfaction, time from bed ready to transfer to ward, and number of staffed acute-care beds. Most of the measures selected by the Delphi group addressed factors affecting the ED’s efficiency to cope with the incoming flow of patients (throughput measures), followed by measures related to factors addressing the ability of the in-patient and ambulatory care systems to provide care after ED discharge (output measures). The top 10 measures that resulted from consensus among Canadian ED experts also concur with measures that researchers in English-speaking countries outside North America have considered important for documenting ED overcrowding.

The heterogeneity of the measures selected by the reviewed studies and the Delphi group may be the result of confusion between causes, characteristics, and outcomes of ED overcrowding. The nature of the relationship between certain variables or events and ED overcrowding is unclear, and distinctions need to be made between correlational and causal relationships to better understand the problem. A correlational relationship implies that two things perform in a synchronized manner. For instance, the length of time in the waiting room will increase as the state of overcrowding worsens, and vice versa. Most of the variables identified by the Delphi group have this characteristic.

The measure of physician satisfaction is an example of a causal relationship, meaning that one variable is responsible for observed changes in another. It is also unidirectional (i.e., if the ED is overcrowded, it is likely that the staff will be dissatisfied or distressed); the opposite is not always true. As ED overcrowding decreases, perhaps physician satisfaction will increase; other factors (e.g., number of support staff, types of patients, acuity levels, time of day) may also influence physician satisfaction. Even though this measure is a contrast from others, many of the experts participating in the Delphi study think that there is a link between ED overcrowding and physician satisfaction. The feeling by staff of being rushed has been considered to be a measure of ED overcrowding in other studies.4,23,30

6.2 Study Limitations

The systematic review satisfied the criteria to perform a rigorous overview of the measures used in the scientific literature to document ED overcrowding. A clear research question was posed a priori, and a comprehensive search of the literature, including grey literature, was performed. The study
selection process was robust: a defined set of eligibility criteria was developed after consultations with a panel of experts, and selection biases were controlled by having at least two reviewers select studies. There are a few limitations. The strict eligibility criteria required that studies include the terms “crowding,” “overcrowding,” or synonyms in the title, introduction, or methods sections. This may have excluded some studies that might have influenced the choice of measures selected. As a result, the findings may not represent an exhaustive list of measures of ED overcrowding. The systematic review may be criticized for not assessing the methodological quality of the included studies. Assessing the internal validity of any research study is a component of any systematic review, but given that the objective was descriptive, and focused on the frequency of reporting measures for ED overcrowding, the quality component of the studies was beyond the scope of the review.

The Delphi study also has limitations. It can be argued that providing a list of selected measures to the Delphi members may have biased their selection. The structure of the Delphi process minimized this possibility. Also, the TEP meetings were part of the process of defining and refining the measures under consideration. A pilot study was conducted to verify that the list of measures (related and not directly related to ED overcrowding) was exhaustive. Delphi members were also given an opportunity to modify the measures, and to suggest additional ones.

By asking participants to rank measures individually, the possibility of selecting a combination of measures as the best indicator was obviated. The system-level features of overcrowding may mean that appropriate measures of ED overcrowding must go beyond ED processes. Multiple or composite measures may be required to accurately record and report ED overcrowding; few published studies have evaluated such measures.

The definition of a measure as “something that changes as a result of changes in overcrowding” has allowed for a mix of measures that may not be conceptually compatible with one another, or that may not correspond well with overcrowding. Because there is no gold standard for measuring overcrowding in all situations and across all systems, all the indicators are proxy measures of the condition. Whether they will measure overcrowding depends on how strongly they are associated with it; not all indicators are equally good.

The Delphi method has several limitations. The model described here used an e-mail questionnaire, whereas classic Delphi studies use the post. Despite the low response rate of e-mail surveys, the research team considered an electronic survey to be appropriate, given the large volume of correspondence that participants would receive. The high response rate suggests that the format did not discourage recipients from responding. It was also believed that using a familiar document format would help minimize the time spent by participants. The structure of the Delphi technique helped to reduce the risk of non-response: participants were contacted directly, and were informed of the role that they were to play as part of a group responsible for developing measures of overcrowding.

The selection of the Delphi group was based on the opinions of the TEP members; this is another limitation, as is their availability within the allocated period. The Delphi group and the TEP were predominantly composed of emergency physicians, and an ED nursing perspective was under-represented in this study. As a result, the responses obtained with the Delphi method may over-represent the opinion of emergency physicians. They may not represent the opinions of the spectrum of professionals who work in the ED who are also affected by the problems associated with ED
overcrowding. Furthermore, the fact that the TEP was composed of Canadian experts may have restricted the choice of potentially relevant international measures.

None of the measures were adjusted based on the ability of institutions to collect information, even though other authors have suggested that availability of information and ease of quantification are crucial aspects of any measure of overcrowding. This may make it difficult for policy makers to implement the 10 measures for the reporting of ED overcrowding.

Though this study has reported the measures as the result of a consensus-generating process, there is debate over how consensus should be determined. The number of rounds required for consensus and the numerical values (e.g., the standard deviation) that might indicate consensus are suspect indicators, and there is no universally accepted method of consensus. There is support for the methods used here: choosing to stop after two rounds and using a “statistical group response” as an indicator of the group judgment (usually the median value and quartile ranges after the final round) have been discussed in a study on the Delphi study’s theoretical foundations.

6.3 Generalizability of Findings

The set of measures to document ED overcrowding identified from the medical literature may be influenced by variations in time, the availability of the measures, and the settings where the studies were conducted. Factors that are most important in academic, urban EDs may not be the same as those in suburban EDs. The importance of some measures may vary according to local criteria, and even change over time. Aspects of ED overcrowding that are most salient at one point are not necessarily the ones that will be most salient later. This should be considered when interpreting the results of this report.

The TEP and the Delphi participants were predominantly health care providers and researchers in large, urban, academic hospitals. Thus, the applicability of their selected measures to smaller, non-academic hospitals would need to be examined. Regional variations in the causes and impact of ED overcrowding across Canada may not be captured by the set of measures that have been rated highly. The heterogeneity of the panel members, however, added to the richness of the discussion, and may have increased the generalizability of the final indicators selected.

6.4 Health Services Impact

Some jurisdictions and EDs across Canada may routinely collect enough data in provincial or national ED administrative databases that enable them to measure more indicators than identified in this report. The inconsistent use of definitions, indicators, and measures of ED overcrowding may create a contradictory picture that fails to capture the multi-dimensional nature of the problem. The measures and indicators presented in this report may help research and development, and guide the improvement of uniform ED data collection systems to track overcrowding across Canada. Definitions will need to be developed to ensure a uniform understanding and use of the measures and indicator.

6.5 Knowledge Gaps

The findings from the literature review point to gaps, and suggest future directions for research. Studies on ED overcrowding often use heterogeneous definitions of measures, and even the term
“ED overcrowding” is used inconsistently. It may not be in the title of an article, yet it is measured, though perhaps only in one domain.

Future research can build on what has been learned here. A logical next step is to develop a shared composite set of indicators of ED overcrowding, including input, throughput, and output components. Such work would yield contributions to the conceptualization and measurement of ED overcrowding, and improve the services provided in the ED.

Few attempts have been made to determine the properties of the measures used to document ED overcrowding. More investigations are required to provide a greater understanding of the sensibility, feasibility, reliability, validity, and responsiveness to change of the measures being used in research. Without a greater knowledge of the operating characteristics of measures, the results will remain difficult to interpret, and be of limited value to policy makers, ED administrators, ED staff, and patients.

7 CONCLUSIONS

This report has identified measures or indicators that may be used to evaluate and monitor ED overcrowding in Canada. A variety of measures are used in the published literature, but their use says more about their availability than their importance to clinicians. The variety of measures also reflects the fact that the most worrisome “symptom” of overcrowding varies across jurisdictions and over time. Hence, the measure that is “best” for a given hospital or jurisdiction is likely to change, and the jurisdiction may want to vary its measures over time to reflect the variable nature of overcrowding and its context.

While most measures identified in the systematic review involved throughput issues, the Delphi technique identified output (access block) issues as being of greatest importance to clinicians. The evidence provided here suggests that limited consideration has been given to the choice of measures to document ED overcrowding. To better understand the problem of ED overcrowding in Canada, policy makers, ED administrators, staff, and patients may find the measures identified in the Delphi technique of value. Using uniform definitions and measures may help with intra- and inter-institutional comparisons, and help in designing and implementing interventions aimed at reducing overcrowding in EDs across Canada.
8 REFERENCES


68. Channan P, Bullard M, Alibhai A, Saunders D, Rowe BH. Reasons why patients leave without being seen from the ED. CJEM 2003; 5(3).


88. Espinosa JA, Treiber PM, Kosnik L. A reengineering success story: process improvement in emergency department x-ray cycle time, leading to breakthrough performance in the ED ambulatory care (fast track) process. Ambul Outreach 1997; 24-7.


130. McAfee AT. Effect of implementation of a computerized order entry system on emergency department patients' length of stay. Acad Emerg Med 2003; 10(5):504-505.


140. Purnell L. Reducing waiting time in emergency department triage. Nurs Manage 1995; 26(9):64Q.


150. Richardson DB, Bryant M. Confirmation of association between overcrowding and adverse events in patients who do not wait to be seen. Acad Emerg Med 2004; 11(5):462.


41(4):467-76.


186. Vertesi L. Does the Canadian emergency department triage and acuity scale identify non-urgent patients who can be triaged away from the emergency department? CJEM 2004; 6(5):337-42.


APPENDICES

Available from CADTH’s web site
www.cadth.ca