Triage narratives – exploring a widely collected but dramatically underexplored source of emergency department data

by

Christopher Picard

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Nursing

Faculty of Nursing
University of Alberta

Abstract

Emergency departments (ED) around the globe rely on triage to sort which patients can wait safely for care when the demand for service outstrips the resources available. The triage process is typically performed by experienced emergency nurses and can be the earliest documented assessment for patients receiving hospital-based care. Because the ED is the most common entry point for patients, triage documentation, specifically triage narratives, are nearly ubiquitous for hospitalized patients. The data collected during triage can be structured (information such as age and sex) or unstructured (in the case of free-text narratives). In Canada, ED data are reported as a national dataset, much of the reported data is collected at triage, but little of the national dataset reflects nurses own work and the triage narratives are not included. The goal of this paper-based thesis is threefold: first, to describe the role of ED nurse in data collection; second, to provide a scoping review protocol for examining triage narratives; and, to describe how triage narratives have been used in the academic literature.

The first portion of this thesis reviews the systems used in Canada for ED data collection. This chapter highlights how little nursing data is collected and proposes that national organizations need to demonstrate that there is utility in collecting these triage data. It calls on the National Emergency Nurses Association to take concrete first steps to press for ED nurse data to be included in national ED data registries.

In the second phase of this thesis, I provide an outline and rationale for how to review one specific component of triage nurse data: the unstructured narrative. The narratives, which have been included in other reviews of clinical documentation have show promise for a number of uses, but they have not been reviewed as a unique form of documentation. The scoping review protocol outlined each step that should be taken in conducting a review on the topic to ensure

that subsequent reviews are rigorous. It offers a template upon which future reviews can use and improve upon.

The third component of this thesis are the results of the scoping review, which mapped the literature around triage narratives and described the narrative structure and its uses. In this review we screened 18,074 studies published between 1990 and 2022 in CINAHL, MEDLINE, Embase, Cochrane, and ProQuest Central. 96 studies from this initial sample examined triage nurse narratives. Theses studies represent over 12 million ED visits and 2438 EDs. We grouped reviewed studies as using narratives for one of the following: case identification, as input variables for predictive modeling, and for quality improvement. Despite the breadth of research there was not much data on the structure of narratives: 27 studies reported the keywords that they used to search triage narratives, 7 offered descriptions of the narratives (typically the number of words used), but only one study described the narratives in any detail. Unfortunately, none of the studies that described the structure of narratives also described the nurses who wrote them – of the 9 studies that included a sample size for the nurses, none described the narratives. Additionally, fewer than 10% of studies (n=8) used reporting guidelines.

This thesis demonstrates that there are myriad uses for triage narratives, and potential for additional synthesis of triage narrative studies. It also details how there is a striking paucity of research describing the structure of narratives or the nurses who generate them. There is a clear opportunity to describe triage narratives. In doing there is an opportunity to highlight the importance of this data type and to strengthen the validity of studies that have used them to conduct research. Canadian EDs are well positioned to do support this work given the national triage system and existing data collection structures.

KEYWORDS nursing; artificial intelligence; machine learning; triage; review, narrative

Preface

(Mandatory due to collaborative work)

The components of this thesis have been published with the assistance of the several collaborators.

Chapter 1 of this thesis has been published as Picard, C., & Kleib, M. (2021). Advancing Emergency Nurses' Leadership and Practice through Informatics: The unharnessed power of nurses' data. *Canadian Journal of Emergency Nursing*, 43(3), 13–17.[1] I was responsible for manuscript concept formation, preparation, and revision. M. Kleib was the second author, and academic supervisor, and was involved with manuscript preparation and revisions.

Chapter 2 of this thesis has been published as Picard, C., Kleib, M., O'Rourke, H. M., Norris, C. M., & Douma, M. J. (2022). Emergency nurses' triage narrative data, their uses and structure: a scoping review protocol. *BMJ open*, *12*(4), e055132.[2] All authors contributed to the design of this protocol. C.T. Picard, M. Kleib and H.M. O'Rourke initiated the project. The protocol was drafted by C.T. Picard and was refined by M.Kleib, C.M.Norris, H.M.O'Rourke, and M.J.Douma.

Chapter 3 of this thesis has been submitted as **Picard, C**., Kleib, M., O'Rourke, HM., Norris, CM., & Douma, M. J. (2022). Emergency nurses' triage narrative data, their uses, and structure: a scoping review. *The journal where the finished manuscript will be submitted is yet to be determined but will be a higher impact open access journal (i.e., Annals of Emergency Medicine, Journal of Medical Internet Research, or the Journal of the American Medical Informatics Association)*.

Acknowledgements

I would like to acknowledge all of the people who supported me through the process of creating this thesis: my committee, my colleagues, and my family. To my supervisor, Dr. Manal Kleib, thank you for your patience and guidance, you have been a pathfinder in the field of informatics and have helped to illuminate my road to success. To my committee, Dr. Hannah O'Rourke, thank you for your specific, focused guidance, and methodological expertise; Dr. Colleen Norris, thank you for your unwavering support and ability to recognize the importance of triage narratives. To my supportive colleagues, Matthew Douma, Richard Drew, and Domhnall O'Dochartaigh: thank you for joining me on this journey; your support both academically and professionally have been instrumental in my success. Foremost thank goes to my family: Walter and Oscar, you have sacrificed time to support me; though you don't know it your natural ability to find fun and make the most out of every moment has made this process a success. To Kelly, without your support and love this would have been impossible. You're a constant source of perspective and inspiration I love you and I've loved being able to walk this road with you - I look forward to all of the new adventures that await.

Table of Contents

Table of Figures

Introduction

Canadian Emergency Departments (EDs) collect patient data using the Canadian Emergency Department Information System (CEDIS).[3] This information is collected nationally by the Canadian Institute for Health Information (CIHI) in the form of the National Ambulatory Care Reporting System (NACRS).[4] These data have been used to track disease patterns and offer valuable insights into the status of Canadian healthcare. Emergency Department (ED) nurses collect a substantial amount of this data (20 to 50% of the required data fields), meanwhile only 5% of NACRS fields collected reflect nursing assessments, interventions, or times to patient.[4,5] ED nurses have carried a disproportionate responsibility for ED information collection but their efforts are largely unrecognized when the data they generate are analyzed or when the systems they use are developed. This research hypothesized that ED nurses collect significant amounts of data at triage that are used for clinical and research purposes, but these data are not recognized as reflecting nurses' work. My thesis involved reviewing the current state of ED data collection in Canadian healthcare and then systematically reviewing the published literature on one key component of triage data: the triage narrative.

In the first chapter we explored the issue of the ED data collection and Canadian nurses' roles in the process. This paper offered an overview of how the Canadian triage Acuity Scale (CTAS) and Canadian Emergency Department Information System (CEDIS) as clinical decision support tools use standardized languages, and how ED nurses perform a key role in Canadian ED data collection. We discussed how nurses' make significant contributions to data collection, but have little of their work captured in national data sets, are under-represented in the CEDIS and CTAS national working groups and how ED nurses have fallen behind other nursing specialties in incorporating informatics competencies into their practice. We highlighted the steps other nursing specialties have taken to highlight the important work nurses complete and discussed how ED nurses could follow these frameworks. We concluded our paper by calling on the Canadian National Emergency Nurses Association (NENA) to take concrete steps to recognize the importance and utility of nursing data and to take concrete steps to addressing the disparity.

The second chapter identified one specific component of triage data to investigate: the triage narrative. The paper in this second chapter was a scoping review protocol for examining triage narratives. The protocol began by outlining some of the ways these triage narratives have been used, and highlights some of the aspects of previous reviews that could be improved upon. Specifically, how previous reviews that included triage narratives used searches that were not comprehensive or systematic, combined triage narratives with other sources of narrative data, and used searches that were restricted to narratives that focused on a narrowly defined patient population. Our protocol justified the search by highlighting that no existing work has reviewed the uses and structures of triage nurses' narratives. Our protocol outlined the goals of our review: to map the existing literature for the uses and structures of triage narratives, and to identify whether there is sufficiently similar data to perform future systematic reviews or meta-analyses. The protocol outlined the steps we took in performing the scoping review and will allow future researchers to replicate the study by providing: the search terms, the screening forms, and the data extraction templates used.

The third chapter of this thesis focused on the findings of our scoping review. We began this paper by reviewing the importance of triage narratives, by discussing how triage narratives were included in previous reviews, and how the absence of a review examining the structure and from of triage narratives was a gap in the existing knowledge base. We then described our methods and outlined how we followed the steps described in our protocol (chapter two) while conducting our study. Our review identified 96 studies that examined triage narratives and emphasized how the majority of studies were conducted after 2008 and in the United States, Australia, and Canada. We noted that the methods of conducting these studies have changed over time (machine learning has become predominant after 2017); how studies can be grouped into three categories of narrative use (for quality improvement, for case identification, or for use as a predictor variable); and how there is a paucity of research directly reporting on the structure of triage narratives. There were a number of studies that demonstrated how triage narratives improved identification rates when searching for specific populations compared to more traditional methods such as diagnosis codes, and these studies may be sufficiently comparable for further systematic review. We noted that there were few studies using reporting guidelines and that future studies are needed to address the heterogeneity in data reporting. Future research is needed to define the structure of triage data.

The fourth chapter of this thesis highlighted the ongoing and potential uses for our existing work and opportunities for future research that may be informed by our research. We grouped these uses and opportunities within the contexts of nursing education, nursing practice, nursing research, and nursing policy development. I concluded the thesis by discussing the limitations of our program of research.

Chapter 1: Advancing emergency nurses' leadership and practice through informatics: The unharnessed power of nurses' data

Abstract

Collection of data in healthcare is vitally important to inform clinical decisions, resource planning and monitor effectiveness of care. The Canadian Emergency Department Information System and Canadian Triage Acuity Scale are primary tools for collecting such data. Although emergency nurses use these tools to collect significant patient and healthcare data on a daily basis, their understanding of the purposes and implications for collecting these data is unknown. Emergency nurses' limited awareness about informatics, and the underrepresentation in information and communication technology strategic initiatives and research are barriers to their realizing the full benefits of information and communication technology. The National Emergency Nurses Association is well positioned to provide the leadership required to move nurses from being data collectors, to information users by maximizing their potential to advance Canadian emergency nursing practice through informatics.

Keywords: informatics, datasets, triage, emergency nursing

Key Takeaways

1. Canadian ED nurses use CEDIS and CTAS to generate vast amounts of comparable electronic patient data nationally.
2. Data generated during triage such as narrative notes have been used internationally for disease surveillance and research. In Canada the data has not been well explored.
3. Although Canadian ED nurses are significant collectors of data, they are underrepresented in national informatics working groups.

4. There is a need for national leadership to describe nursing informatics competencies and increase nurse participation in informatics.

Manuscript

Introduction

Collection of data in healthcare is vitally important to inform clinical decisions, resource planning and monitor effectiveness of care. The Canadian Emergency Department Information System (CEDIS) and Canadian Triage Acuity Scale (CTAS) nationally are primary tools for collecting and storing such data in emergency departments (ED). Although nurses use these systems on a daily basis, there is a paucity of research describing nurse understanding of these systems and the implications of data collection for national-level healthcare planning. There is also a mismatch between how nurses collect these data and use it in the context of care, with nurses being underrepresented in both research and policy development. This paper will briefly describe these data systems and identify opportunities for nurses to assume leadership in ED informatics.

CEDIS and CTAS as Foundational Pillars of Canadian ED Informatics

The CEDIS dataset uses the standardized International Classification of Disease (ICD) codes, with specific data elements unique to ED practice.[3,6] Standardized clinical terminology or language for CEDIS and the National Ambulatory Care Reporting System (NACRS) data standard fulfill data reporting requirements, as mandated by the Canadian Institute for Health Information (CIHI), and enable 84% of Canada's emergency departments to collect patient data in a format that is useable to those who need it.[7] The breadth of CEDIS data reported to CIHI varies between hospitals; but the minimum dataset includes: patient-, provider-, and facility-specific information.

CEDIS data has been used to perform disease surveillance, track work that is being done in EDs, identify quality indicators and create decision support tools.[8] CEDIS data and historical trends allow administrators to predict workload, compare costs, throughput, and patient demographics between sites for benchmarking purposes.[8] Managers use data to monitor patient acuity levels, department flow, and monitor for pandemics.[8] Bedside, clinicians generate CEDIS data through live time mapping of patient acuity level and location, and through point of care order entry, vital signs tracking, and electronic charting.[9] Arguably, one of the most tangible and best-validated benefit of CEDIS is from electronic triage assistance using the Canadian Triage Acuity Scale (CTAS).

CTAS is one of the most comprehensively studied English-language triage tools in existence.[10] It is used internationally and when large study samples were assessed, CTAS was the only triage system to have zero deaths in the lowest acuity level of patients studied.[11] The CTAS represents a clinical decision support tool that uses information entered into CEDIS data fields to aid care providers' clinical judgment when assessing patients in the ED. This information includes elements such as coded primary complaint, narrative summary of the triage assessment, and vital signs to help nurses assign an acuity score. These acuity scores offer guidance on how long patients can safely wait, how frequently patients are reassessed, to which areas of the department patients are stratified, and which patients should be seen first. CTAS-aided triage has been proven superior to clinical gestalt[12] with excellent predictive validity for length of stay, severity of illness and resource utilization,[13] and, therefore, 95% of Canadian EDs use CTAS as their preferred triage system.[9] Clearly, using these systems requires more than technical proficiency to be able to appreciate the role of information and communication technology (ICT) as a tool for data and information management in complex healthcare environments such as the ED.

Describing the Collection-Use Gap in Canadian ED Data

Although Canadian ED nurses use the CEDIS and CTAS to collect significant patient data on a daily basis, their understanding of the purposes and implications of these data is largely unknown. To date there is a surprising paucity of nurse representation in ED informatics research, quality improvement, or professional advocacy. For example, nurses perform nearly all triage and collect 20 to 50% of the required CEDIS data fields, yet only 5% of these fields are dedicated to capturing the work nurses do.[4,5] The work that nurses do, and which is infrequently captured in these systems, includes…Despite being a large group of care providers and users of ED datasets,[14,15] nurses have limited representation in initiatives aimed at enhancing ED information systems, such as the CEDIS National Working Group (NWG).[16] Initially, this initiative was intended as a joint project between CAEP and the National Emergency Nurses Association (NENA).[6] Nonetheless, ED nurses represent less than 13% of the CEDIS NWG membership, with a majority of physicians steering this group.[16] Additionally, ED nurses have limited participation in ED informatics research in Canada and are under-represented within research and steering committees charged with setting direction on the ICT use in EDs, despite being primary users of these technologies. For example, when the author list of a large review was assessed, nurses constituted only 15% of CTAS researchers, with only one study out of 32 of CTAS having had a nurse as the principal investigator.[10]

Given their high involvement in collecting CEDIS data and using CTAS, an opportunity exists to highlight ED nurses' contributions, and to examine the value and impact of this data on practice and patient outcomes. NENA is well positioned to provide the leadership required to move nurses from being data collectors, to information users by maximizing their potential to advance emergency nursing practice through informatics.

Informatics Competency as A Catalyst for Change in Emergency Nursing

Informatics tools have the potential to improve consistency of care, patient outcomes, cost efficiency, and nursing knowledge.[17] Nursing Informatics is defined as the "science and practice (that) integrates nursing, its information and knowledge, with the management of information and communication technologies to promote the health of people, families, and communities worldwide" (para. 4).[18] Historically and to this day, nurses' work is all about data and information. At the point of care, ICT tools such as clinical information systems are essential for managing and synthesizing healthcare and patient data needed to support clinical decision-making for all healthcare providers including nurses.

According to the Canadian Nurses Association (CNA) and Canadian Nursing Informatics Association (CNIA), technology and innovation are continuously transforming healthcare and nursing practice.[19] For nurses to adapt to this new culture of digital health, they must acquire and maintain competency in informatics so that they are able to optimize practice and keep pace with advances in technology.[19] Informatics competent nurses are better able to advocate for incorporating data elements unique to nurses' interventions in existing and emerging clinical information systems so that nursing knowledge and contributions to improving patient and system outcomes are more visible.[15,19,20]

For example, within Canadian emergency nursing, nurses have a unique opportunity to show the value of nursing narratives inputted into triage datasets, which is evolving internationally, but not explored within the Canadian context yet. For example, nurses' triage narratives have been studied in a systematic way,[21] and used for real-time bio-terrorism surveillance,[22] and for epidemiological purposes such as identifying injury[23] and drug use patterns.[24] These studies have shown that triage narratives can be used in isolation,[25] or in combination with ICD codes for epidemiological research;[26,27] with some research suggesting that the triage narratives may even be superior to coding or ICD codes for identifying clinical cases.[28] Despite the time invested by Canadian triage nurses in recording these data during triage assessments, to date there has been no push for CIHI to include it as part of mandatory data reporting. Initiatives to incorporate triage narratives into nationally collected data could create a whole new source for epidemiological data, a source that starts at the earliest moments in patient care. Yet, these opportunities are currently being missed out due to limited awareness about emergency nursing informatics.

In 2012, the Canadian Association of Schools of Nursing (CASN) developed and approved the Nursing Informatics Entry-to-Practice Competencies for Registered Nurses in Canada.[29] These competencies include the following dimensions: Information and knowledge management, professional and regulatory accountability in using digital technologies, and the ability to use various digital health technologies in the delivery of patient care. Although these competencies have been in existence for over a decade now, a survey by Canada Health Infoway of nurses providing direct care (n=2058) revealed only 30% of these nurses were familiar with these competencies, just 7% actually applied the competencies in their practice, and 67% were not familiar.[30] In another recent study involving practicing registered nurses in Alberta (n=2,844), results revealed the two main areas nurses reported struggling most with were: information and knowledge management and the use of ICT in the delivery of patient care, core elements of informatics competency.[31] Although no specific research is available yet on ED nurses' perceived competency in informatics, these studies suggest that these patterns are likely evident in this group.

As a self-regulated profession, the emergency nursing standards of practice provide "the lens by which the public views and evaluates nursing performance and to which nurses are held accountable" (p. 4).[32] Currently, the Emergency Nursing Scope and Standards of Canadian Practice and the Emergency Department Core Competencies emphasize advanced practice, the use of CEDIS and CTAS in triage, and research among others skills. However, these documents fail to align these requirements with informatics concepts and competency.[32–34] Given the extensive use of technology in ED and the intricate role ED nurses play in the collection of critical patient and healthcare data in this complex environment, it is of utmost importance that NENA takes a proactive approach to addressing informatics competency needs among this group.

Steps to Close the Gap

As a beginning step toward this goal, NENA can endorse the CNA and CNIA joint position statement on nursing informatics (NI).[19] Endorsement provides Canadian ED nurses with a direction to understand how informatics is integral to the nursing role. The position statement offers important insights on the value of using standardized clinical languages such as the International Classification for Nursing Practice (ICNP) and InterRAI, providing a roadmap for ED nurses to think about data elements being collected within the CEDIS and whether these capture nurses' work or not. NENA may also consider defining specific strategies based on the NI position statement to expand ED nurses' competency in informatics. This planning, however, would require an understanding of ED nurses' current state of informatics competency. While there are many validated tools such as the SANIC, Self-Assessment of Nursing Informatics

Competencies Scale;[35] and the Technology Informatics Guiding Education Reform, TIGER-based Assessment of Nursing Informatics Competencies;[36] the Canadian Nurse Informatics Competency Assessment Scale (C-NICAS) is more relevant to the Canadian context because it is based on CASN's entry-to-practice NI competencies for Canadian nurses.[37] Administering this tool for assessing Canadian emergency nurses' perceived informatics competency and factors impacting development of informatics competency in these nurses in future research, which could be sponsored by NENA, would be an important first step to determine current state and identify strategies to address gaps.

CASN's leadership in increasing informatics capacity in Canadian nursing education through defining NI competency requirements for registered nurses and development of educational resources is remarkable.[38] For example, CASN's digital health resource, publicly accessible on CASN's website, provides com-prehensive learning resources to support self-directed learning in informatics; nonetheless, not many nurses are aware of these resources.[38] While these resources were primarily intended to support nurse educators in integrating entry-to-practice NI competencies in nursing education, these learning opportunities can be of benefit to all nurses across practice settings. Increasing awareness about the CNA and CNIA joint position statement and other relevant informatics resources including the C-NICAS and CASN digital health resource may encourage nurses to participate in self-assessment to identify competency gaps, which may subsequently encourage them to obtain formal or informal education in informatics. NENA, as a national voice for ED nurses, is uniquely positioned to begin engaging ED nurses in this process of learning and valuing the role of informatics in modern day nursing practice.

Conclusion

Canadian ED nurses utilize ICT tools, specifically CEDIS and CTAS, on a daily basis to input important patient and healthcare data. Yet, to date, Canadian ED nurses' awareness about informatics is unexplored. There is also limited nursing representation in ICT strategic initiatives and research within EDs themselves. These barriers may be preventing nurses from realizing the full benefits ICT offers to improve patient and system outcomes, and nursing knowledge development. Describing and advancing ED nurses' informatics competency should be a key priority. ED nurses' contributions, and their ability to examine the value and impact of healthcare data is essential to ensuring nurses are not merely inputting data in hospital clinical information systems datasets as a task, but rather they are able to use these data to improve patient outcomes. NENA is well positioned to lead this change through endorsing the CNA & CNIA joint position statement on nursing informatics to provide guidance to ED nurses.[39] Research to understand the state of informatics preparedness in ED nurses is vital to inform future planning. The C-NICAS, a validated Canadian-based informatics competency assessment scale, could be applied to facilitate self-assessment and continuing education in informatics.

Chapter 2. Emergency nurses' triage narrative data, their uses and structure: a scoping review protocol

Abstract

Introduction: The first clinical interaction most patients have in the emergency department occurs during triage. An unstructured narrative is generated during triage and is the first source of in-hospital documentation. These narratives capture the patient's reported reason for the visit and the initial assessment and offer significantly more nuanced descriptions of the patient's complaints than fixed field data. Previous research demonstrated these data are useful for predicting important clinical outcomes. Previous reviews examined these narratives in combination or isolation with other free-text sources but used restricted searches and are becoming outdated. Furthermore, there are no reviews focused solely on nurses' (the primary collectors of these data) narratives.

Methods and analysis: Using the Arksey and O'Malley scoping review framework and PRISMA-ScR reporting guidelines, we will perform structured searches of CINAHL, Ovid MEDLINE, ProQuest Central, Ovid Embase and Cochrane Library (via Wiley). Additionally, we will complete forward citation searches of all included studies. No geographical or study design exclusion criteria will be used. Studies examining disaster triage, published before 1990, and non-English language literature will be excluded. Data will be managed using online management tools; extracted data will be independently confirmed by a separate reviewer using pre-piloted extraction forms. Cohen's kappa will be used to examine inter-rater agreement on pilot and final screening. Quantitative data will be expressed using measures of range and central tendency, counts, proportions, and percentages, as appropriate. Qualitative data will be narrative summaries of the authors' primary findings.

Patient and public involvement: No patients involved.

Ethics and dissemination: No ethics approval is required. Findings will be submitted to peer-reviewed conferences and journals. Results will be disseminated using individual and institutional social media platforms.

Manuscript

Introduction

The first interaction most patients will have with a care provider in the Emergency Department typically occurs during triage. A nurse will conduct a brief patient assessment, record demographic information, and assign a visit code and acuity level.[10] The decisions nurses make during this triaging process are used to stream patients into different parts of the department and ensure they do not deteriorate if medical care is delayed.[10] The triage process originated in battlefield hospitals during the Napoleonic wars,[40] but didn't see widespread incorporation into healthcare until the 1960s,[41] a formal taxonomic approach until the 1980s,[42] or national implementation until 1994, when Australia became the first country to do so.[43] Most countries now use both formalized triage processes and validated triage tools;[10] with many systems basing their model on the Australian system.[10,44]

Because triage is typically the first clinical encounter for most patients, the triage record is the first source of in-hospital documentation. In some jurisdictions (such as Canada) national standards guide both the triage process and ED data collection,[32,45] and the vast majority (85%) of EDs collect and report minimum data sets to a national repository.[3,5,6] These minimum data sets help administrators assess workload trends, compare costs, measure throughput, monitor patient acuity, and perform pandemic surveillance.[3] Minimum data sets require documentation of a set amount of structured data, but also allow for the optional reporting of significantly more data, such as triage narratives. Despite their routine collection in some jurisdictions, these narratives are not part of the 'minimum' data set,[1,5,46] and until recently have not been used in a consistent manner.

Unstructured triage narratives differ from structured data in that they aren't restricted to one of several fixed options. These narratives capture subjective data such as the patient's own reported reason for visiting the emergency department and the nurse's assessment; and, as such, offer significantly more nuance than an itemized 'presenting complaint', or procedural or diagnostic codes. While unstructured data have historically been difficult to utilize, recent advances in artificial intelligence (AI) have been applied to rapidly assess large volumes of triage narratives with strategies including natural language processing,[47] computerized keyword searches, [48] and computer aided expert analysis.[49] The results of previous studies suggest that narrative data can be useful for predicting important clinical outcomes such as: the need for admission to hospital, [25] or receiving a diagnosis of sepsis,[26] or influenza.[49] In addition to being predictive of clinical course, narrative data have been used for epidemiological purposes such as tracking characteristics and rates of injury,[50] and substance misuse patterns.[51,52] Furthermore, when triage narratives have been added to more traditional approaches for estimating disease prevalence (i.e. International Classification of Disease - ICD codes) they have been able to identify significantly more cases,[51,52] with one study showing that triage narratives can more than more than quadruple (90% versus 21%) the number of cases (road collisions involving drugs or alcohol) identified when compared to ICD codes alone.[51] The ability to identify cases is of particular importance for rare or complex presentations that are not easily captured using existing diagnostic codes.[48] These triage narratives have been clinically operationalized by the Emergency Nurses Association, who have incorporated AI into triage decision support.[53] The clinical implementation of AI into triage suggests that the research field is poised to expand significantly, and as such a mapping of the current literature base is needed.

To date, at least three systematic reviews have included triage narratives in their examinations of unstructured clinical data. Two systematic reviews (an initial review[23] and subsequent update[54] authored by the same group) identified 56 papers examining (AI) approaches to analysing injury surveillance narratives. The reviews included studies that assessed narrative data from any source and any administrative health database, provided that it was assessed using (AI) algorithms. The reviews included a number of studies that used triage narratives, and found that narrative data offered not only an alternative method for performing surveillance, but also offered rich contextual information not otherwise available.[54] A third systematic review, which examined ED syndromic surveillance of influenza identified 38 studies that tracked illness, several of which used both nursing triage narratives and nurse assigned complaint codes.[55]

More recently there have been scoping reviews of (AI) that have examined their utility in aiding in medical diagnoses[56] and for clinical decision support[57] Both reviews examined triage and unstructured narratives data, but they included these with disparate types of data. Additionally, neither study referenced the previously mentioned systematic reviews, and both missed a significant number of studies identified in the systematic reviews that examined AI. Some of these omissions may have been intentional given that one review excluded studies that were linked to specific presentations (such as influence, chest pain or trauma)[57] or because the other review included only those that explicitly mentioned the use of synonyms for “artificial intelligence” in their title, were not directly relevant to medicine (e.g., public health) or “did not report outcomes or evaluations” (e.g., non-intervention epidemiological studies).[56]

These reviews are foundational, and our work will attempt to add to them. The systematic reviews suggest that there is limited utility in examining narrative data in general, concluding that a "lack of guidelines for recording information… and subsequent inconsistencies in the documentation"[23] complicate the use of the data. These conclusions may be either a reflection of the limitations present in AI at the time, or the result of generalizing about the ability of narrative data to reflect more narrowly defined and specific presentations (e.g., injuries or communicable diseases). The authors of the systematic review limited their searches by excluding grey literature,[23,54] and examining only studies indexed on PubMed.[55] The scoping reviews, while more current, have limited their search breadth by using restrictive search strategies that exclude studies linked to specific presentations. Our work will add to other scoping reviews that have examined triage narratives as part of their data; but will offer an updated overview of the early systematic reviews, and will focus on examining the clinical, epidemiological, and prognostic value of the narrative data generated by the clinicians who are pushing to incorporate AI into the very earliest stages of clinical documentation: emergency nurses.

Objective

Using the Arksey and O'Malley[58,59] scoping review framework and PRISMA-ScR reporting guidelines,[60] the goal of our review will be to identify, map, and describe the peer reviewed literature examining nursing triage narratives in order to identify if there is sufficient literature to conduct future systematic reviews on: the use of triage narratives for estimating prevalence for other conditions and populations, the use of artificial intelligence at triage, and to determine whether there is sufficient literature to meta-analyse comparisons of prevalence estimates using narratives and other coding sources. To further these goals, we will focus on the following objectives:

1. To describe the populations and locations where the data originated,
2. To determine the types of literature that have been generated on the topic,

3. To describe the authors reported purpose and conclusion in the literature.

Methods and Analysis

Search Strategy

The review will be conducted between February and May 2022. We will identify relevant literature through structured searching of online databases and forward citation searches of included studies. With the assistance of a medical librarian our review will use controlled vocabulary (Appendix 1) to perform a broad search of the following databases: PubMed, CINAHL, MEDLINE, and ProQuest. To ensure comprehensive coverage, citation searches will be performed for included studies and Ovid Embase and the Cochrane Library will also be included. Because of the long history of triage and the potentially disparate systems that existed prior to national standardization, the search will be limited to studies published after 1990, four years prior to the implementation of the first nationally recognized triage system.[43]

Inclusion and exclusion criteria

In order to be included in this review, identified literature will need to meet all of the following inclusion criteria:

Setting: The literature must address triage that occurs in an emergency department. Although many services will perform triage to determine who needs care first, the focus of our review is on the narratives generated by triage nurses working in emergency departments. Emergency care can occur in hospital-based emergency departments, freestanding emergency departments, and urgent care centers; which can be referred to in a number of ways: Emergency Room (ER), Urgent Care Center, Accident and Emergency (A&E), Emergency ward, and casualty station.[61]

Narrative data: Included literature must assess triage narratives. These are data that are not collected in a categorical manner; they are in a narrative free form. These can be viewed as analogous to a "chart note".[62] There are multiple phrases used for this, the most commonly used terms are: "summary", "narrative", "free-text or text", "presenting complaint", "syndrome or symptom", or "unstructured description(s)".[23]

Primary Population: Because the primary population of interest is nurses, only narratives generated by nurses will be included. To maintain maximum sensitivity, we define a nurse as someone "authorized by the appropriate regulatory authority to practice nursing",[63] regardless of their registered title ex: Registered Nurses, Nurse Practitioners, Midwife, Psychiatric Nurses, or Practical Registered Nurses.

Study Type: Any peer-reviewed primary literature identified in the search will be included, regardless of design. Reviews, commentaries, conference proceedings, and results available in abstracts form only will not be included.

Language and Full Text availability: Only English language literature will be included.

Literature will be excluded for any of the following reasons:

Disaster specific triage tools: There are a number of different triage tools used during disasters. Triage in these settings is predicated on speed,[64] and does not typically include either a detailed assessment of the patients nor an unstructured narrative of the assessment. As such they will be excluded from this review.

Providers other than nurses: There are other healthcare providers who may perform triage, for example paramedics. While there may be some concordance between different providers[65,66] in order to examine a homogenous practice group, only nurse generated narratives will be included in this review.

Reference & Data Management

Literature identified through the structured searches will be loaded into Covidence (Veritas Health Innovation) where duplicates will be removed, and screening will occur. Title, abstract, and full-text screening will be performed independently by two reviewers using Covidence and will be guided by pre-piloted screening forms (Appendix 2). Any disagreements on inclusion will be settled by consensus, or if needed by a tie-breaker from a third researcher. Citations will be managed using Zotero (Corporation for Digital Scholarship). Continuous and categorical data will be extracted into a pre-built online spreadsheet and word processing programs by the lead researcher and independently confirmed by another using pre-piloted extraction guides (Appendix 3).

Data Extraction

Data extraction will be guided by the central objectives of the study.[58] First, to describe the populations involved in the studies, we will extract data related to: the location of where the studies are being conducted, the number and types of hospitals and departments being assessed, the numbers and types of patients being assessed, and the numbers of nurses involved in the sample and on the research team. Second, to determine the types of literature that has been generated on the topic of interest and the methods used by each of the studies, we will extract data related to the target populations, research designs, and data extraction and hypothesis testing techniques. Finally, we will describe the authors reported purpose for conducting the study or generating the literature, summary of findings or conclusion, and whether any reporting guidelines were followed. The purpose of the literature may be the objectives of the study or rational for position statement or guiding document, the summary of findings will be a summary of the authors' reported conclusion or recommendations. Because AI specific reporting guidelines have only recently been published[67] any reporting guidelines used by the authors will be extracted. The data extraction will be guided by a pre-piloted guiding document (Appendix 3) and will occur in two ways: categorical or continuous data will be extracted into an online spreadsheet and narrative data will be extracted into a word processing document (Appendix 3). Prior to data extraction a sample of 10 studies will be extracted and compared to assess for agreement between reviewers and to determine if the extraction tool needs revision.

Data Synthesis

Data synthesis will aim to map data in a manner that facilitates comparisons and the identification of gaps in the literature.[58] Extracted data will be presented in tabular format to facilitate the summation and comparison of the nature and scope of existing literature to identify promising topics for future systematic reviews aimed to synthesize findings. Quantitative synthesis of continuous variables such as sample sizes will use counts and measures of range and central tendency (mean or median). Categorical data such as year and location of study, the sources of data, approach to data extraction, and methods of hypothesis testing will be expressed as counts and percentages, as the authors' reported purpose for conducting the study and as a summary of their findings will be expressed narratively. Although quality appraisal can be performed,[60] it is not typical to include a critical appraisal within a scoping review as the primary aims are to map the nature and features of the literature, not to provide an unbiased assessment of the findings of the body of literature.[58,59] While we chose to present authors' recommendations from the primary studies in this review, we do so to give a broad sense of what research in the area has generated to date, and to identify areas for future systematic reviews, which will synthesize study findings and also appraise the quality of the studies included in the review.[58] We will, however, report on the proportion of studies that follow formal reporting guidelines. Inter-rater agreement on pilot screening criteria and extraction tools, and final screening and extraction results will be assessed and using Cohens Kappa. PRISMA ScR criteria will be used for reporting findings.[60]

Patient and Public Involvement

No patients involved

Ethics and Dissemination

Ethics

This study will be examining previously published academic literature and publicly available grey literature, as such there is no requirement for ethical approval.

Dissemination

Findings from this scoping review will be presented in several ways: (1) Findings will be submitted for presentation at local, and national academic and emergency nursing conferences. (2) We will request the opportunity to present the findings from this review to the Canadian Triage Acuity Scale and Canadian Emergency Department Information System National Working Groups. (3) A final manuscript will be submitted to an open access peer-reviewed journal for publication.

Manuscript Appendices

Appendix 1. Database Search Strategies

Table 1. Ovid MEDLINE search strategy

1	exp Medical Records Department, Hospital/ or exp Dental Records/ or exp Records/ or records or entries.mp
2	(note* or text*).ti,ab
3	exp "Information Storage and Retrieval"/ or exp informatics / or exp Artificial Intelligence/ or natural language.mp. or exp Language/ or exp Pattern Recognition, Automated/ or exp Neural Networks, Computer
4	(text mining or **data mining or machine learning or natural language processing or semantic* or narrative text*** or text descrip* or **text field***).mp.
5	1 or 2 or 3 or 4
6	emergency treatment/ or emergency medicine/ or exp emergency health service/ or evidence based emergency medicine/ or emergency nursing/ or exp emergency care/ or emergency ward/ or (emergicent* or ((emergenc* or ED) adj1 (room* or accident or ward or wards or unit or units or department* or physician* or doctor* or nurs* or treatment*or visit*))).mp. or (triage or critical care or (trauma adj1 (cent* or care))).mp.
7	exp Nursing/ or exp Nursing Staff/ or nurs*.mp.
8	6 or 7
9	Triag*.mp
10	(Austral* Triage Scale or Canadian Triage Acuity Scale or Canadian triage acuity scale or Emergency Severity Index or Manchester Triage Scale or South African Triage Scale or CTAS or ESI or MTS or SATS).ti,ab.
11	9 or 10
12	5 and 8 and 11
13	limit 12 to (yr="1990-Current" and English)

Table 2. EMBASE Search strategy

1	exp Medical Records Department, Hospital/ or exp Dental Records/ or exp Records/ or records or entries.mp
2	(note* or text*).ti,ab
3	exp "Information Storage and Retrieval"/ or exp informatics / or exp Artificial Intelligence/ or natural language.mp. or exp Language/ or exp Pattern Recognition, Automated/ or exp Neural Networks, Computer/
4	(text mining or **data mining or machine learning or natural language processing or semantics or narrative text*** or text descrip* or **text field***).mp.
5	1 or 2 or 3 or 4
6	emergency treatment/ or emergency medicine/ or exp emergency health service/ or evidence based emergency medicine/ or emergency nursing/ or exp emergency care/ or emergency ward/ or (emergicent* or ((emergenc* or ED) adj1 (room* or accident or ward or wards or unit or units or department* or physician* or doctor* or nurs* or treatment*or visit*))).mp. or (triage or critical care or (trauma adj1 (cent* or care))).mp.
7	exp Nursing/ or exp Nursing Staff/ or nurs*.mp
8	6 or 7
9	Triag*.mp
10	(Austral* Triage Scale or Canadian Triage Acuity Scale or Canadian triage acuity scale or Emergency Severity Index or Manchester Triage Scale or South African Triage Scale or CTAS or ESI or MTS or SATS).ti,ab.
11	9 or 10
12	5 and 8 and 11
13	limit 12 to (yr="1990-Current" and English)

Table 3. CINAHL search strategy

1	(MM "Charting+") OR (MM "Coding+") OR (MM "Nursing Orders") OR (MM "Shift Reports") OR (MM "Medical Records") OR (MM "Information Management") OR (MM "Data Analytics")
2	(MH "Decision Making, Computer Assisted+") OR (MH "Data Mining") OR (MH "Data Collection, Computer Assisted") OR (MH "Data Analysis, Computer Assisted+") OR (MH "Artificial Intelligence+") OR (MH "Minimum Data Set+") OR (MH "Health Informatics+")
3	(MH "Emergency Service+") or (MH "Emergency Medicine") or (MH "Physicians, Emergency") OR (MH "Emergency Nurse Practitioners") or (MH "Emergency Nursing+") or (MH "Emergency Patients") or "casualty department*" or ((emergenc* or "ED") N1 (room* or accident or ward or wards or unit or units or department* or physician* or doctor* or nurs* or treatment* or visit*)))
4	"Austral* Triage Scale" or "Canadian Triage Acuity Scale" or "Emergency Severity Index" or "Manchester Triage Scale" or "South African Triage Scale" or CTAS or ESI or MTS or SATS or (MM "Triag*")
5	1 or 2 and 3 and 4 (restrict to peer review, with abstract, published 1990-2021)

Table 4. ProQuest Search strategy

1	noft(records, hospital) OR noft(records, nursing) OR noft(records, medical) OR noft(medical records, computerized) OR noft(information storage and retrieval systems) OR noft(Informatics) OR noft(Artificial Intelligence) OR noft(natural language) OR noft(machine learning)
2	noft(Emergency Treatment) OR noft(Emergency Medicine) OR noft(emergency medical services) OR noft(emergency service) OR noft(Hospital) OR noft(trauma centers) OR noft(Evidence-Based Emergency Medicine) OR noft(Emergency Nursing) OR noft(casualty department) OR noft(nurse) OR noft(nurs*)
4	noft(Austral* Triage Scale) or noft(Canadian Triage Acuity Scale) or noft(Emergency Severity Index) or noft(Manchester Triage Scale) or noft(South African Triage Scale) or noft(CTAS) or noft(ESI) or noft(MTS) or noft(SATS) or noft(Triag*)

Table 5. Cochrane Search strategy

1	[mh “Medical Records Department, Hospital”] or [mh “dental records”] or [mh “Dental Records”] or [mh ”Records”]
2	(note* or text *):ti,ab,kw
3	[mh “Information Storage and Retrieval"] or [mh “informatics"] or [mh “Artificial Intelligence"] or [mh “natural language"] or [mh “Language"] or [mh “Pattern Recognition, Automated"] or [mh “Neural Networks, Computer"] or [mh “text mining"] or [mh “data mining"] or [mh “machine learning"] or [mh “natural language processing"] or [mh “semantics"] or [mh “narrative text*"] or [mh “text descrip*"] or [mh “text field”]
4	1 or 2 or 3
5	[mh “Emergency Treatment"] or [mh “Emergency Medicine"] or [mh “emergency medical services”] or [mh "emergency service, hospital"] or [mh “trauma center"] or [mh “Evidence-Based Emergency Medicine"] or [mh “Emergency Nursing"] or [mh “Emergencies"] or [mh “emergent*"] or [mh “casualty department*"]
6	(emergenc* NEAR/1 (room* or accident or ward or wards or unit or units or department* or physician* or doctor* or nurs* or treatment* or visit*)):ti,ab,kw
7	[mh “Nursing”] or [mh “Nurs*”]
8	6 or 7
9	[mh ”Triage”]
10	(”Austral* Triage Scale” or “Canadian Triage Acuity Scale” or “Emergency Severity Index” or “Manchester Triage Scale” or “South African Triage Scale” or “CTAS” or “ESI” or “MTS” or “SATS”):ti,ab,kw
11	9 or 10
12	4 and 8 and 11
13	limit 12 to (yr="1990-Current" and English)

Figure 1. Screening Form: Nurses' Narrative Triage Data: A Scoping Review Protocol

First Author (year): **Date of screening:** **Assessor's Initials:**

Article Title:

DECISION: ☐Include ☐Exclude ☐Reference Review ☐Additional Review

Complete this form sequentially. Instructions are written in red:

1. Study type	Does the article discuss emergency department data? If the answer is 'yes', proceed to question 2 If the answer is 'no' mark 'exclude' on the top of the form.
2. Aim	Does the article discuss unstructured narrative data in any way? If the answer is 'yes', proceed to question 3 If the answer is 'no' mark 'exclude' on the top of the form.
3. Population	A. Does the article discuss nurses, nursing, triage documentation, or presenting complaints? If the answer is 'yes', proceed to question 4 If the answer is 'no', proceed to question 3B: B. Does the article discuss unstructured emergency narratives generated by providers other than nurses? If the answer is 'yes' mark as 'reference review' on the top of the form. If the answer is 'no' mark 'exclude' on the top of the form.
4. Study Type	A. Is the article a primary report from published?[1] If the answer is 'yes', proceed to question 5 If the answer is 'no', proceed to question 4B: B. Did the article report on a review (scoping, systematic, narrative)? If the answer is a 'yes' mark as 'reference review' on the top of the form. If the answer is 'no', mark as 'exclude' on the top of the form.
5. Language and Full Text availability	Is the text available in English? If the answer is 'yes', mark 'Include' If the answer is 'no', mark as 'additional review' on the top of the form.

Research Question: To map the literature examining nurses collected triage narratives: where data has been generated, what types of literature examine the issue; and to describe author reported purposes and conclusions.

Review Method: Scoping Review

Definitions:

1. **Emergency Department:** There is a surprising lack of consistency in terminology. Emergency care can occur in hospital-based emergency departments and freestanding emergency departments and urgent care centers. They can be referred to in many ways, these include: Emergency Room (ER), Urgent Care Center, Accident and Emergency (A&E), Emergency ward, casualty station.[1][61]

2. **Unstructured narrative data:** These are data that are not collected in a categorical manner; rather in a narrative free form. These can be viewed as analogous to a "chart note". There are multiple phrases used for this, most commonly: Summary or narrative Free-text or text, Presenting complaint, Chief Complaints, Syndrome/symptom, description(s).[2][23]

3. **Primary Population:** The primary population of interest are nurses. Nurses for the purpose of screening include all skill levels (or bands) of nursing, these include: Registered Nurses (NP), Registered Psychiatric Nurses (RPN), Licenced Practical Nurses (LPN).

4. **Language and Full Text availability:** Only full text English-language literature will be included; conference proceeding and abstracts will be excluded.

References

[1] Khangura JK, Flodgren G, Perera R, Rowe BH, Shepperd S. Primary care professionals providing non-urgent care in hospital emergency departments. Cochrane Database of Systematic Reviews. 2012(11).[61]

[2] McKenzie K, Scott DA, Campbell MA, McClure RJ. The use of narrative text for injury surveillance research: a systematic review. Accident Analysis & Prevention. 2010 Mar 1;42(2):354-63.[23]

Figure 2. Data extraction form and instructions

Data will be extracted from studies into one table to facilitate comparisons. Multiple reports of a single study will be marked and grouped together in the data extraction tables to prevent double-counting the results originating from a single sample or intervention. Please complete all fields. If there are no data for a given field, please mark "Not Applicable" or "Not reported".

Table Items:

1. Study ID: This will include the Author and Year, Digital Object Identifier, or Covidence assigned ID number.

2. Extractor - First and Last initial of the author extracting the data.

3. Sample Data

    Data Source

    a. **Country/State/City:** Where the study data originated from. Extract as much information as possible in this field starting from least to most specific, separating levels with commas. If there are multiple responses at the same level separate them with a forward slash. Ex: Canada, Alberta, Edmonton/Calgary.
    b. **Triage System:** Extract the name of the system, include specifics on versions or electronic if offered. (Ex: CTAS 2018, eCTAS)
    c. **Source of Data:**
        i. **Endorsing Organizations:** Extract the full names of any endorsing organization. (Ex: CTAS guidelines - National Emergency Nurses Association, Canadian Paediatrics Society, Canadian Association of emergency Physicians).
        ii. **Source of Primary Data:** Refers to where the primary data were accessed from. For example, a provincial or national registry, or from hospital electronic records, these can include Regionalized Emergency Data Sets (Canadian Institute for Health Information, National Hospital Ambulatory Medical Care Survey, National Trauma Data Bank, etc.).

    Population

    d. **Number of Emergency departments:** Extract the number of hospitals or urgent care centres reported. If the hospital reports more than one department at one site record each as a unique department (Ex: University of Alberta Hospital Emergency, Stollery Hospital Emergency)
    e. **Number of records:** Extract the number of patient records or visits that underwent final analysis. (Repeat visits by the same patient are counted individually)
    f. **Number of patients:** If the study records the number of providers collecting data record the number included in the final analysis.

4. Study Data

    Design

a. **Qualitative or Quantitative**: Select one, for mixed or multiple methods record "both". For position statements or guidance documents record "not applicable"
b. **Design type**: Record the authors reported design.
c. **Population of interest**: Summarize the authors' population of interest.

Data extraction methods

d. **Automated:** If the authors used any form of automated technology (ex: machine learning, artificial intelligence) to extract data record the type used. (Ex: bag of words, keyword, index searching)
e. **Not automated:** If data were manually extracted summarize the approach used (ex. Expert review and analysis of narrative).

Hypothesis testing

f. **Internal uses:** Describe what the authors used the data for (ex: making predictions about admission).
g. **External comparisons:** If the authors compare narrative data to other data sources (ex: ICD codes) summarize the data external data types.
h. **Statistical testing methods**: Describe the statistical testing techniques.

5. Data collected

Description of data collected

Extract any description of the narrative triage data collected (Ex: number of words in a narrative, number of errors in a narrative, completeness of records, etc.).

What were the authors' objectives?

Extract what the authors report as the primary intention of their study. For primary research this could be the hypothesis being tested or phenomena being described. For guidance documents this could be a position, recommendation, or practice standard.

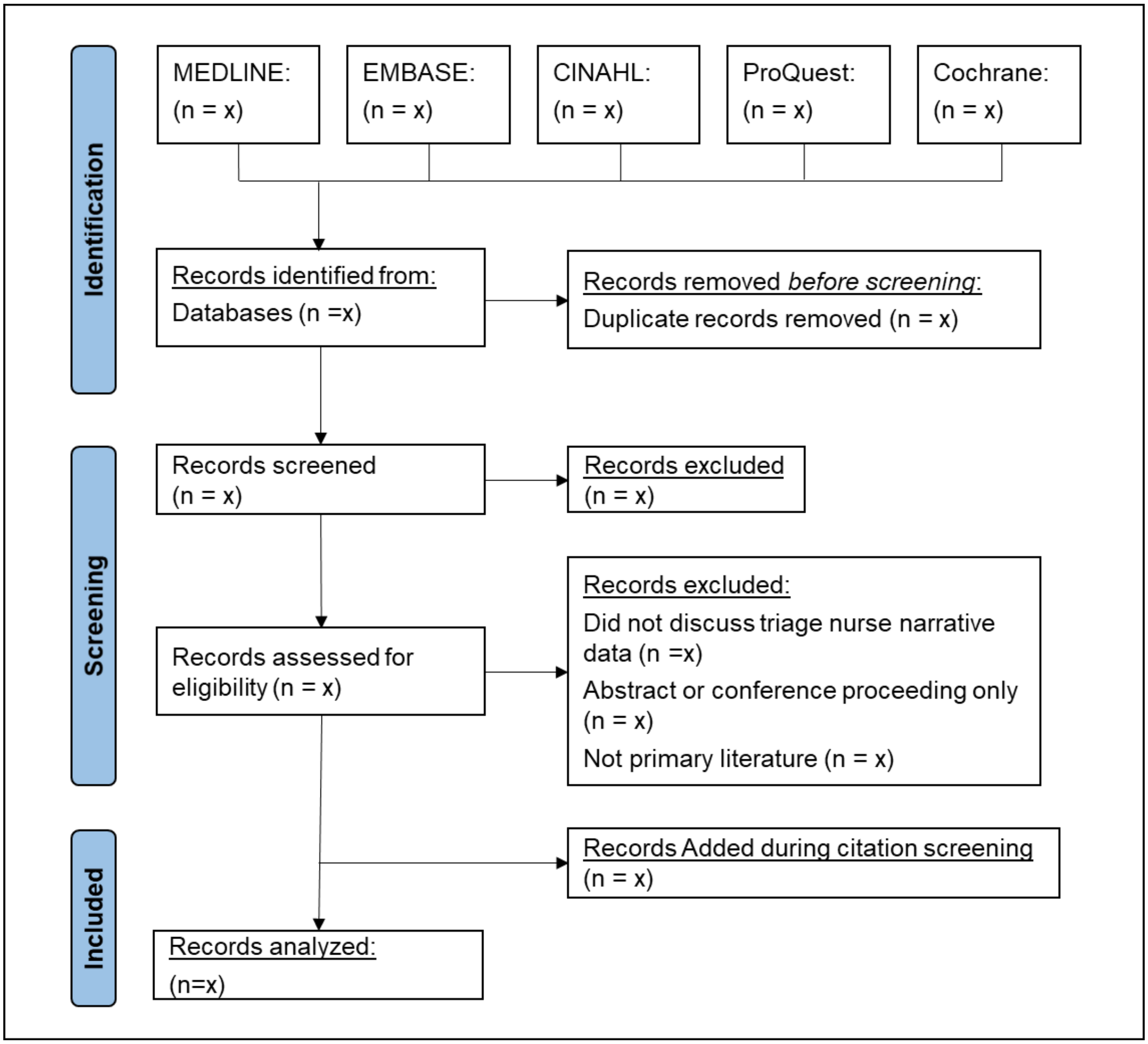
What were the authors findings or conclusions?

Extract any recommendations made by the authors. This could include that further research is needed, or suggestions for future use of narrative data examples could include quality assurance, research, etc.

6. **Presence of quality appraisal:** Extract any reported formal reporting or quality frameworks, or any other quality assessments measures.

1. Study ID	2. Extractor	3. Sample Data	4. Study Data	5. Findings
		Data Source a) **Country/State:** b) **Triage System:** c) **Source of Data:_** *Population* d) **Number of emergency departments:** e) **Number of records:** f) **Number of patients:** g) **Number of providers:**	*Design* a) **Qualitative or quantitative design?** b) **Design type:** c) **The population of interest:** *Data extraction methods:* d) **Automated (specify method):** e) **Not automated:** *Hypothesis testing* f) **Internal uses:** g) **External comparators:** h) **Statistical testing methods (if used):**	*Description of data collected:* *What were the data used for?* *What were the authors' objectives?* *What were the primary findings?* *Were reporting guidelines used?*

Figure 3. PRISMA diagram for the proposed search



Chapter 3. Emergency nurses' triage narrative data, their uses, and structure: a scoping review.

Abstract

Background: Globally, Emergency Departments (EDs) use triage to ensure that the most acute patients are seen quickly and safely. This process is typically nurse-led and results in documentation that can be collected in large databases. Free text narratives generated as a part of this record have been previously studied for specific conditions but never reviewed in a comprehensive manner.

Objective: This paper describes the results of a comprehensive search of unstructured triage narratives using a scoping review methodology. It maps the literature and describes what the data have been used for, what information is available on the form and structure of narratives, and highlights where there are similarities and opportunities for future research.

Results: We screened 18,074 studies published between 1990 and 2022 in CINAHL, MEDLINE, Embase, Cochrane, and ProQuest Central. We identified 96 studies that directly examined the use of triage nurse narratives. There were over 12 million ED visits, drawn from over 63 million health records, generated by 2438 EDs included in the review. More than 80% (n=79) of these studies were performed in the United States (n=43), Australia (n=31), or Canada (n=5). Thirty-nine studies (41%), most of which were published after 2017, used machine learning to incorporate triage narratives into research. Triage narratives were grouped as being used in three ways: for case identification, as input variables for predictive modeling, and for quality improvement. Of the 96 included studies, thirty (31%) described triage narratives in some fashion: twenty-seven (28%) used keywords and seven (7%) offered more comprehensive discussion. There was inconsistent data reporting across studies and only eight (8%) declared using a reporting guideline.

Conclusions: The breadth of studies identified suggest there is widespread routine collection of triage data, and that a significant number of studies examine triage narratives. These data have been used successfully to identify and describe cases rare disease and to perform disease surveillance for more common presentations. They have also been used to inform quality improvement and most recently as inputs for predictive modelling. Analyses of triage data have changed over tie and in 2017 machine learning became the prevalent method for examining triage narratives. Despite the common use of triage narratives as source data in studies these narratives, and the nurses who generate them, are poorly described in the literature and data reporting is inconsistent. Additional research is needed to describe the structure of triage narratives; to determine whether the use of triage narratives in improves the sensitivity of studies looking to describe disease prevalence, and to either develop triage specific data reporting guidelines, or improve the use of existing guidelines. While machine learning methods have been gaining popularity and have recently been incorporated into triage decision support tools, further research is required to determine how these models operate and whether they can be used prospectively to support clinical care.

KEYWORDS nursing; artificial intelligence; machine learning; triage; review, narrative

Manuscript: Emergency nurses' triage narrative data, their uses, and structure: a scoping review.

Introduction

In many countries the Emergency Department (ED) is the most common entrance point for people seeking hospital care. There are an estimated 40.4 ED visits per 100 people;[68] or 142 million visits to Canadian and American EDs per year.[7,68] The large volume of visits requires EDs to sort and prioritize patients. This sorting is accomplished through triage, which is typically the first step in ED care. During triage patients are sorted according to treatment priority and care needs and the process ensures that patients with the most acute needs are seen first. Modern is performed by a health care professional (typically a nurse) using a formally validate tool.[10] Records of the triage interaction, which contains both patient- and visit-related data, entered into extensive health record databases which may be collected nationally. Advancements in electronic health record archiving have expanded the amount of and diversity of data being collected including: ED process timestamps, patient vital signs, screening assessments and triage provider free text descriptions.[69,70] These electronic health record databases are used to track ED volumes and to guide local and national policy decisions.[57] Furthermore, advancements in data storage, machine learning and computer interpretation allow the data to be examined for a range of purposes using a variety of methods.[57,71] Despite the ubiquity of triage and triage-related data collection, the potential impact of triage data research using triage narrative data remains largely unrealized.[1,72]

Background

Triage is the process of sorting patients, originated during the Napoleonic wars,[40] and was introduced into civilian practice in the 1960s.[41] Triage was formalized in the 1980s,[42] and was first implemented as a national system in 1994.[73] Most countries now use a modern formal triage system[10] and these systems have been associated with improved patient safety and service efficiency.[74] The process of ED triaging is a nurse function in most healthcare systems and is typically performed by specially trained and experienced nursing staff.[75] The triage assessment typically consists of a brief history and physical assessment of the patient, followed by the assignment of a visit category and priority level.[75] The patient triage record created in this exchange can include significant amounts of structured and unstructured data.

Data collection in the emergency department

There are several countries with standardized mandatory collection of ED data. Canadian [3,6] and Australian[46] EDs report a triage minimum data set and these both include structured complaint code fields. In addition to these nationally coordinated triage data collection efforts there are also regional databases that collect additional information at triage for local monitoring of injuries or for syndromic surveillance,[76] and these systems have been used to guide responses to local issues such as toxic drug supplies, sporting events, and influenza outbreaks. The range of triage data collected in these systems vary but the data types can be grouped into structured and unstructured data with each type having its own strengths and weaknesses.

Structured data are collected in a pre-formatted fashion, which force the triage nurse to select from one of several options, and restricts the types of data that can be entered into any given data field. Examples of structured data collected at triage would include: arrival times, vital signs, demographic information (age and sex), insurance status, categorical chief complaints, and numerical triage acuity score. These data have been used to track visit volumes, estimate disease prevalence (e.g., The number of patients arriving with influenza like symptoms), and to guide operational decisions such as staffing (such as patient to provider ratios and shift start times). Structured data are by far the most commonly used data that is generated during triage,[57,70]; these data can be used to make predictions about wait times, outcomes, and resource utilization.[57] Structured data are more readily available (due to their routine collection) and simple to work with than unstructured data, but their use is not without limitations. Structured data have been criticized for not being granular enough, both from the perspective of communicating how many nurses collected the data and from the lack of richer contextual detail that can only be gleaned from unstructured narratives.[57]

Unstructured data differs from structured data in its composition; they are typically a free text written note, often called a "narrative",[77] that reflects the triage nurse's assessment and/or the patient's reported reason for visiting the ED. Unstructured narratives collected at triage can offer additional contextual details about: the patient's chief complaint in their own words, descriptions of events associated with the ED presentation, or the nurse's physical exam findings.[78] These narratives may vary in length and structure based on the electronic record system and triage tool used, but would share the common trait of being some of the earliest documented patient information used to inform subsequent ED patient care. Unstructured narratives can be valuable sources of for case identification.

Two systematic reviews focused on injuries,[23,54] have examined whether unstructured clinical narratives (including from triage) could be used for large scale injury surveillance. These reviews, initially performed in 2010,[23] then updated five years later,[54] sought to identify and summarize the extent to which narrative data were used gather injury information, the methods for extracting these data, and to highlight how best to interrogate these data fields. These studies identified 2831 studies over more than 18 years using structured searches and citation searching from 5 included studies. Of the 56 studies included, 13 used ED triage data. This study showed that the use of narrative data has increased with time, and that analyzing the data required automatic or manual extraction of data using keywords or machine learning techniques. The authors compared narrative text derived case identification to gold standard case definitions and expert evaluation and suggested that narratives helped provide important contextual information and identified additional cases. They were critical of the heterogeneity of the narratives assessed in the collected studies, and called for improved data collection methods. Their call for improved consistency in narrative data collection are justifiable, however the heterogeneity they noted may be partially explained by the wide range of administrative dataset types interrogated, and a more homogenous set may have been possible if triage narratives alone were examined.

Two additional review studies, both published in 2013,[55,76] focused their analyses on studies using triage narratives for syndromic surveillance systems, programs that monitor for disease outbreaks. The first, a systematic review, that was very narrow in its search, by Hiller at al., screened 89 studies identified through a structured search of PubMed for syndromic classifiers that were used to detect influenza in ED triage data sets. Syndromic classifiers, which are computer programs that use chief complaint narratives to group patient visits into one of several categories, are used to monitor for changes (outbreaks) in disease related presentations. The authors identified 38 studies that met their inclusion criteria of: i) examining clinical data, that was ii) generated in the ED data, which iii) examined influenza as one of the illnesses being reported on. When they examined the methods used in each study the most common method was to use chief complaints fields with 32 studies using it as a primary means for identifying cases. Hiller et al., remark that data collected in its native format (such as ED triage narratives) allowed for large-scale data collection which, in turn, facilitated comprehensive data collection that was well suited for research and program evaluation. The review, unfortunately, restricted their search to only one database: PubMed, only and likely does not offer a comprehensive view of the literature. Hillier et al., did however highlight that even for a fairly specific presentation pattern there was a significant body of literature to draw from. Unfortunately, they did not offer details on the structure of, or methods for extracting, chief complaint classification data.[55]

The second 2013 review, about syndromic surveillance, by Conway et al., offers significantly more detail on the structure of these syndromic surveillance systems and their data, but their search was not systematic. In this review Conway et al., examined 17 studies drawn from an undisclosed number of initial studies and identified 15 chief complaint classifier systems of interest. These systems were built to monitor for variations in disease prevalence using data that is "nearly ubiquitous…routinely generated…and available electronically".(pg2)[76] The authors describe (geographically) where each system was in use, the process used by each to classify all visits into groups of syndromes, and the relative strengths and weaknesses of each system. The review noted that all but one system (from Canada) was American, that classifiers used differing degrees of computer text parsing to assign patients into groups that ranged from 4-9 syndromes, and grouped the methods of each into either keyword, statistical, or linguistic methods. They highlight how statistical methods (machine learning) are reliant on large volumes of data, but are robust to the "noisy" inputs found in narrative text. Keyword and linguistic methods, by comparison, were described as disadvantaged because regional or temporal variations in triage vocabulary would demand time intensive adjustments to the keyword-based search strategies. These drawbacks are balanced however, by the transparency the method offers, something that is not obvious or easy to articulate when describing machine learning studies. The authors make a compelling argument that triage narratives are of great utility for disease surveillance; they note that syndrome classifiers needed to pre-process data to remove variations. Overall, the authors were far less critical of variations in initial data quality, than other study authors. Whether this is the result of a more homogenous input is not known, and the authors do highlight that there is a need for common syndromic definitions, and that definitions, along with data sharing and aggregation, are needed to improve the utility of these data.

Despite the use of triage data for different purposes, there is a common recommendation from review studies over the last 12 years to improve the consistency and quality of ED triage narratives. There have been some efforts put forth to create common data definitions for these narratives,[79] and to create national ED nursing data sets[80], yet much of the data, including triage narratives, are not as widely collected as structured data;[1] and there is a paucity of literature examining what if any structures are common to triage narratives. This scoping review addresses these concerns and examines the published peer-reviewed literature to identify what ED triage narratives data have been used for, to identify studies that may be sufficiently similar to compare, and to identify where additional research may be needed. In addition, this scoping review systematically examined the research to determine what, if any, structures underlie these narrative data and describe what the data have been used for.

Objectives

The objectives of this review are to:

i) Describe the literature that has been generated using ED nurses' triage narratives,

ii) Describe the objectives and findings of these studies

iii) Determine if there is sufficient data to systematically review the structure or descriptions of triage narratives

iv) Determine if there is enough consistency in the data to perform further review of the included studies for any other outcomes.

Methods

This review uses the scoping framework proposed by Arksey and O'Malley.[58,59] The search, screening, data extraction, and data analyses were published a priori,[2] and the PRISMA-ScR extension was used to guide reporting.[60] To identify studies that examined unstructured narratives in the ED we conducted a search using controlled terminology for the main topics of health records narratives, EDs, and triage. Input from a medical librarian was used to refine the search terms and pre-specified filters were used for the concepts of emergency.[81–84] To maximize the breadth of retrieved studies a comprehensive search of the following databases was conducted: CINAHL, Ovid MEDLINE, Ovid Embase, Cochrane Library (via Wiley), and ProQuest Central. The search was limited to peer-reviewed literature published after 1990: 4 years before the first nationally implemented triage system.[73] Reference lists of select excluded studies: those that examined the free-text narratives of emergency physicians and review studies that included triage narratives, were hand-searched for inclusion.

Data were downloaded into Covidence (Veritas Health Innovation) for screening. Title and abstract screening were performed independently, and in duplicate, by CP and MJD in two stages (title plus abstract, then full text) using pre-piloted screening forms. Any peer-reviewed studies that examined unstructured (free-text) narratives,[23,62] generated within an ED,[61] by a nurse,[63] were included. Studies that examined disaster triage systems, studies that did not have full text (abstracts only) and non-English studies were excluded. Kappa statistics were used to gauge agreement during screening and all conflicts were settled by consensus.

Data extraction

Data were extracted into Microsoft Excel (2019, Redmond, WA) (CP) using pre-piloted forms and results were independently confirmed by a second reviewer (MJD). Counts and proportions are used to describe categorical and numeric values. Categorical values extracted included study variables such as: study design, country of origin, triage system used, and whether machine learning was used. Numeric data extracted included: the year of publication, the numbers of EDs that data were drawn from, the number of visits or patients included in the initial and/or final samples; and the number of nurses examined in each study. For studies that reported data as minimum values (i.e. *"there were over three million of visits"*)[51,85–91] values were recorded as the minimum stated value (i.e. *three million*).When studies reported using quality or reporting frameworks, we report the tool by name. The main conceptual categories of each study: the objectives, design, population, and results were described.[92] We summarized the descriptions of the triage narratives and keywords when the narratives were reported in the study; when five or fewer keywords were used, they were recorded verbatim.

Data analysis

Due to the wide distribution of data our estimates of central measures are made using both median (with interquartile ranges) and minimum and maximum counts. Statistical analyses were preformed using IBM SPSS statistics 25 (Armonk, NY, IBM Corp.). Citation management was performed with Zotero (Corporation for Digital Scholarship). Study objectives were categorized dichotomously (yes/no) for having used machine learning (defined as any form of artificial intelligence), and the study objectives were grouped into exclusive categories according to the primary use of the triage narratives: case identification, as a predictor variable, or for quality improvement.

Results

Overview of articles:

There were 25,091 studies identified in the initial search; after de-duplication 18,074 studies underwent title and abstract screening. Proportionate agreement between reviewers (CP, MJD) during screening was 97.4% for excluded, and 98.1% for included articles (κ= 0.250). 214 studies underwent full text review. One hundred and thirty-one studies were excluded at this stage with the primary rationale being that studies did not specify whether narratives were generated by a nurse or at triage (n=67). All included studies, and any review articles identified in the primary search, that discussed triage or ED narratives underwent citation screening: an additional 13 studies were added at this stage. (Figure 1).

Description of studies

Nearly all of the studies (n=80, 83%) used retrospective designs (Supplement 1). Data were typically drawn (in part or entirely) from electronic databases, except in earlier studies, in which data were manually abstracted from paper charts.[45,93,94] 63 (66%) studies used data housed at the hospital level, 33 (34%) used data extracted from regionalized databases. All studies reported on the unstructured narratives generated at triage; but there was significant variation in the types (and details of) additional data reported. The most commonly collected non-triage-narrative data were patient demographic data, such as: age (n=63, 66%), sex (n=60, 63%), and vital signs (n=29, 30%); visit details, such as: chief complaint codes (n=57), discharge status (n=53), and arrival date (n=35, 59%) and time (n=32, 33%); and ED data, namely: the triage system used (n=41, 43%), (Supplement 2). There was only a weak relationship between the number of items collected and time, with 12% (r^2 = 0.122) of variation being attributable to time (r (94)=0.35, $p<0.001$). 88 studies (92%) reported on the number of EDs included, 78 (81%) on the initial data set size. For included cases 76 (79%) reported these values as the number of visits, 28 (29%) as the number of patients. There were only 9 (9%) studies[45,85,94–100] that reported on the number of nurses who generated the narratives that were used in the study.

The median study size included 12,103 visits (IQR=803; 150,089), or 391 patients (IQR=391; 76,069); drawn from an initial sample of 60,231 visits (IQR= 2,943; 461,435) from 2 EDs (IQR=1; 12) (Table 1). There was a large spread in the numbers of visits and departments examined with the included sample sizes that ranging from less than 100 to over 2 million visits from initial databases ranging from 100 to over 14 million visits drawn from one to 496 EDs (Table 1). There was an increase in the numbers of studies, sample size, and frequency of machine learning use in each 6-year period between 1998 and 2021 with more than 60% (n=59) of studies having occurred in the last 6 years. Median sample sizes increased dramatically after 2009 from 7,951 (IQR= 518; 55,952) to 160,944 (IQR= 19,418; 501,758); so too did the use of ML after 2017, with 30 of the 39 studies (77%) that utilized machine learning occurring in this period (Table 1). We noted that machine learning was used more commonly in predictive studies (21/25, 84%) compared to those using narratives for case identification (17/58, 29%) or quality (1/13, 8%) (Figure 2). Geographically the United States (n=43, 45%), Australia (n=31, 32%), and Canada (n=5, 5%) accounted for over 80% of published papers; only one study was reported from each of South America and Africa (Supplement 3). The studies coming from countries with official languages other than English[96,99,101–104] were from countries that adopted or adapted existing triage systems. Interestingly however other countries with large English-speaking populations are either under-represented (Great Britain, New Zealand) or not represented at all (South Africa, Wales, Ireland, Scotland). (Supplement 1).

Study Objectives

The most common objectives for studies that used triage narratives were to perform case identification (n=59, 61%), to use narratives as a predictor variable in a machine learning model (n=21, 22%), or to use narratives for quality improvement (n=16, 17%) (Supplement 1). Studies categorized with case identification as their primary objective sought to describe incidence or prevalence estimates, or to describe populations of interest. Studies that used narratives as a predictor variable made predictions on patient acuity scores, resource use, or for specific diagnoses by using narrative as a covariate. Quality improvement studies used triage narratives to increase clinician or system safety and were subdivided as pertaining to reliability, accuracy, and validity; or safety and efficiency. Reliability, accuracy and validity studies examine inter-rater reliability or consistency and were used in validating formal triage systems. Tools were also examined for their reliability and consistency in use with special populations or for their performance with specific presentations. Safety and efficiency papers examined narratives to improve data quality, or reduce error and effort (Table 2). Machine learning consisted of several models, and we used an inclusive approach by combining all ML, NLP, and other AI models. We noted the frequency of ML use to be accelerating and that the ML was more commonly used in predictive studies (21/25) compared to those using narratives for case identification (17/58) or quality (1/13) (Figure 2).

Descriptions of triage narratives.

The quality and structure of the triage narratives used as inputs for each of the included studies were poorly articulated. Of the 96 included studies, less than one third (n=30, 31%) described the narrative in any manner. Discussion about the narratives in these studies occurred in one of two ways: as a description of the triage narrative itself, or more commonly, as a description of the keywords used to search within the narrative. (Table 3)

There were seven studies that provided some form of description of the narrative.[51,53,85,89,105–107] Triage narrative descriptions relied primarily on counts of the characters and words used in the typical triage narrative. Length of triage narrative entries ranged from 40 characters[85] to 127 characters (although entries of less than 30 characters were excluded);[107] and 14.6 words[89] to 35 words (including abbreviations)[51] (Table 3). One study described the narratives in terms of clinical features,[53] but it is unclear how these differ from the more commonly reported chief complaints or whether they can be compared to other studies.

There were 27 studies that reported on the specific keywords that were present within the narratives.[22,50–53,85–87,89,106–125] The numbers of keywords ranged from 1[110] to over 160,[51] with a median reported number of 11 keywords (IQR=3-24.5) used (Table 3). However 3 of the 27 studies (11%) did not report the exact number of keywords used.[51,85,86] Some studies did report these data and used express keywords with correct spellings,[108,110,113,116,120,124,125] or intentional variation such as misspellings, abbreviations, or truncations.[51,86,109,112,119]. One study searched for terms using both English and French keywords.[111]

There were nine studies[45,85,94–100] that reported on the number of nurses who generated the narratives.[45,85,94–100] The total number of nurses assessed in these studies was 3,844. The median sample size was 50 (IQR=10,50); but sample sizes ranged from 2[95] to 3,538[98]. These nine studies represent only three percent of the total sample (n=367,946). More importantly however only one study reported on both the structure of triage narratives and reported on the number of nurses included in the sample.[85]

Only one study, by Travers and Haas,[105] explored triage narratives in depth by describing the structure of the narrative and regional variations. This three-centre retrospective cohort study used verbatim triage chief complaint narratives drawn from EDs in the south eastern United

States. Using a corpus of 13,494 unique chief complaint narratives drawn from 39,038 visit they used manual and automated techniques to identify chief complaint concepts using the Unified Medical language System. Concepts that were not readily classified using machine learning models were described in both form and function, and the authors detail the function of the punctuation, acronyms and abbreviations, truncations, modifiers, and qualifier words used in triage narratives.[105]

Although quality appraisal can be incorporated into scoping reviews[60] we did not perform a formal quality appraisal for the included studies because our primary aims were to describe the literature, as opposed to an assessment of each study's findings.[58,59] Consequently, we restricted our quality appraisal assessment to whether or not study's used reporting guidelines. Of the 96 studies, quality reporting guideline use were only reported in 8 studies, all of which were published after 2018. The most commonly used guidelines were Enhancing the QUAlity and Transparency Of health Research (EQUATOR) Network guidelines (Supplement 4). Four studies used reporting guidelines specifically for predictive models,[107,126–128] and only one reported using a quality framework to guide data cleaning and the protection of patient information.[127]

Discussion

Compared to other studies our review screened and identified significantly more studies than previous reviews. From 5 databases we screened over 18,000 studies from a 32-year search period. Previous reviews have search and screening processes that were difficult to replicate, either because of a lack of reporting on the search methods and inclusion criteria,[76] lack of PRISMA diagrams,[23,54] or because triage narratives were combined with other data types.[23,54,55,69–72] To facilitate future reviews of triage narratives we attempted to keep our search strategy as transparent as possible, both by registering our search, and using standardized screening templates, and extraction forms.[2] Our initial search intentionally included a very large number of studies, given the scoping nature of our review, and was intended to maximize the sensitivity of our search. The lack of specificity in our search, as well as the lack of sensitivity in previous searches, demonstrates that identifying triage narrative in the academic literature is difficult and that more straightforward ways of identifying pertinent studies is needed. Authors of studies examining ED triage narratives may need to increase their efforts in clearly flagging this in their keywords, title, or abstract.

In our study the most commonly reported data, besides the triage narrative, were patient: age, sex, chief complaint category, discharge status, and triage acuity. This is quite similar to a 2020 systematic review of machine learning for clinical decision support in the ED which identified the top reported data to be: age, sex, vital signs, triage acuity, and chief complaint category. [57] Similar to other review studies we found an acceleration in the number of studies conducted over time: a 2018 review of clinical information extraction studies by Wang et al.,[69] noted the same increase. Our study adds to this by also showing that there was a sharp increase in the number of included cases after 2008. Our findings also support the observation of Wang et al., that studies utilizing machine learning (they examined natural language processing specifically) lagged behind studies of health record data, and we noted this trend continued until 2017, when it become more common in the literature. Wang et al., tabulated the top sources of electronic health record narratives and determined that of the 58 reported the most common sources were: discharge summaries (n=26, 45%), progress notes (n=15, 26%), admission notes (n=9, 16%), operative notes (n=5, 9%), and primary care notes (n=3, 5%). Triage notes were not mentioned by Wan et al., our review however identified 5,[47,109,129–131] that used machine learning during the same period. Similar to our review, Wang et al. determined that most of the studies were performed in the United States. Unlike our study however, they identified proportionally fewer studies from Australia (1%, n=3/263); our study in comparison identified that over 50% of studies originated from Australia during the same period (n=10/18)   two of which could likely have been included in their study.[95,109] Our results differ in part because, unlike Wang et al., who explicitly examined machine learning, we did not restrict the search in the same manner and rather focused on unstructured narratives as a primary search concept.

Previously discussed reviews and several other studies included in this review established that triage narratives can improve case identification when used in isolation or when added to diagnosis codes.[23] Narrative use for these purposes was reported as a straightforward process in several included studies which demonstrated how the inclusion or exclusion of different data types changed the sensitivity and specificity of searches.[28,47,52] Refinement of these techniques may improve the sensitivity of searches and have significant impacts on disease prevalence estimates, in particular for rare illnesses which may not be well captured with discharge codes,[132] which have been the one of the more common methods of tracking population illness.[133] The methods used in keyword-based case identification studies may be particularly well positioned for these uses given their clearly defined and reproducible methods and long history of being used for these purposes, with studies of this nature being amongst the first studies to use narratives collected on a large scale.[88,105] The potential improvements offered to sensitivity and specificity of case identification may justify a systematic review of the studies included in this review. Additionally, future research could focus on clearly defining, in a reproducible manner, the improvements that narrative data analysis offer for case identification studies.

Predictive studies will likely continue to become a more important are of research and other reviews have examined narratives for these purposes and have noted that the majority of these studies pertain to triage. [57] This review demonstrated that there are a wide variety of predictions that can be made using triage data. On this front nurses are among early adopters given that machine learning techniques are being actively incorporated into triage[53] but it is not certain the degree to which unstructured narratives will influence this prospective use of machine learning support given that there seems to be fewer studies using natural language processing for machine learning studies.[69] Regardless of the methods there is likely an imperative that nurses continue to push for large scale collection of triage data given that studies frequently call for more multicenter collection of nurses data. [57]

The need to collect nursing data has been made before[1] and triage has been identified as one of the most important areas for quality improvement.[134] There have been a number of studies to review QI efforts at triage[72] and calls to include the narratives in these efforts,[135] but significant work is still needed. A renewal of early efforts to establish a minimum ED nurse data set[80] efforts to create common definitions for different narrative elements is needed still needed.[79] Additional research is needed to describe the structures of triage narratives in general. This work is required to determine if there is a common structure to the data as our results show that of the 30 studies that offered a description of the narratives, only one did so with any significant depth. A fulsome description is also required to ensure that the domain specific (triage nursing) contextual data is not lost through text normalization (a typical early step in data cleaning for models) given the unique use of punctuation and abbreviation in triage narratives, a concern raised by Travers and Haas.[105] to Finally, studies need to more clearly what the narrative pertains given that there are regional variations to the breadth and depth of information that is collected at triage.

The most commonly used reporting guidelines: The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) and REporting of studies Conducted using Observational Routinely-collected Data (RECORD) guidelines were published in 2007[136] and 2015,[137] respectively. However, less than ten percent (n=8/96) of studies reported using any kind of reporting guidelines, despite the fact that 86 percent (n=83) of these studies were reported after 2007. Recently published reporting guidelines such as Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD)[138] should help with reporting guideline use, and 2021 saw the highest proportion of studies use a reporting guideline (n=3/13). Use of reporting guidelines will help to reduce the heterogeneity noted in reporting, furthermore the number of studies identified as containing triage narratives may justify the creation of a triage specific reporting guideline.

Limitations

Study search and ID – three terms (ED, triage, narrative) were searched – we used an inclusive search approach so there was a great deal of studies identified in the search phase. The agreement during screening was fair, and agreement likely suffered from the large number of studies reviewed and the due to the need for full text reading in order to determine if the narrative was nurse generated. Future refinements to the search strategy may enable a less wide-reaching search and more clearly defined methods for identifying nurse generated narratives may decrease the need for full text reading, thereby improving the agreement scores during screening. Additionally, clear methods for identifying when narratives are generated by nurses' may prevent the grouping of text from different clinician groups, which may result in more studies being positively identified as originating from triage nurses.

Conclusion

This review identified 96 studies that used triage narratives to perform either: Quality improvement, case identification, or predictions about clinical outcomes. We have described how narrative uses are changing with time to use larger samples and more machine learning methods for interrogating the data. In addition to this we have provided a strong argument that there is a serious lack of research into the structure of the triage narratives themselves, and that the vast majority of studies have not used reporting guidelines. Future research should focus not only on the outcomes of their study, but should endeavour to describe in detail the data source that they use as an input. Future researchers should strive to follow reporting guidelines to improve the quality of data reporting and increase the ability to pool and compare study findings. Emergency nursing scholars have the opportunity to push for national collection of triage data; in doing the data can be compared between regions and common structures of the data better articulated.

Figure 4. PRISMA Diagram

Tables and Figures

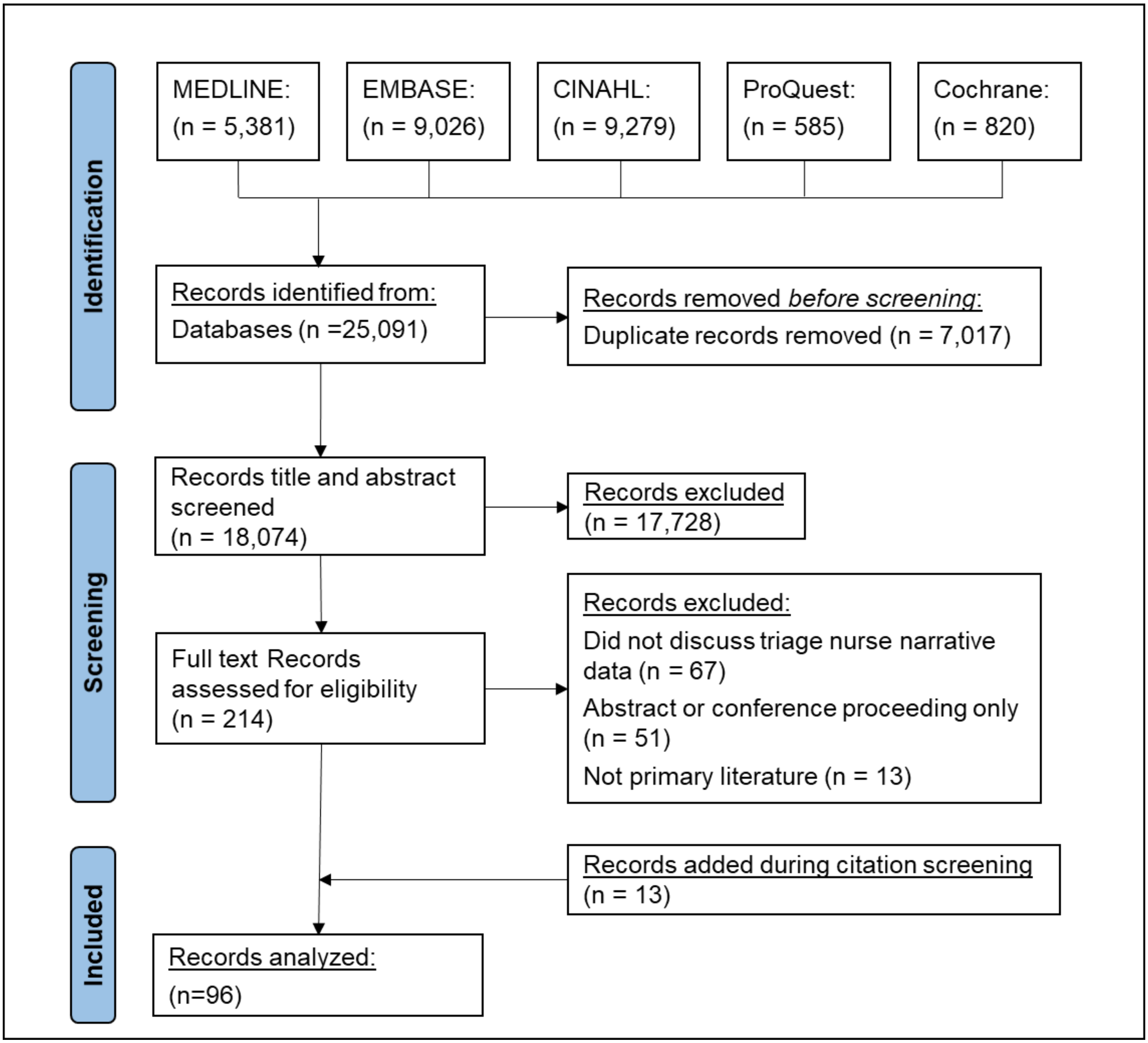


Figure 5. Study characteristics by publication year

Year	Number of studies	EDs included (studies reporting)	Initial sample (studies reporting)	Included visits (studies reporting)	Included patients (studies reporting)	Number of nurses (studies reporting)	Number of studies using ML methods
1998	1	1 (1)	104 (1)	104 (1)	104 (1)	- (0)	0
1999	2	2 (2)	100 (2)	100 (2)	100 (2)	24 (2)	0
2000	0	- (0)	- (0)	- (0)	- (0)	- (0)	-
2001	2	497 (2)	98,672 (2)	84,000 (1)	- (0)	- (0)	0
2002	1	1 (1)	11,861 (1)	- (0)	305 (1)	- (0)	0
2003	2	5 (2)	43,078 (1)	17,413 (2)	- (0)	- (0)	1
2004	3	23 (3)	1,021,949 (3)	21,949 (2)	73,115 (2)	- (0)	2
2005	4	14 (3)	579,032 (3)	1,510 (2)	86,079 (2)	- (0)	3
2006	1	1 (1)	46,602 (1)	45,329 (1)	- (0)	50 (1)	0
2007	1	1 (1)	521 (1)	419 (1)	- (0)	- (0)	0
2008	2	95 (2)	119,479 (2)	5,956 (2)	- (0)	- (0)	1
2009	2	14 (2)	3,556,352 (2)	1,089,984 (1)	389 (1)	- (0)	0
2010	1	2 (1)	263,937 (1)	19,252 (1)	- (0)	- (0)	0
2011	1	6 (1)	794 (1)	794 (1)	- (0)	2 (1)	1
2012	5	182 (5)	12,810,122 (3)	71,427 (4)	519 (1)	27 (2)	1
2013	3	4 (2)	348,895 (1)	41,624 (1)	798 (1)	- (0)	0
2014	3	282 (3)	16,074,953 (3)	43,114 (2)	38,479 (1)	3,538 (1)	1
2015	3	74 (3)	13,051,141 (2)	310,353 (3)	- (0)	- (0)	1
2016	4	109 (3)	13,194 (3)	2,972 (2)	369 (1)	- (0)	0
2017	7	345 (5)	2,450,310 (5)	2,287,592 (7)	- (0)	- (0)	2
2018	9	18 (8)	195,014 (8)	59,801 (8)	183 (2)	10 (1)	3
2019	12	641 (10)	5,453,665 (10)	3,426,182 (10)	153,145 (3)	193 (1)	7
2020	14	29 (14)	4,183,453 (12)	3,372,239 (10)	1,029,147 (7)	- (0)	10
2021	13	92 (13)	3,090,691 (10)	1,318,752 (12)	422,081 (3)	- (0)	6
Sum (studies)	96 (96)	2,438 (88)	63,413,919 (78)	12,220,866 (76)	1,804,813 (28)	3,844 (9)	39 (39)
Median	2.5	2	60,231	12,103	391	15	1
25^{th}	1	1	2,943	803	240	10	0
75^{th}	4.25	12	461,435	150, 089	76,069	50	2
Min, n	0	1	50	29	29	2	0
Max, n	14	496	14,000,000	2,100,000	412,858	3,538	10

Interquartile range (IQR), Machine Learning (ML)

Note: Median, IRQ percentiles, minimum and maximum values are calculated based on individual study sample sizes; results that are reported by year are pooled.

Figure 6. Triage narrative uses

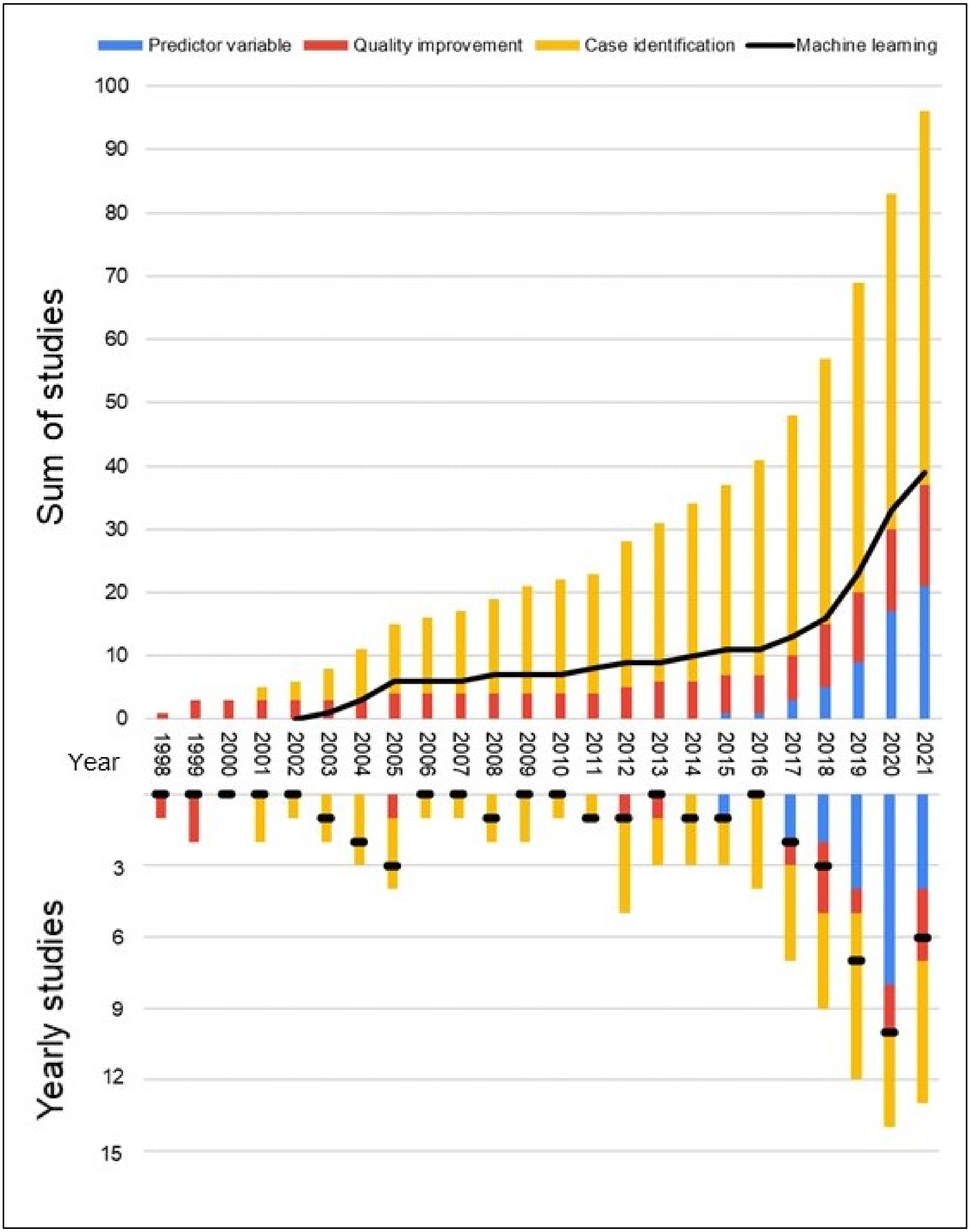


Figure 7. Summary of study objectives.

Study category	Types of papers in category	Explanation
Quality improvement	Accuracy, validity, and reliability	Studies used triage narratives from previous ED visits as a research instrument. These studies would have nurses or physicians re-score visits and compare the scores in order in order to calculate the reliability, validity, accuracy, or agreement or providers or specific triage tools. [45,94,96,99,128,139,140]
	Safety and Efficiency	These studies examined quality as completeness of triage data,[93] how time sensitive presentations were handled at triage,[141] and to identify or improve errors in acuity or category assignment[96,99,128,142]. Other studies focused on improving triage and measured the amount of duplicate or redundant information within triage narratives;[143] or the efficiency,[88,99,100] accuracy,[100] and completeness [48,100,103,144] of narratives.
Case Identification	Syndromic classification	These studies had a primary objective of developing, describing, or comparing syndromic surveillance systems. These systems would attempt to group all patients from a single large cohort into one of several broadly defined groups in order to assign a reason for visit category. [47,85,88–90,95,97,98,105,130,145–148]
	Estimate incidence or describe a population	Triage narratives have been used as an alternate means of identifying general or specific presentations. General grouping included cases related to: drugs or alcohol,[24,28,51,52,86,112,123,125,149–152] sports,[27,48,121,122,153] motor vehicle collisions,[50,87,119,122,124] mental health related presentations, [107,111,154–158], environmental injuries,[91,110,114] infections,[22,26] assaults,[113,159] and animal bites.[91,108,109] Narratives seem to be particularly good at identifying rare cases.[113,116,120,132,160] Narratives have also been used to provide granular data about patients such as: temporal information,[161] to complete missed vitals,[162] and to provide contextual details such as events leading to an injury.[28,50,51,121,122,153,163]
Prediction	Acuity or resource use	Predictions are being made to forecast the resources uses by patient, either in general,[117] of specifically. Specific predictions include: need for admission,[115,131,164–166] triage acuity,[25,53,102,104,106,127] length of stay,[164] critical illness,[127] and mortality,[102,167,168]
	Specific diagnoses	Triage narratives were used as a covariate for machine learning models that made predictions on specific resource or admission needs. Admission destinations and resource of interest included: Advanced diagnostic imaging use,[101,126,169] mental health admission,[118] Intensive Care Unit admission (ICU),[170] or neuro-intensive care unit admission.[171]

Figure 8. Descriptions of the structure of triage narratives.

Study ID	Description of the triage narrative[i]	Number of keywords[ii]	Keyword topics
Travers, DA. 2003.[105]	There is a description of the characteristic components of narrative chief complaints that were not matched by ML: punctuation, truncations, modifiers and qualifiers are discussed	-	-
Chapman, WW. 2004.[22]	-	5	There were specific fever related keywords offered.
Day, FC. 2004.[89]	The mean length of the triage narratives was 14.6 words (SD=7.9) in each database.	3	Shortness of breath and difficulty breathing are discussed.
Thompson, DA. 2006.[85]	The Maximum narrative was 40 characters.	>100	Keywords for chest pain, syncope, earache, and others.
Indig, D. 2010.[51]	The average triage note was 35 words (abbreviations) per presentation; there was a secondary text field not discussed.	>160	Drug and alcohol keywords.
Bregman, B. 2012.[108]	-	2	Mammal bite related terms and their animals.
McKenzie, K. 2010.[109]	-	50	Work, worker, and work-related keywords and truncations.
Vallmuur, K. 2013.[52]	-	18	Alcohol related keywords.
Mitchell, RJ. 2015.[50]	-	32	Alcohol and vehicular collision related keywords
Luther, M. 2016.[110]	-	1	Presentations with the keyword “heat”.
Rahme, E. 2016.[111]	-	16	Suicide related keywords were searched in both English and French.

Study ID	Description of the triage narrative[i]	Number of keywords[ii]	Keyword topics
Whitlam, G. 2016.[112]	-	12	Alcohol related keywords.
DeYoung, K. 2017.[86]	-	>150	Cannabis related keywords.
Kondis, J. 2017.[113]	-	2	“Crying” and “fussy” were the search keywords reported although variations were also included (but not specified).
Morano, LH. 2017.[114]	-	11	Heat injury related keywords.
Zhang, X. 2017.[115]	-	25	A list of keywords predictive of patient admission.
Chu, KH. 2018.[116]	-	1	“Headache”.
Gligorijevic, D. 2018.	-	24	Mixed keywords for a variety of presentations.
Goldman-Mellor, S. 2018.[118]	-	8	Mental health and substance use related keywords.
Hargrove, J. 2018.[119]	-	23	Vehicle collision related keywords.
Negabhushan, M. 2018.[120]	-	2	Specific chest pain feature keywords “ripping” and “tearing”.
Chen, M. 2019.[121]	-	2	“Tramp” and “bounce” are specified but other terms may have been used.
Eley, R. 2019.[122]	-	14	Bicycle related keywords.
Marx, GE. 2019.[123]	-	8	Marijuana related keywords.

Study ID	Description of the triage narrative[i]	Number of keywords[ii]	Keyword topics
Trivedi, TK. 2019.[124]	-	3	Electric-scooter related brand names “bird” and “lime” as well as “scooter” were the keywords.
Sterling, NW. 2020.[106]	The mean length of triage narrative was 143.17 (SD 77.8) characters (excluding spaces) or 64.3 (SD=35.2) words in length.	-	-
Vernon, N. 2020.[87]	-	3	Electric-scooter related keywords and their variations were searched. “Scooter”, “e-scooter”, and “electric-scooter” are offered as specific terms. .
Ivanov, O. 2021.[53]	The average clinical features per text entry was 12.79. There was no discussion about character or word counts.	-	-
Rahilly-Tierney, C. 2021.[125]	-	3	“Heroin” and “overdose” were specified as inclusion terms, “detoxification” as an exclusion term; although there may be additional terms included but not specified.
Rozova, V. 2021.[107]	The average triage note was 127 characters (notes with less than 30 were excluded).	40	Suicide related keywords.

i – Studies reporting only the process of cleaning and normalizing unstructured narratives are not included.

ii – Variations in spelling, abbreviations, bigram duplications, and negation terms are counted if specified.

Figure 9. Summary of included studies

First author, Study year	Country	Narrative use	Study objective	Summary of findings
Kabir, A. 1998.[40]	AUS	Quality improvement	To compare the completeness of documentation occurring with paper charting versus computer assisted triage.	This retrospective observational study of consecutive children presenting for acute asthma to an ED compared the degree of information gathered during computer assisted triage it to non-computer triage and physician documentation. Using descriptive statistics, and kappa values they reported on: visit characteristics, triage details, nursing observations, medical details, asthma severity, nursing observations, determining these were better recorded in the paper chart than the computer chart.
Beveridge, R. 1999.[41]	CAN	Quality improvement	To examine the reliability of the triage acuity assigned by different observers.	This retrospective analysis of routinely collected data with prospective rescoring for validations examined 50 patient records (10 from each acuity level) and re-scored them using summarized visit details that included: the presenting complaint, mode of arrival, vitals, and the verbatim triage narrative. Between and within group comparisons (nurses, physicians) were made using kappa coefficients, and 1 and 2-way ANOVA testing. Between group and within group acuity scores had high degrees of agreement, ANOVA testing showed that physicians assigned higher triage scores overall.
Goodacre, SW. 1999.[42]	GBR	Quality improvement	To perform triage quality improvement by examining interrater agreement on triage.	This retrospective used manual chart reviews to rescore triage visits and compared triage nurse accuracy. Full chart data (including triage narratives) were used to reassign triage acuity levels. Acuity scores were compared using Kappa, Sensitivity, Specificity, NPV, PPV, Sensitivity/recall and F-scores. Kappa scoring showed fair to moderate inter-rater agreement. Predictive power of scores did not improve with after developing of triage guidelines although there were modest improvements in agreement scores.
Aronsky, D. 2001.[35]	USA	Case identification	To reduce the frequency of free text triage chief complaints narratives by developing a list of coded chief complaints.	This was a three-step quality improvement project that used retrospective triage narrative data to develop a chief complaint category list, prospectively examined whether the list reduced the frequency of free text chief complaint documentation, and described staff satisfaction with the project. Using descriptive statistics and control charts the authors described the final complaint list, frequency of free text narrative charting, and responses to Likert scale questionnaire responses. The intervention resulted a final chief complaints category list that included 54 coded variables, free text charting decreased

First author, Study year	Country	Narrative use	Study objective	Summary of findings
				from 23% to 1%, the number of staff using free text decreased from 45% to 9%, and ED staff found the intervention beneficial.
Burt, CW. 2001.[91]	USA	Case identification	To estimate and compare the incidence and characteristics of sports and non-sports-related injuries to USA EDs.	This retrospective cohort study used the triage narratives from visits identified using diagnosis codes to describe the accuracy of injury diagnosis codes and assign activities. They found that males were more than twice as likely to have sports related injuries in the 5-24year old patients' group (48.2 versus 19.2/1,000 persons). Basketball and cycling were the most commonly noted activities associated with injuries.
Howe, A. 2002.[92]	GBR	Case identification	To identify victims of assault, to compare the number of identified cases to the number recorded by police, and to identify the location and method of assault.	This was a retrospective observational study of patients who presented to the ED after assault. The authors interrogated physician and triage nurse narrative fields using "assault" keywords to identify cases. Cases identified using narrative field searches were compared to those identified using hospital and police maintained categorical codes. Descriptive statistics and Mann-Whitney U were used to describe and compare: the number of cases identified, location of assault, method of assault, and rates of identification between groups and methods. 2.6% of ED patients presented after an assault34% of cases were identified by both triage nurses and physicians. Triage nurses identify 57%, physicians 77% of assaulted patients. Differences in detection were not related to the patient's age or sex.
Begier, EM. 2003.[93]	USA	Case identification	To compare the triage narrative derived syndromic chief complaint code to discharge diagnosis codes.	This retrospective cohort study reviewed the syndromic codes that were assigned by two expert reviewers (using the triage narratives) to the recorded discharge diagnosis. The overall agreement for the syndromes of: "death", "gastrointestinal", "neurological" "rash", "respiratory" "sepsis", "unspecified infection", and "other." To the discharge diagnosis varied significant kappa values raged from 0.085 to 0.684 for different syndromes.
Travers, DA. 2003.[73]	USA	Case identification	To use machine learning models to build a concept-oriented set of ED nursing terminologies from triage narratives.	This was a prospective comparison of different natural language processing models that used retrospectively collected triage narratives to create triage "concepts". Expert evaluation was used to examine the model's accuracy in mapping concepts to Unified Medical Language System concepts. the authors detailed the steps used to normalize data and used descriptive statistics to describe the number of concept matches made after each round of narrative data normalization.

First author, Study year	Country	Narrative use	Study objective	Summary of findings
Chapman, WW. 2004.[51]	USA	Case identification	To compare the sensitivity and specificity of three machine learning models for detecting fever from triage narratives.	This study used retrospective electronic health record data to compared the sensitivity and specificity of three machine learning algorithms for detecting fever from triage and discharge narratives. Using patients identified with an ICD-9 discharge diagnosis of fever models were trained to use narratives and coded chief complaints to detect fever. Sensitivity, specificity, and positive likelihood ratios were used to compare different models to expert clinician ability to detect fever. Machine learning models using narratives were most sensitive, models using keywords the most specific when compared to expert clinicians.
Day, FC. 2004.[36]	USA	Case identification	To use word clusters to automatically link free text with reason for visit categories based on ICD codes.	This was a retrospective of ED data from a state database that used word clusters to automatically link triage narratives with reason for visit categories. Using patient age, sex, free text chief complaint, and diagnosis (based on ICD codes) cases were identified. Descriptive statistics were used to describe salient keywords identified, the character count of narratives, and the percentages of correctly coded chief complaints by age. Sensitivity/specificity analyses were used to build inclusion and exclusion keyword clusters and to compare clusters to ICD code categorization.
Mikosz CA. 2004.[80]	USA	Case identification	To compare the agreement between two coding programs and their associated free text triage notes.	This was a retrospective cohort study that compared the triage chief complaints assigned to free text narratives by two different machine learning algorithms. Overall agreement varied by syndrome (Kappa = .419 - 0.711).
Baumann, MR. 2005.[94]	USA	Quality improvement	To determine the reliability and validity of ESI for pediatric populations.	This study retrospectively examined triage EHR data and had to assess the reliability and validity of ESI scores for pediatric patients. Patient and visit characteristics were described. The relationship between resource use and ESI scores were examined, inter-rater agreement of re-triaged scores. Weighted agreement ranged from good to excellent. Hospitalization, ED length of stay, and resource utilization were strongly associated with ESI category.
Chapman, WW [1]. 2005.[81]	USA	Case identification	To evaluate a machine learning algorithm's ability to classifying free text data into syndromic categories for	This retrospective observational study examined EHR data to compare physician "gold standard" coding to a machine learning model's ability to classifying free text data into syndromic categories. machine learning classifier AUC scores ranging from 0.95 to 1.0.

First author, Study year	Country	Narrative use	Study objective	Summary of findings
			surveillance purposes.	
Chapman, WW [2]. 2005.[82]	USA	Case identification	To examine a machine learning programs accuracy in classifying syndromic categories and to determine if the program can be used to examine chief complaints in triage narratives.	This retrospective observational study used ICD codes from EHR data to determine if a machine learning model could accurately classify syndromic categories using triage chief complaint narratives. The models accurately classified 92% of neurologic and respiratory syndromes and 99% of botulinic syndromes. The models sensitivity ranged from 30% to 75%, with specificities greater than 90%, and positive predictive values of 12% to 44%.
Muscatello, DJ. 2005.[95]	AUS	Case identification	To test a syndromic surveillance system based on routinely ED data.	This study describes the development and implementation of an automated near real-time syndromic surveillance using routinely collected ED data. The evaluation compares a machine learning model that examined patient demographic, presenting complaints, and triage narratives to physician-assigned ICD and departure codes. The machine learning model assigned syndromic classifiers to a greater proportion of ED visits 93% (range: 58% – 100%) compared to the availability of physician completed ICD codes 76% (range: 68% – 86%). During the assessment period the EDs had a median of 997 (range: 941–1077) visits per day and there were no outbreaks recorded.
Thompson, DA. 2006.[31]	USA	Case identification	To use word clusters to automatically link free text with reason for visit categories.	This retrospective study used EHR data to automatically link free text triage narratives word clusters with reason for visit categories based on ICD codes using a machine learning model. Demographics and visit characteristics are described; 87.5% of visits were successfully classified.
Gillam, C. 2007.[96]	AUS	Case identification	To examine an injury surveillance system and compare nurse injury coding to the narrative text in order to determine levels of agreement and sources of error.	This study assessed the validity of an injury surveillance system using EHR data for all patients presenting with injuries. Researchers extracted narrative description of the circumstances leading to the presentation, intent, and cause using triage nurse narratives. Agreement between the triage nurses' coding and those assigned using the narrative were compared for each data element. Of the included cases (n=419) triage nurses and expert agreement was 91.9% for intent and 79.2% for cause.

First author, Study year	Country	Narrative use	Study objective	Summary of findings
Indig, D. 2008.[97]	AUS	Case identification	To compare provisional medical diagnosis and nursing triage text for detecting alcohol-related ED presentations and to explore visit related data to examine why alcohol-related presentations are under-detected.	This study compared two methods for detecting alcohol-related ED presentations: diagnosis codes and nursing triage text, and compared patient and visit characteristics examine missed formal diagnoses and describe why alcohol-related ED presentations are under-reported. Triage free-text fields, retrieved from a surveillance database, were searched using alcohol-related terms. Multivariate Log regression analysis of demographics, diagnostic info, service delivery characteristics, and free text (triage) was used to identify factors predictive of medical versus triage identified cases. Approximately 4.5% of ED presentations were alcohol-related, 24% were identified through diagnostic codes and by triage narrative. Presence of diagnostic coding was more for patient arriving by ambulance or those with signs of aggression; and less likely for patients with injuries.
Irvine, AK. 2008.[98]	USA	Case identification	To describe a program designed to extract temporal information from triage notes.	This study reports on a program designed to extract temporal information from triage notes. It prospectively compares a machine learning model to an expert reviewer's ability to extract information from triage narratives. The most common class of time statements were relative (e.g., 1 week ago), there was perfect agreement between time coders, decision tree out-performed models outperformed naive bayes machine learning in precision and recall while assigning data.
Indig, D. 2009.[99]	AUS	Case identification	To examine different methods for detecting alcohol-related ED presentations and to compare these patients to ED patients identified as risky drinkers by a questionnaire.	This multiple step study compared different methods (triage narratives, ICD codes, and questionnaire) for identifying alcohol-related ED presentations and to compare the characteristics of each group. Proportions of patients identified using each method were compared and multivariate logistic regression was used to determine which characteristics are predictive of alcohol related presenting complaints. ICD codes had a higher specificity and lower sensitivity than triage narratives for identifying total number of presentations and percentage of risky drinkers presenting for an alcohol related complaint.
Mitchell, R. 2009.[100]	AUS	Case identification	To identify sports injuries from triage narrative text.	This retrospective cohort study sought to identify sports injuries using triage narratives from a national ED database. A 5% sample of cases flagged using keywords were manually audited to confirm true and false positive rates for each keyword used. Within the sample some sports had higher precision (Ex. Rugby) than others (Ex. Squash).

First author, Study year	Country	Narrative use	Study objective	Summary of findings
Indig, D. 2010.[32]	AUS	Case identification	To use triage narrative to detect drug and alcohol related presentations and describe their service delivery characteristics.	The purpose of this study was to use retrospective EHR data (triage narratives) to detect drug and alcohol related presentations and describe their service delivery characteristics. The triage narrative identified 90% of drug and alcohol related presentations, physician assigned ICD detected 21% of drug and 25% of alcohol related visits.
Wagholikar, AS. 2011.[101]	AUS	Case identification	To use a rule-based approach to classify unstructured triage narratives into symptom groups.	This study compared the sensitivity and specificity of a text classifying algorithm to expert categorization in identifying chest pain from triage narratives. the algorithm was 99.3% sensitivity, 80.0% specificity, and had an F-score of 0.9 when compared to expert assignment for identifying chest pain.
Bregman, B. 2012.[52]	USA	Case identification	To use triage narratives to identify and characterize visits to EDs for animal bites.	This retrospective cohort study identified animal bite rates using triage narrative searches for the word "bite". Over 6,000 animal bite visits were identified on a yearly basis. 70% of animal bites were from dogs, 13% from cats; the characteristics of visits were similar to other surveillance systems.
Grossmann, FF. 2012.[43]	CHE	Quality improvement	To examine the validity, reliability, and accuracy of triage acuity scores for adult patients and to describe the reasons for acuity errors.	This prospective cohort study examined triage data (original narratives) retrieved from hospital records of patients greater than 65 years of age. Visits were rescored by two blinded triage experts for agreement, acuity scores were compared to resource use, length of stay, disposition, need for lifesaving intervention, and mortality. Agreement between original score and expert re-triage was high (weighted kappa=0.934, 95% CI 0.913 to 0.954). Under triage occurred in 117 cases, with inaccurate interpretation of vital signs being the most common cause.
Malmström, T. 2012.[102]	FIN	Case identification	To classify use free text and diagnostic codes to produce ED specific complaint classification codes.	This retrospective descriptive study with prospective validation and focus group interviews. The first phase described the process used to produce an algorithm that groups complaint categories using triage narratives and diagnostic codes. The second phase prospectively examine the implementation of the algorithm. The final phase was evaluated by an expert panel. The presenting complaints summarized from free text fields were similar to those of the algorithm. The algorithm included 89 presenting complaints and ED staff found it easy to use.
McKenzie, K. 2010.[53]	AUS	Case identification	To compare different methods for identifying potential work-related injury	This retrospective cohort study described and compared methods for identifying potential work-related injury cases using triage narratives. It examined the sensitivity, specificity, and positive predictive value of different keywords used in

First author, Study year	Country	Narrative use	Study objective	Summary of findings
			cases using textual data and to compare the predicative power of each method to coded injury surveillance data for work related injuries.	text interrogation against coded injury surveillance in a national database.
Rhea, S. 2012.[103]	USA	Case identification	To use routinely collected ED data to characterise heat related injury visits.	This retrospective cohort study used routinely collected ED data and weather data to characterize heat related injury visits. Visits containing a heat-related illness ICD-9 code has their chief complaint and triage narrative data examined for potential exposure categories. Descriptive statistics were used to describe visit characteristics and distributions, patient characteristics, and regional temperatures. Visits were most likely in June, Heat related visits increased by 1.4 for each 1-degree Fahrenheit) increase from 90F to 98F and by 15.8 for each 1F increase from 98F to 100F. Distributions of causes (sport or work) were age related. Admission rates increased after age 65.
Genes, N. 2013.[162]	USA	Case identification	To transform free text vital sign data into numeric scores.	This retrospective cohort and prospective derivation study examined an algorithm's ability to transform narrative triage vital sign data into numeric scores. 98% of values assigned from narrative data had perfect quality scores. Errors rates did not increase with periods of high triage throughput.
Mosley, I. 2013.[105]	AUS	Quality improvement	To describe the triage patterns associated with using rapid care protocols in patients who present within two hours of acute stroke symptom onset.	This retrospective observational study investigated factors associated with triage acuity assignment and initiation of stroke protocols for patients presenting to the ED within 2 h of symptom onset. Using EHR data researchers identified patients with an ED diagnosis of stroke or TIA and using logistic regression examined demographic and situational factors (including triage narratives) associated with a triage category of 1 or 2. 94% of patients diagnosed as a stroke or TIA who presented within 2 hrs were correctly identified. In all cases not correctly identified the patient was diagnosed with intracranial haemorrhage.
Vallmuur, K. 2013.[54]	AUS	Case identification	To compare different approaches for identifying alcohol involvement in	This retrospective cohort study compared methods for identifying alcohol involvement in injury-related youth ED presentations. Using triage narrative keywords and ICD codes researchers searched triage narratives to determine which keywords were most frequently

First author, Study year	Country	Narrative use	Study objective	Summary of findings
			youth who present to the ED for injuries and to describe the text used to triage these patients.	used and identify visit characteristics and patient demographics associated with alcohol use. 6.4% of injury presentations had documented of alcohol involvement, patients 18 to 24-year-old, females, indigenous youth, Saturday or Sunday presentations, and presentations between midnight and 5 am were more likely to be alcohol related. Generic alcohol terms such as ethanol or alcohol were the most common terms used in the triage narrative.
Haas, SW. 2014.[106]	USA	Case identification	To develop and test a syndromic classifier using triage notes and chief complaints.	This was a retrospective cohort and model validation study. It used previously collected and classified syndromic database data to update existing syndrome definitions. Using triage notes and coded surveillance data that was manually classified by three clinicians' researchers used the chief complaint, triage narrative, diagnosis, vital signs and admission status to determine if the updated definitions correctly classified visits into syndrome groups. Descriptive were used to describe triage narratives. Agreement was measured using kappa statistics. Performance measured using sensitivity and specificity. The triage narratives contained an average of 22 words in the initial dataset; 25 words (151 characters) in the follow-up dataset. Kappa for the two studies were 0.76 and 0.82 respectively. The new definitions improved the sensitivity and specificity for each syndromic cluster.
Liljeqvist, HTG. 2014.[107]	AUS	Case identification	To compare the accuracy of different methods of identifying mental health-related ED visits.	This retrospective cross-sectional study compared the accuracy of using different classification methods identify mental health-related visits using EHR data. The accuracy of different methods of identifying models: triage narrative data, ICD codes, and SNOWMED Codes were evaluated by comparing their case identification rates to expert coding using descriptive statistics, kappa scores, and by calculating each model's sensitivity and specificity. Agreement between clinician's classification and model classification ranged from moderate to almost perfect (Kappa 0.73, 95% CI 0.58 - 0.87), models were moderately sensitive (68% 95% 95% CI 46%-84%) and highly specific at (99%, 95% CI 0.98-0.997). Positive predictive value was 81% (95% CI 0.57 – 0.94) and negative predictive value was 98% (95% CI 0.97-0.99).
Rhea, SK. 2014.[38]	USA	Case identification	To describe the incidence of animal bite injuries.	The aims of this retrospective cohort study were to describe the incidence of animal bite related ED visits. Using surveillance codes, chief complaint codes, ICD codes, and triage notes

First author, Study year	Country	Narrative use	Study objective	Summary of findings
				researchers coded visits by animal (bat, bear, beaver, chipmunk, coyote, fox, groundhog, hedgehog, mole, opossum, otter, prairie dog, raccoon, skunk, squirrel, wolf, or woodchuck) and determined the incidences rates for each. By age of 10, patients had a 1 in 50 risk of dog bite injury. Dog bite rates were highest for patients ≤ 14 years of age; cat bites and scratches were most common in patients > 79 years of age. The lifetime risk of a cat bite or scratch (requiring ED care) was 1 in 60. Postexposure rabies prophylaxis was given to 1,664 of 38,971 visits.
Handly, N. 2015.[83]	USA	Predict outcomes	To determine whether machine learning algorithms using coded triage chief complaint data outperformed algorithms that did not in predicting hospital admission.	The objective of this two phase (derivation and validation) retrospective cohort study was to determine whether machine learning models using coded triage chief complaint data outperformed algorithms that did not in predicting hospital admission. In the derivation phase a machine learning model identified 213 chief complaint codes from the triage narratives, and identified other predictor variables such as: age, sex, race, visit characteristics (arrival time and day) and triage acuity. In the validation phase the sensitivity/specificity of models without and with coded triage complaints were 64.0% (95% CI 63.7–64.3)/87.7% (95% CI 87.4–88.0), 59.8% (95% CI 59.5–60.3%)/91.7% (95% CI 91.4–92.0) respectively. In the derived phase the sensitivity/specificity models without and with coded triage data were 60.7% (95% CI 60.4–61.0)/87.7% (95% CI 87.4–88.0) and 59.8% (95% CI 59.5–60.3)/90.6% (95% 95% CI 90.3–90.9) respectively.
Hides, L. 2015.[108]	AUS	Case identification	To describe the characteristics of alcohol related ED presentations.	This retrospective cohort study describes the visit characteristics of alcohol related injury presentations in the ED. Keyword searching of the triage narratives identified 7,381 of 12,264 cases. Descriptive statistics were used to report frequencies and proportions. Chi-square values were used to compare groups according to use of alcohol, demographics, location of injury, visit characteristics (time/date), and triage urgency. The patient cohort identified using triage narratives accounted for 38% of alcohol related injury presentations. The most common cause of injuries was falls and violence in males; intoxication and self harm in females.
Mitchell, RJ. 2015.[55]	AUS	Case identification	To use surveillance database data to identify injury	This retrospective interrogated a nationally injury surveillance data base with various keywords then manually reviewed one percent of identified cases to estimate the likelihood of true positives. They

First author, Study year	Country	Narrative use	Study objective	Summary of findings
			risk factors of road users by age group.	also compared collision characteristics (e.g., vehicle speed) and injury risk factors (e.g., non-restraint use) by age group and road users' type (i.e., motor vehicle drivers, motor vehicle passengers, motorcyclists, pedal cyclists and pedestrians). Descriptive statistics were used to describe demographics, and number of cases identified with each keyword.
Gray, SE. 2016.[109]	AUS	Case identification	To examine the completeness of triage data collected for fitness-related injuries.	This retrospective cohort study of the Victorian Emergency Minimum Dataset used fitness facilities injuries to examine the completeness of coded data. Using an anonymized ten-year sample of data on patients the authors created new variables (degree of specificity, location, activity, and product) using narrative data. Descriptive statistics were used to compare the frequency of each new variable. Cross-tabulations were used to compare the narrative coded data to existing categorical codes. Of the 2,936 identified cases: two percent could not have any additional information coded, and 95.8 % had at least one piece of information missing. There was injury type and body region coding in 92.6 and 96.6 % of cases, but only in 27.1 and 75.4 % of narratives. The causal variable was specifically coded in 47. 7 % of cases and in 45.9 % of narratives.
Luther, M. 2016.[56]	AUS	Case identification	To compare triage visits to environmental data to determine whether a change in heat warning threshold would change the predictive power of a surveillance tool.	This retrospective chart review of electronic hospital records used keyword searches of triage narratives to identify heat-related presentations. Heat-related presentation volumes were compared to the mean temperature of the three preceding days. Patient demographic and visit related details where collected. Using these data, the authors determined that temperatures above 30C were associated with increased risks, that extremes of age and outdoor activity was associated with increased risk. The symptom most strongly associated with presentations was syncope.
Rahme, E. 2016.[57]	CAN	Case identification	To describe the visit characteristics of patients who present to the ED for a suicide attempt, to identify factors associated with admission, and to validate ICD-10 codes as a method	This retrospective cross-sectional study used ICD codes and triage narratives to identify patients who were seen in the ED and hospitalized following a suicide attempt. The purpose of the study was to describe patient and visit characteristics, identify factors predictive of need for hospitalization, and to compare the number of cases identified using ICD (10) codes for ''intentional self-harm". 5746 cases were identified; 369 were fully reviewed, 281 (76%) were identified using the triage keyword method. Of these, 176 were treated in the ED and 193

First author, Study year	Country	Narrative use	Study objective	Summary of findings
			to detect suicide attempts in routinely collected data.	required hospitalizations, 46% of patients received an ICD intentional self-harm code. 46% of cases were poisonings (the most frequent method), half of patients were less than 34 years old, 53% were female, and 75% had a previous history of mental health complaints.
Whitlam, G. 2016.[58]	AUS	Case identification	To evaluate the precision of routinely collected ED data in identifying acute alcohol-related harms.	This retrospective study randomly sampled 1,000 ED visits with an ED diagnostic code of alcohol harms and reviewed the triage narrative to confirm the diagnostic code and classify the visit as 'acute' or 'chronic' harm. Predictive factors for alcohol related presentations (ex, age, sex, time of arrival, etc.) and the predictive value of triage narratives were for acute harm were calculated. The PPV of diagnostic codes for acute alcohol harm was 53.5%. Variables predictive of acute harm were: ambulance arrival (aOR = 3.4, 95% CI 2.4–4.70, age (12–24 vs. 25–39 years: aOR = 3.4, CI 2.2–5.3), admission (not admitted) status (aOR 2.2, 95% CI 1.5–3.2), and arrival time (between 2200hrs and 0600hrs) (aOR 2.1, 95% CI 1.5–2.8).
Berendsen Russell, S. 2017.[110]	AUS	Quality improvement	To describe a method for reducing duplicate and redundant clinical terms in ED documentation systems.	This retrospective study describes a method for reducing duplicate and redundant information in ED electronic charting systems. Using narratively entered chief complaint data researchers identified 64,849 unique complaints from 1.7 million visits. Of these complaints 450 terms were used more than 100 times. Of the terms used 177 (39.3%) matched current definitions. The authors categorically grouped complaints into alternate clinically meaningful groups: cardiovascular (chest pain, arrhythmias), respiratory (shortness of breath, cough), gastrointestinal (abdominal pain, vomiting), and injury (fractures, trauma).
DeYoung, K. 2017.[33]	USA	Case identification	To use structured and free-text emergency data to develop and validate a syndromic definition for marijuana related ED visits.	This retrospective cohort study used structured triage variable and unstructured narratives to develop a syndromic case definition for marijuana related ED visits. Different ways of identifying cases were compared (triage narrative, discharge codes, discharge narrative, chief complaint, clinical impression) were compared to expert review (gold standard). There were 6 major keywords identified that yielded PPV of 82.5-100% PPV for case identification in different clinical fields. The triage notes and chief complaints have PPVs of 88.0 and 82.5%, respectively.
Horng, S. 2017.[111]	USA	Predict outcomes	To describe the utility of free text data in identifying patients with	This retrospective study describes a method for reducing duplicate and redundant information in ED electronic charting systems. Using narratively entered chief complaint data researchers identified

First author, Study year	Country	Narrative use	Study objective	Summary of findings
			suspected infections.	64,849 unique complaints from 1.7 million visits. Of these complaints 450 terms were used more than 100 times. Of the terms used 177 (39.3%) matched current definitions. The authors categorically grouped complaints into alternate clinically meaningful groups: cardiovascular (chest pain, arrhythmias), respiratory (shortness of breath, cough), gastrointestinal (abdominal pain, vomiting), and injury (fractures, trauma).
Kuramoto-Crawford, SJ. 2017.[112]	USA	Case identification	To evaluate and describe whether unstructured triage chief complaint data can identify suicide-related ED visits.	This retrospective study examined all ED visits with an unstructured narrative chief complaint or diagnosis that mentioned a suicide related term. These cases were examined and numbers of cases identified using each field were compared to describe the patient characteristics and numbers of cases identified using each method. The numbers of cases identified by method did not vary according to patient sex or age. Chief complaint narratives identified 62% of cases, discharge diagnoses identified 38% of cases.
Kondis, J. 2017.[59]	USA	Case identification	To determine the incidence of abuse related fractures for infants with prior ED visits for fussiness.	The aim of this retrospective cohort study was to determine the incidence of diagnosing fractures suggestive of abuse in infants who presented to the ED with "fussiness". Triage key word searches of electronic health records and imaging reports from radiology databases infants younger than 6 months were used to identify cases. Descriptive statistics were used to describe the frequency and distribution of ICD codes, chief complaints, presence of fracture, intervals between visits, outcomes of radiology variables in 18 patients. Of the 16 patients with fractures suggestive of abuse: the mean age was 2.5 months (SD=1.2). The mean interval between the initial and second presentation was 27 days. The most common diagnoses were multiple fractures, extremity fractures, and rib fractures.
Morano, LH. 2017.[60]	USA	Case identification	To improvements the current heat-related illness syndrome classification definition.	This retrospective validation study's aim was to update the heat-related illness syndrome definition used in a syndromic surveillance database. Using heat-related illness diagnostic codes and triage narrative keywords researchers identified and characterized: the presentation characteristics of heat-related illnesses, compared the numbers of cases identified using existing and updated definitions, compared the PPV for each search strategy, and described the correlation coefficient for different data sources (triage note, chief complaint, and ICD codes). The existing definition identified 8,928 ED visits; the updated definition identified an additional 598 ED visits.

First author, Study year	Country	Narrative use	Study objective	Summary of findings
				Of the triage narrative (keyword) identified visits (n=4,006), 3216 (80.3%) were identified using the following terms: "heat ex" (n=1674, 41.8%), "overheat" (n=646, 16.1%), "too hot" (n=594, 14.8%), and "heatstroke" (n=302, 7.5%).
Zhang, X. 2017.[61]	USA	Predict outcomes	To compare the predictive power of machine learning models that use natural language processing (using triage narratives), to those that do not, in predicting admission to hospital.	This cross-sectional study examined a probability sample of nationally representative ED data ED from a two-year period. Data available at triage (including the triage nurse narrative) was used to develop machine learning models (with and without unstructured narrative data) to predict hospital admission. Variables derived using natural language processing from the unstructured triage narrative were used to explain 75% of the variance in hospital admission. machine learning models using structured data only had AUCs scores of 0.823-0.824 (95% CI 0.817-0.829 to 0.818-0.830) For logistic regression and neural-network models, respectively. machine learning models using unstructured data alone had AUC of 0.742 (95% CI 0.742-0.764). Models using both types of data had AUCs of 0.844 to 0.846 (95% CI 0.836-0.852 to 0.839-0.853) for logistic regression and neural network models, respectively.
Chu, KH. 2018.[62]	AUS	Case identification	To compare the incidence and outcome of intracranial hemorrhage in the general population and in ED patients presenting with headache.	This retrospective cohort study's aim was to describe the proportion of patients who presented to the ED with headache who received a diagnosis of subarachnoid hemorrhage and to compare it to the incidence in the general population. Using a regional ED database and vital statistics bureau birth, death and marriage registry. Patients with an ICD codes for subarachnoid hemorrhage, triage narrative including headache keywords, and a death registry subarachnoid hemorrhage cause of death were compared. Descriptive statistics were used to describe ICD codes, mortality, demographics, and incidence. The incidence of subarachnoid hemorrhage was 9.9 (95% CI 9.5–10.4) per 100,000 person-years. The in-hospital mortality rate was 23.8% (95% CI 22.0–25.8%). 1.9%, (95% CI 1.8–2.0) of ED patients with headache were diagnosed with subarachnoid hemorrhage.
Gligorijevic, D. 2018.[63]	USA	Predict outcomes	To develop a machine learning model that can predict resource requirements for ED patients	This retrospective cohort study used the triage narratives from ED visits to develop a machine learning model to predict resource utilization for ED patients. Visit characteristics, patient characteristics, vital signs, and insurance status were used as inputs for the machine learning model. Using these combined sources, the

First author, Study year	Country	Narrative use	Study objective	Summary of findings
				machine learning model achieved an AUC of 88% for identifying resource intensive patients, an accuracy of 44% in predicting the category of resources, and was 16% more accurate than nurses in its predications.
Goldman-Mellor, S. 2018.[64]	USA	Predict outcomes	To describe presenting complaints syndrome classifier and to examine its ability to predict mental health diagnoses using discharge ICD codes.	This retrospective cohort study describes a presenting complaints syndrome classifier and its ability to predict mental health diagnoses using diagnosis codes. A machine learning model was used to automatically classifies free text into chief complaints to identify cases as belonging to a syndrome group. Syndrome groups were compared to diagnosis codes. Agreement between the machine learning model and discharge codes and the ability to identify cases was high (Kappa=0.92; Sensitivity=100%, Specificity=98.6%).
Hargrove, J. 2018.[65]	USA	Case identification	To describe the effects of different case definitions on identifying motor vehicle crash injury cases.	This retrospective cohort study used triage narratives to automatically classify patients presenting complaints into mental health syndromes. The authors compared discharge ICD codes to the classifier models identified cases in order to: identify cases, describe the typically presenting complaints in these patients, and to examine the classifier model's ability to predict mental health diagnoses. Descriptive statistics were used to characterize presentations. The classifier model has excellent agreement with the ICD identification approach (Kappa=0.92), 100% sensitivity, and 98.6% specificity in identifying cases.
Hendin, A. 2018.[113]	CAN	Quality improvement	To describe the outcomes of patients older than 65 years who received lower acuity scores.	This retrospective validation study's objective was to describe and compare different motor vehicle collision case definitions. Using existing motor vehicle collision codes, triage texts suggestive of motor vehicle collision, or a combination of both researchers described demographic variables (sex and age), arrival means (ambulance, walk-in, other), insurance status (private insurance, Medicare, Medicaid, self-pay, or other), disposition (home, admitted, transferred, left without being seen, died). Patients identified using motor vehicle collision codes were more likely to be male, arrive by ambulance, and be admitted when compared to the other search methods.
Nagabhushan, M. 2018.[66]	USA	Case identification	To determine how frequently patients with confirmed acute thoracic-aortic	This retrospective cross-sectional study examined emergency records to estimate the prevalence of chest pain described as "tearing" or "ripping" by patients who received a diagnosis of aortic aneurism rupture. Complete review of the

First author, Study year	Country	Narrative use	Study objective	Summary of findings
			dissection present to the ED with chest-pain described as “ripping’ or “tearing”.	emergency records (including triage narratives) failed to identify any cases in which the patients used “ripping” or tearing” descriptors. The authors suggest that the complaints which have been historically held as prototypical may not be commonly used by patients.
Petruniak, L. 2018.[114]	CAN	Quality improvement	To describe the variables associated with triage acuity classification in patients with sepsis.	This retrospective cohort study described patient and contextual variables associated with high- versus low-acuity triage scoring in patients diagnosed with sepsis. Patients admitted with sepsis, severe sepsis, or septic shock (ICD codes) had their triage records examined. Triage narratives were examined and variables such as: unwell appearance, communication barriers, number of comorbidities, number of prescription medication, and place of residence were assigned and used as variables in addition to: CTAS score and ED census. Logistic regression found communication barriers and cognitive impairment (OR 5.7; 95% CI 2.15-15.01), acute confusion (OR 3.4; 95% CI 1.3-8.2), unwell appearance (OR 3.4; 95% CI 1.7-7.0), and hypotension (OR=0.98; 95% CI 0.96-1.0) predictive of higher acuity classification. ED census, heart rate, respiratory rate, and temperature were not predictive of triage acuity assignment.
Rice, BT. 2018.[115]	UGA	Case identification	To develop and validate a chief complaint list for ED visits in Uganda.	This retrospective cohort study developed and validated a chief complaint list for emergency care in Uganda. Using triage narratives from consecutive visits to a rural ED authors developed groups and compared rater assignment of the derived categories. 555 chief complaints were used to initially categorize 95.8% of visits. This complaint list was further refined using a consensus process to yield a longlist of 451 total complaints and shortlist of 83 complaints. Interrater reliability using the shortlist was of complaints was 71.5% agreement, (kappa of 0.70), over 80% of all visits were categorized into 24 primary complaints.
Soufi, MD. 2018.[44]	IRN	Quality improvement	To design and evaluate a clinical decision support system that can be used at triage to improve the speed and accuracy of triage acuity assignment.	This multi-step study used observations, unstructured interviews, and survey data to design a clinical decision support tool and a crossover intervention to validate the tool. Keywords from triage narratives were used as predictor variables in a machine learning model that achieved an accuracy rate of 99.44% in test data, and helped improve the completeness of documentation from 76.72% to 98.5%.
Chen, M. 2019.[67]	AUS	Case identification	To compare the injury patterns	This retrospective cohort study compared injury and treatment patterns for trampoline related

First author, Study year	Country	Narrative use	Study objective	Summary of findings
			and treatment costs in patients injured on trampolines at commercial and residential locations.	injuries that occurred at home versus a commercial trampoline park. The authors searched the triage narratives from a children's hospital over a one-year period using trampoline specific keywords to identify cases. Demographics, incident location, injury pattern, and cost data were derived from visit case notes, imaging reports, and surgical reports comparisons were made between injuries sustained at home versus a commercial facility. of the 392 cases identified 68.9% occurred at home; 19.4% were from a commercial location. The median age of those injured at home was significantly younger than commercial locations (5.6 vs. 12.8 years; P < 0.001), There were significantly more females injured at commercial facilities than at home (61.8% vs. 48.2; P = 0.03). 27.3% of injuries resulted from a fall from the trampoline. Fractures were the most common injury (39.5%); 17.4% required admission, and 12.8% surgery.
Choi, SW. 2019.[48]	KOR	Predict outcomes	To develop and evaluate machine learning models to predict triage acuity.	This retrospective cross-sectional study trained and compared machine learning models to triage acuity levels using categorical data, free text triage narratives only, and both categorical data and triage narratives. Categorical data included demographics (sex and age), arrival characteristics (date/time, coded chief complaint, means of arrival, vital signs. The free-text triage narrative was an unstructured nursing note, one to three sentences in length that summarized the reason for presenting to the ED. The AUC of the model using categorical data was greater than models trained on free-text data only, the machine learning models with the highest AUC were trained on the entire dataset.
Eley, R. 2019.[68]	AUS	Case identification	To describe characteristics of patients presenting to the ED for bicycle collisions.	This retrospective cohort study used triage narratives to characterise bicycle collision characteristics. Researchers used keywords to identify cases and derive injury patterns, object collided with, and trends over time. The authors describe visit demographics and typical injury patterns.
Greenbaum, NR. 2019.[116]	USA	Quality improvement	To describe the effects of a machine learning driven autocomplete on the efficiency and quality of triage documentation.	This multi-phase mixed-methods study used: a retrospective cohort to derive a triage narration ontology, a prospective quality improvement method to implement of a machine learning driven autocomplete function into triage documentation, and a before-after assessment of documentation quality (defined as the frequency in which nurses needed to use unstructured data fields). The authors described patients and clinical

First author, Study year	Country	Narrative use	Study objective	Summary of findings
				characteristic. The machine learning autocompletes improved documentation completeness and data quality (based on expert consensus), while simultaneously reducing the numbers of keystrokes needed to document a triage assessment.
Jones, R. 2019.[89]	AUS	Case identification	To analyze methamphetamine related ED presentations.	This retrospective observational study analyzed methamphetamine related ED presentations in patients aged 13-59 who presented with an injury. Triage narratives were interrogated using keywords to identify cases and describe visit characteristics. The relationship between: drug type, demographics (age, sex, ethnicity), triage acuity level, and year were examined. 84.4% of methamphetamine-related visits had a high triage acuity score. Compared to other visits methamphetamine-related visit more frequently required police or ambulance involvement.
Lee, SH. 2019.[37]	USA	Case identification	To compare machine learning algorithm's ability to assign a syndromic category using chief complaint data only.	This retrospective validation study compared the performance of different machine learning models for disease surveillance by predicting a coded chief complaint using triage narratives. This study used chief complaint and discharge diagnosis narratives and compared them to coded diagnosis codes. The authors found that different machine learning models were better at predicting certain syndrome using different data sources and suggest that deep learning machine learning models may improve automatic coding of unstructured narrative data.
Marx, GE. 2019.[69]	USA	Case identification	To compare the performance of syndromic classifiers and physician review in identifying marijuana-associated ED visits using medical records.	This retrospective, observational validation cohort study assessed the validity of using syndromic surveillance data to detect marijuana-associated ED visits. The authors compared syndromic queries to physician-reviewed medical records in identifying marijuana-related visits. Keyword searches of triage narrative were compared to diagnostic code identified cases. Keyword identified cases, and diagnostic code + keyword identified cases were compared to expert identified "true cases". Descriptive statistics were used to report on patient demographics (age, sex, residential address), route of marijuana use, toxicology testing results, and if there was documented polysubstance use. For each visit, marijuana-specific diagnostic codes and keywords noted in the chief complaint or triage narrative were recorded. 453 of 44,942 total ED visits were identified as marijuana related; 188 (45%) of identified cases were true cases. All true positive cases were identified as marijuana-related using

First author, Study year	Country	Narrative use	Study objective	Summary of findings
				either diagnostic codes or triage keywords. PPV of each method varied by hospital (36% to 64%). 109 (58%) of true positive cases were men and 178 (95%) of cases used marijuana intentionally, visits were more likely to report versus smoked marijuana.
Nanda, G. 2019.[117]	AUS	Case identification	To compare different machine learning methods for identifying medical records that will require human review and assignment of rare cause-of-injury codes.	This two-stage retrospective descriptive, prospective cohort study compared the ability of different machine learning models to identify cases needing human review while assigning injury codes. The authors used four machine learning models to examine the triage narrative and determined that each model was able to detect rare cases needing human review.
Sterling, NW. 2019.[118]	USA	Predict outcomes	To examine machine learning algorithm's ability to predict patient disposition using triage narrative alone.	This retrospective cohort study examined the ability of machine learning models to predict ED patient disposition. Using descriptive statistics and by comparing AUC the authors described the performance of three models (bag-of-words, paragraph vectors, and topic distributions) in predicting whether patients were discharged home, admitted to hospital, or left prior to completion of treatment. The mean character count for triage narratives was 143.17 (SD=77.8) and contained 64.3 words (SD=35.2). AUC of models were Bag-of-words= 0.737 (95% CI 0.734 - 0.740), Paragraph Vectors=0.785 (95% CI 0.782 - 0.788), Topic distribution=0.687 (95% CI 0.684 - 0.690).
Trivedi, TK. 2019.[70]	USA	Case identification	To identify injuries related to electric scooters and characterize their presentation patters and clinical outcomes.	This retrospective cohort study used keyword searches of triage narratives to identify patients presenting with injuries associated with electric scooter use. Flagged cases were reviewed, patient and visit characteristics were described. 58.2% of visits were male, the mean [SD] age was 33.7 [15.3] years. Injuries were described and the majority of patients ([94.0%) were discharged and helmet use rates were noted to be low.
Xingyu Zhang, M. 2019.[74]	USA	Predict outcomes	To describe the effects of socioeconomic, demographic and clinical factors on diagnostic imaging use and to develop a machine learning algorithm that can predict diagnostic	This study was a two-stage retrospective-descriptive and validation study that examined the association between ED visits and diagnostic imaging use in pediatric patients. The authors used machine learning to compare models that used: structured (vital signs), unstructured (free-text triage narratives), or a combination of structure and unstructured data. There were 27,665 visits included, 8394 (30.3%) visits received diagnostic imaging - 6922 (25.0%) visits received an X-ray and 1367 (4.9%) computed

First author, Study year	Country	Narrative use	Study objective	Summary of findings
			imaging in pediatric ED patients.	tomography scan. The model using structured variables had a c-statistic of C=0.71 (95% CI 0.70–0.71) for any imaging use, C=0.69 (95% CI 0.68–0.70) for X-ray, and C=0.77 (95% CI 0.76–0.78) for computed tomography. machine learning models including only unstructured information had higher predictive power C=0.81 (95% CI 0.81–0.82) for any imaging use, C=0.82 (95% CI 0.82-0.83) for X-ray, and C=0.85 (95% 95% CI 0.83–0.86) for computed tomography. Including both structured and unstructured data improved predictive power C=0.82 (95% 95% CI 0.82–0.83) for any imaging use, C=0.83 (95% 95% CI 0.83–0.84) for X-ray, and C=0.87 (95% 95% CI 0.86–0.88) for computed tomography.
Zhang, X. 2019.[119]	USA	Predict outcomes	To predict advanced diagnostic imaging use the data immediately available during ED triage.	This retrospective cohort study was a secondary analysis that examined the association between advanced diagnostic imaging utilization and the structured and unstructured information available during ED triage of adult patients. The authors compared models using structured, unstructured (triage narrative), and structured plus unstructured data. Structured data included demographics (age, sex, ethnicity), residence (private residence, nursing home, homeless, or other), insurance status, conveyance (ambulance or not), visit characteristics (arrival time, previous 72hr visit, initial versus follow-up visit), vital signs (pain scale, temperature, heart rate, respiratory rate, blood pressure, pulse oximetry), chief complaint codes, medical history, and triage acuity score. Unstructured data (reason for visit triage narrative) were extracted using natural language processing. There were 139,150 visits included, 21.9% resulted in advanced diagnostic imaging utilization: 16.8% who received computed tomography, 3.6% who received an ultrasound, 0.4% who received magnetic resonance imaging, and 1.2% of received multiple diagnostic imaging types. The AUC ranged from 0.69 to 0.83 depending on the diagnostic imaging exam assessed and the variables included in the predictive model: the inclusion of unstructured data improved the accuracy of all models.
Bacchi, S. 2020.[120]	AUS	Predict outcomes	To compare different machine learning models discriminatory power in predicting length of stay and	This retrospective, observational validation cohort study examined whether length of stay and discharge destination could be predicted using natural language processing and machine learning models. The authors examined the discriminatory power predicting whether patients would have admissions (≤2 days or>2 days) using the grouped

First author, Study year	Country	Narrative use	Study objective	Summary of findings
			discharge destination.	unstructured data derived from nursing triage and physician admission notes. The artificial neural network model had the highest predictive power of r admissions >2days (AUC=0.75). For the prediction of home as a discharge destination all machine learning models performed similarly.
Fernandes, M [1]. 2020.[121]	PRT	Predict outcomes	To describe a machine learning algorithms ability to predict the risk of mortality for patients at triage.	This retrospective cohort study evaluated a machine learning models' ability to predict which patients are at risk of requiring cardiopulmonary resuscitation or dying within 24 hours from triage. machine learning models using structured data only were compared to those that included unstructured data. Structured data included vital signs (temperature, heart and respiratory rate, blood pressure, pulse oximetry, Glasgow Coma Scale, blood glucose level, and pain scale), coded chief complaints, triage acuity level, Demographics (age, sex), arrival information (ambulance vs. walk-in, stretcher vs. wheelchair or ambulatory, time of triage) number of prior ED visits, and the number and type of exams performed at triage. Unstructured data was the triage narrative for chief complaint. Of the 235826 patients examined 1121 required resuscitation or died. The predictive power of models was: logistic regression AUC=0.93, random forests AUC=0.95, and extreme gradient boosting AUC=0.96. The predictor variables with highest importance were Glasgow coma score, patient age, pulse oximetry and arrival mode. The models using structured and unstructured data had higher recall most accurately identified patients at risk of death or requiring resuscitation.
Fernandes, M [2]. 2020.[122]	PRT, USA	Predict outcomes	To describe a machine learning algorithms ability to identify patients with a high risk of intensive care unit admission.	This multinational retrospective cohort study examined the ability of a machine learning model to predict patients at high risk for intensive care admission from EDs in Portugal and the United States. machine learning models used different sets of variables: triage acuity, triage acuity + Clinical variables, triage acuity + clinical variables + chief complaint, and clinical variables + chief complaint. For both hospitals, the logistic regression model had the best performance, with AUC of 0.91 (95% CI 0.90-0.92) for the United States hospital and 0.85 (95% CI 0.83-0.86) for the Portuguese hospital. Vital sign derangement (Heart rate, pulse oximetry, respiratory rate and blood pressure) were the biggest predictors of ICU admission. machine learning models using clinical variables and the chief complaint

First author, Study year	Country	Narrative use	Study objective	Summary of findings
				presented had improved recall for patients who are at risk for ICU admission.
Jones, PG. 2020.[123]	NZL	Quality improvement	To determine the accuracy of coded chief complaint coding using triage narratives.	This retrospective cohort study examined the accuracy of chief complaint coding using expert re-triage of visits. Triage narratives were used by experts to assess the accuracy of assigned chief complaints by comparing actual and assigned chief complaints. Results were compared for concordance, by sex, age, ethnicity, and by presenting complaints. 852 of the 1,000 visits sampled were eligible, coded chief complaints agreed with clinical notes in 514 (60.3%) of cases. Overall agreement occurred in 732 (85.9%) cases. Patient age, sex and ethnicity did not influence coding accuracy.
Joshua, W. 2020.[75]	USA	Predict outcomes	To compare the ability of machine learning models to ESI or vital sign triggers to predict critical illness using triage data.	This retrospective cross-sectional and prospective comparison study compared the predictive power of different machine learning models in identifying critically ill patients. The models used data collected at triage: vital signs, triage acuity scores (ESI < 3), and free-text chief complaints as inputs to identify patients who were admitted to the ICU or died within 24hours. 13.7% of patients were critically ill; the AUC for the simplest predictive models (vital sign and triage acuity) were 0.521 and 0.672 (95% CI 0.519-0.522 and 0.671-0.674), respectively. For machine learning models AUCs were 0.803 (95% CI 0.802–0.804) for logistic regression, 0.820 (95% CI 0.818–0.821) for gradient boosting, 0.811 (95% CI 0.807–0.815) for structured data neural network, and 0.851 (95% CI 0.849–0.852) for neural network models using both structure and textual data.
Klang, E. 2020.[45]	ISR	Predict outcomes	To develop a machine learning model that predicts a patients need for non-contrast head CT.	This retrospective cohort study developed a machine learning model to predict the need for non-contrast head computed tomography during triage. The models compared used coded and free text chief complaints derived from the triage narrative. Coded data was limited to triage specific data: demographics (age and sex), presentation data (admission date/time); conveyance (either walk-in, ambulance), number of previous visits, triage acuity score, and coded chief complaint. Free text complaints were derived from a two-word complaint recorded by the triage nurse. Of the included 595,561 visits computed tomography rate was 11.8%. Of each unique variable chief complaint had the AUC (0.87). The best model showed an AUC of 0.93

First author, Study year	Country	Narrative use	Study objective	Summary of findings
				(95% CI 0.931–0.936) for predicting non-contrast head CT usage at triage level.
Klug, M. 2020.[46]	ISR	Predict outcomes	To evaluate a machine learning model for predicting mortality and triage level.	This retrospective cohort study trained and evaluated the performance of a machine learning model for predicting mortality using triage data. The authors evaluated each of the following as individual predictor variables: demographics (age and sex), presentation data (admission date/time); conveyance (either walk-in, ambulance), number of previous visits, triage acuity score, home medications, comorbidities, and coded chief complaint. Early and short-term mortality data (within 2 and 30 days of triage, respectively) was derived from electronic health records and vital statistics registries. Of the 799,522 ED visits the early mortality rates were 0.6% and 2.5%, respectively. The machine learning Model utilizing the full data set had an AUC of 0.962 for early and 0.923 for short-term mortality. The machine learning model using the nine most predictive variables (age, arrival mode, chief complaint, five primary vital signs, and triage acuity) had an AUC of 0.962 for early mortality.
Mor, S. 2020.[47]	ISR	Quality improvement	To describe how frequently family history of coronary artery disease is collected at triage and to compare the clinical characteristic of patients with and without family history.	This retrospective cohort describes how frequently clinicians recorded family history of coronary artery disease and whether it is predictive of ST elevation myocardial infarction. The following clinical risk factor information was collected: demographics (age, sex, and ethnicity), family history of coronary artery disease, and medical history (diabetes, dyslipidemia, hypertension, previous cardiac events, and smoking). Outcome variables examined included triage acuity score assignment, time to: nursing care, ECG, physician assessment and diagnosis, hospital length of stay, readmission rates, and mortality. Compared to physicians' nurses collected family history less frequently (98.8% vs. 5.7%).
Stapelberg, NJC. 2020.[124]	AUS	Case identification	To validate a machine learning algorithm that can identify suicidal and self-harm presentations to the ED.	This retrospective cohort study used visit data from 2 EDs to validate a machine learning algorithm to identify suicidal ideation and self harm presentations. The psychiatrist evaluated dataset was used to determine factor weights and the resulting machine learning algorithm achieved a sensitivity of 0.95 and a specificity of 0.92 in identify cases.
Robinson, J. 2020.[125]	AUS	Case identification	To describe the machine learning language processing	This retrospective cohort study details the development of a regional self-harm surveillance system. Cases will be identified using diagnostic and billing codes and free text triage notes. A

First author, Study year	Country	Narrative use	Study objective	Summary of findings
			models and data elements that were used to develop a self-harm surveillance system.	machine learning model will identify ED visits for self harm using data collected during ED visits: demographics (age, sex, residential address), visit characteristics (date and time), clinical characteristics (categorical chief complaint, free-text triage narratives, diagnostic codes), and treatment information (time to assessment, disposition). These fields will be used to identify cases, mechanism of injury, and medication involvement data. Mechanisms of injury will be grouped as: Injuries (intentional injury, use of firearms or explosives, sharp objects, falls from a height, and other); drug related (medication or drug, alcohol, poison/chemical/noxious substances); and hanging, strangulation, suffocation, drowning, or submersion.
Roquette, BP. 2020.[126]	BRA	Predict outcomes	To compare different machine learning models' ability to predict admission using both structured and unstructured triage data.	This retrospective observational study evaluated a machine learning models predictive power for patients requiring hospital admission using both structured and unstructured data available at triage for patients under 18 years of age. machine learning model were trained with structured clinical variables and unstructured (textual) clinical narratives. Structured variables included: Demographics (age, sex, city of birth,), registration details (insurance statues, number and time from previous ED visits, home residence), historical variables (number of past image exam requests, number of past laboratory test requests, number of past ED visits, number of prior admissions, past image exam/visit ratio, past admission/visit ratio, past laboratory exam/visit ratio), clinical variables (triage acuity score, categorical symptom,); time variables (arrival time, time to triage); vital signs (heart rate, blood pressure, oxygen saturation, temperature, weight, blood glucose, pain score); and unstructured narrative data that included (drugs, text triage notes, previous triage notes, chief complaint, past visit chief complaint, previous visit medical image exam requests). There was a 5.76% admission rate in the studied period. The machine learning model achieved an AUC of 0.892.
Sterling, NW. 2020.[71]	USA	Predict outcomes	To compare the predictive power of machine learning models to human raters in predicting triage	This retrospective cohort and prospective model validation study trained and validated a machine learning model to predict ED resources requirements using triage acuity as a proxy. Using natural language processing Spielberg of triage narratives, the researchers compared the predictive power of machine learning models to

First author, Study year	Country	Narrative use	Study objective	Summary of findings
			acuity using triage narratives.	human experts using structured (demographics, age, sex, vital signs, previous visit information, arrival mode, coded chief complaint, and coded medical history variables and unstructured triage narratives. The machine learning model accuracy was comparable to expert raters (65.9% vs 69.0%). machine learning models more accurately predicted low resource visits.
Sveticic, J. 2020.[127]	AUS	Case identification	To compare the sensitivity and specificity of discharge codes in identifying suicide and self harm cases in a dataset that used machine learning of triage narratives to identify cases.	This retrospective cohort study evaluated the reliability of different methods for identifying suicide and self injury ED presentations. Charts identified by a machine learning algorithm as related to suicide were reviewed by an expert clinician and classified as dealing with: a suicide attempt, suicidal ideation, or non-suicidal self-injury. The diagnosis codes, presenting complaints, and patient/visit characteristics were compared. There was significant diagnostic and presenting complaints code heterogeneity. Diagnostic codes had low sensitivity in identifying suicide attempts (18.7%), non-suicide self-injury (38.5%) and suicidal ideation (42.3%) and were biased toward identifying cases associated with female gender, indigenous status, and cases involving knife related presentations.
Vernon, N. 2020.[34]	USA	Case identification	To describe the patterns and characteristics of patients presenting with injuries related to e-scooters.	This retrospective review used triage narratives to identify keywords associated with electric scooter injury ED visits. The authors used keyword searches of triage narratives to identify cases and describe the patterns and characteristics of patients presenting with injuries related to e-scooters. 293 patients were included. The mean patient age was 34.1 (SD=14.0), 56% of patients were male. 100% of patient received at least one diagnostic imaging test (number of tests ranged from 1-9 (mean=2.4). 276 (94.2%) patients were discharged from the ED, 5 (1.7%) were transferred, 1%, 2 (0.7%) were admitted.
Bouchouar, E. 2021.[128]	CAN	Case identification	To describe the process of developing and implementing a syndromic surveillance system.	This two stage study reports on the retrospective design and prospective validation of a syndromic surveillance system. The five stages include: the initial review of ED data; the development of syndrome definitions; The natural language processing model development; validation of machine learning model; and refinement of the machine learning model. ED data examined included: coded chief complaints, triage narratives, discharge diagnosis codes. Validation and refinement measured the positive predictive value (the total number of true-positive over the total number of records). Each data fields

First author, Study year	Country	Narrative use	Study objective	Summary of findings
				influence on syndromic assignment was measured individually and in combination. The narrative terms were common in each syndrome were measured as frequencies. The positive predictive value for the automated detection of each syndrome ranged from 48.8–89.5% to 62.5–94.1%.
Cheung, KY. 2021.[76]	CHN	Quality improvement	To examine the validity of triage by comparing triage nurse assigned acuity scores with expert review and rates of resuscitation in older adults.	This retrospective descriptive, cross sectional study re-examined triage acuity scores by: comparing assigned categories and scores against an expert panel and by examining for whether patients required life-saving interventions in adults greater than 65 years of age. Using the clinical narrative and structured variables (assigned complaint category and vital signs) experts re-triaged patients. Sensitivity, specificity, and predictive values were used to summarise whether patients who received resuscitation were correctly assigned in to the highest triage scores. Percentage agreement and weighted kappa were used assess agreement on triage acuity levels and rates of: correct, over, and under-triage. Overall expert panel agreement was 96.7%. The agreement between clinical triage category and the expert panel assignment was 68.5%; with 16.3% and 15.3% over-triage and under-triage rates, respectively. Weighting kappa between groups was 0.72 (95% CI 0.53–0.91). The sensitivity, specificity and positive likelihood ratio for the need for life-saving interventions were 75.0% (95% CI 47.6%–92.7%), 97.1% (95% CI 94.4%–98.8%), and 26.2 (95% CI 12.5%–54.8%), respectively.
Delany, C. 2021.[129]	AUS	Case identification	To describe the effects of the 2018 Commonwealth Games on the volumes and distributions of drug and alcohol related ED presentations.	This retrospective observational study examined the impact of the 2018 Commonwealth Games on drug and alcohol related ED presentations. Using physician assigned diagnostic codes, and drug and alcohol related keyword searches of nurse recorded free-text triage narratives researchers identified cases. Presentation rates were compared to the same time period in the preceding year. Patient demographics (sex, country of residence,), presentation characteristics (arrival date/time, mode of arrival, triage acuity), and outcomes (discharge date/time, departure destination) were compared. There were 890 (5% of total) drug and alcohol-related ED presentations in the study period with no difference between pre (n = 312), during (n = 301) and post (n = 277) periods (P =0.2). Presentations were more common in younger (median age=35 years, IQR=24–48),

First author, Study year	Country	Narrative use	Study objective	Summary of findings
				male (n = 493,55%), Australian (n = 820, 92%) patients. most patients arrived by ambulance (n = 650, 73%), during the hours of 3PM and 11PM (n = 365, 41%). keyword searching identified 500 additional cases, the most commonly used terms were: 'ETOH' (n = 274, 55%), 'overdose'/ 'OD' (n = 159, 32%), 'intox' (n =63,13%), 'drink' (n = 58, 12%) and 'drug' (n= 45, 9%).
Eyal Klang, MA. 2021.[130]	USA	Predict outcomes	To develop a machine learning model to predict mortality during triage.	This retrospective study developed and compared different machine learning models to predict mortality using structured free-text narrative notes. The study outcome was in-hospital death within 48 hours; structured and free-text data collected within 30 minutes from triage were included in the model. Variables include demographics (age, sex, and ethnicity); arrival mode (walk-in, by ambulance); clinical parameters (vital signs, triage acuity scores, coded chief complaints, laboratory orders); and nursing and physician free text information. The 48-hour mortality rate was 0.2%. the machine learning model AUC values ranged from AUC=0.97 (95% CI 0.96–0.99) to AUC=0.98 (95% CI 0.98–0.99), when combined the models AUC value was 0.99 (95% CI 0.98–0.99).
Ivanov, O. 2021.[50]	USA	Predict outcomes	To determine if different machine learning models could accurately assign triage scores.	This retrospective cohort study sought to determine if different machine learning models could accurately assign triage acuity scores. Using visit characteristics, free text narratives from triage, and patient history data machine learning model assigned triage acuity scores and original triage records were compared to expert clinician re-triage. Triage scores, mis-triage rates and AUC for triage accuracy were compared. machine learning model assigned acuity scores were more frequently accurate than the original nurse assigned score (75.7% vs. 59.8%, respectively p<0.001) and had better discernment on patients on the boundary between level 2 and 3 acuity assignments (80% vs. 41.4%, P <.001).
Klang, E. 2021.[131]	USA	Predict outcomes	To describe different machine learning models' ability to predict neurovascular intensive care unit admission using data available in the first 30min of a visit.	This retrospective cohort study used structured and unstructured data available in the first 30min of an ED visit to predict Neuro-intensive care unit admission. The authors compared structured, unstructured and combined powers of machine learning models. Structured data included: demographics (age, sex, ethnicity, home address); visit characteristics (date/time, means of arrival, number of previous visits/admissions); clinical variables (coded chief complaint, triage vital signs, triage acuity score, and medical history

First author, Study year	Country	Narrative use	Study objective	Summary of findings
				codes); and free text (triage, nursing, and physician texts generated within the first 30minutes). 1900 (0.5%) of the 412,858 visits were admitted into Neuro-intensive care. The median daily number of ED presentations was 231 (IQR 200–256), the median time from triage to admission was 169 min (IQR 80–324). The machine learning models AUC values were 0.90 (95% CI 0.87–0.91), 0.92 (95% CI 0.91–0.94), and 0.93 (95% CI 0.92–0.95) for models trained with unstructured text data only, structured data-only, and the combined data model, respectively.
Lam, T. 2021.[90]	AUS	Case identification	To compare the overdose patterns and distributions for different pharmaceutical opioids.	This retrospective observational study examined ED records to measure and compare pharmaceutical opioids contribution to non-fatal overdoses. Researchers compared the volume of pharmaceutical opioids dispensed to ED records. ED cases were identified using diagnostic coding and keyword searches of triage narratives using of variations of generic and commercial opioid names. Records were reviewed and the following variables were extracted: patient demographics (age, sex, country of birth, home address); ED characteristics (location, type); clinical variables (triage acuity score, admission status, diagnostic codes); and overdose intent (accidental vs. intentional) and were compared by opioid types which were adjusted by annual supply rates - per 100 000 oral morphine equivalents (OME). The highest supply-adjusted overdose rates were seen with codeine (OME = 0.078, CI = 0.073–0.08) and oxycodone (OME =0.029, CI = 0.027–0.030); the lowest with tapentadol (OME = 0.004, CI = 0.003–0.006) and fentanyl (OME = 0.003, CI = 0.002–0.004). 62% of poisonings involved females. Codeine, oxycodone and tramadol overdose were associated with younger patients (59.5%, 41.7% and 49.8%, respectively) and intentional self-harm (65.2%, 50.6%, and 52.8%, respectively).
Metzger, P. 2021.[132]	USA	Quality improvement	To describe the impact of race and language on the triage acuity scores of patients controlled for illness severity.	This retrospective cohort study examined the impact race and language had on triage acuity scores in illness severity adjusted presentations in immunocompromised patients aged 0 to 17 years. Cases were identified by performing keyword searches of triage narratives for the following words: 'central line', 'diabetes' or 'diabetic ketoacidosis', 'hematology', 'neonatal', 'oncology', 'short gut', 'shunt', 'sickle cell', 'transplant' or 'bone marrow transplant'. The following variables were extracted: demographic information (age, sex,

First author, Study year	Country	Narrative use	Study objective	Summary of findings
				insurance status, race/ethnicity, language), clinical variables (vital signs, triage acuity level, coded chief complaint), and treatment variables (time to provider, LOS, disposition and National ED Over-Crowding Study score) and compared by ethnic status (white race vs non-white). There were 10,815 visits from 8,928 patients. White patients were 34.6% of patients. Non-White patients had reduced likelihoods of receiving an emergency (level 2) or urgent (level 3) acuity score (OR= 0.4, 95% CI 0.33–0.49; OR, 0.5, 95% CI 0.45–0.56, respectively) and greater likelihood of receiving a minor (level 5) acuity score. The disparity was not noted when comparing language independently of ethnicity.
Personnic, J. 2021.[133]	FRA	Case identification	To describe the frequency of neurological disorders in a pediatric ED.	This retrospective observational study described the frequency with which patients with neurological disorders presented to a pediatric ED under 17 years of age. This study used keyword searches of the nursing free-text narrative of the chief complaint recorded at triage. Age, sex, chief complaint, history of previous visit, neurology referral, types of investigations (bloodwork, lumbar puncture, and neuroimaging), and medical history data were collected. 1,471 charts with neurological chief complaints were identified (1.8% of total visits); of these patients, 806 (55%) received a final physician assigned neurologic diagnosis, 2% of whom were admitted into pediatric intensive care. Seizures were the most common diagnosis for admitted patients. 40% of patients had at least on ED visit in the preceding 6 months.
Rahilly-Tierney C. 2021.[72]	USA	Case identification	To create data linkages for patients with both ED and ambulance visits for opioid related overdoses.	This retrospective cohort studies compared the unstructured narratives from ambulance services to ED triage in order to make data linkages between the systems for patients who presented with opioid-related overdoses. There was an 82% and 63% match rate for ambulance and ED records, respectively, with a 3% false positive match rate between the systems.
Rodríguez Vico, A. 2021.[134]	ESP	Quality improvement	To evaluate the quality of triage and compare the number of signs and symptoms of acute stroke gathered at triage and by using commonly used stroke scales.	This retrospective cohort study compared the triage note complaints of confirmed acute strokes to other recorded stroke scales within the same patient chart. The most common presenting complaints were aphasia, ataxia, and limb weakness. The authors determined that simplified stroke assessment scales would detect fewer symptoms than more comprehensive tools.

First author, Study year	Country	Narrative use	Study objective	Summary of findings
Rozova, V. 2021.[49]	AUS	Case identification	To develop an automated system that can identify self-harm related ED presentations using triage narratives.	This retrospective cohort study evaluated triage notes for suicidal ideation or self-harm using the following methods using different natural language processing models and keyword searches. These were then classified by expert adjudicators who assigned whether the cases were self-harm, suicidal ideation, or false positives. natural language processing outperformed keyword searches.
Tahayori, B. 2021.[135]	AUS	Predict outcomes	To compare machine learning models to expert reviewers' ability to predict ED disposition using triage narratives.	This retrospective cohort study used natural language processing of triage notes to predict ED disposition. The predictive power of the machine learning model was compared to expert consultant predictions. The machine learning model had an AUC of 0.88. The machine learning model had a predictive power that was close to expert consultants in general and was more sensitivity than expert review.

Figure 10. Data items collected

Study ID	Narrative	Age	Sex[i]	Chief complaint code	Discharge status	Triage acuity score	Triage system	Arrival means	Arrival day	Arrival time	Vital signs	Diagnosis	Admission location	Alcohol or drug use	Diagnostic imaging use	Mortality	Ethnicity[ii]	Length of stay	Pain	Syndrome classification	Treatment details	Labs	Physician narrative	Insurance	Trauma
Kabir, A. 1998.[93]	**X**				**X**			**X**	**X**			**X**													
Beveridge, R. 1999.[45]	**X**					**X**	**X**																**X**		
Goodacre, SW. 1999.[94]	**X**				**X**		**X**																		
Aronsky, D. 2001.[88]	**X**				**X**																				
Burt, CW. 2001.[153]	**X**	**X**	**X**	**X**		**X**								**X**			**X**	**X**					**X**		**X**
Howe, A. 2002.[159]	**X**	**X**	**X**							**X**	**X**			**X**											
Begier, EM. 2003.[145]	**X**			**X**	**X**																**X**				
Travers, DA. 2003.[105]	**X**				**X**							**X**									**X**				
Chapman, WW. 2004.[22]	**X**			**X**																	**X**			**X**	
Day, FC. 2004.[89]	**X**			**X**	**X**																				
Mikosz CA. 2004.[129]	**X**				**X**																				
Baumann, MR. 2005.[139]	**X**	**X**	**X**		**X**	**X**	**X**	**X**				**X**	**X**		**X**		**X**	**X**	**X**			**X**	**X**		
Chapman, WW [1]. 2005.[130]	**X**			**X**																	**X**				

Study ID	Narrative	Age	Sex[i]	Chief complaint code	Discharge status	Triage acuity score	Triage system	Arrival means	Arrival day	Arrival time	Vital signs	Diagnosis	Admission location	Alcohol or drug use	Diagnostic imaging use	Mortality	Ethnicity[ii]	Length of stay	Pain	Syndrome classification	Treatment details	Labs	Physician narrative	Insurance	Trauma
Chapman, WW [2]. 2005.[47]	**X**			**X**	**X**																				
Muscatello, DJ. 2005.[146]	**X**	**X**	**X**	**X**		**X**	**X**		**X**						**X**										
Thompson, DA. 2006.[85]	**X**	**X**	**X**	**X**	**X**	**X**			**X**									**X**			**X**				
Gillam, C. 2007.[163]	**X**													**X**											
Indig, D. 2008.[24]	**X**	**X**	**X**	**X**	**X**	**X**		**X**	**X**	**X**			**X**		**X**	**X**									
Irvine, AK. 2008.[161]	**X**																								
Indig, D. 2009.[28]	**X**	**X**	**X**	**X**		**X**			**X**	**X**	**X**			**X**		**X**								**X**	
Mitchell, R. 2009.[27]	**X**	**X**	**X**	**X**		**X**			**X**				**X**	**X**											
Indig, D. 2010.[51]	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**		**X**		**X**	**X**									
McKenzie, K. 2010.[109]	**X**			**X**	**X**																				
Wagholikar, AS. 2011.[95]	**X**				**X**																				
Bregman, B. 2012.[108]	**X**	**X**	**X**							**X**	**X**														
Grossmann, FF. 2012.[96]	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**			**X**			**X**			**X**	**X**			**X**			
Malmström, T. 2012.[97]	**X**				**X**																				
Rhea, S. 2012.[172]	**X**	**X**		**X**	**X**																				

Study ID	Narrative	Age	Sex[i]	Chief complaint code	Discharge status	Triage acuity score	Triage system	Arrival means	Arrival day	Arrival time	Vital signs	Diagnosis	Admission location	Alcohol or drug use	Diagnostic imaging use	Mortality	Ethnicity[ii]	Length of stay	Pain	Syndrome classification	Treatment details	Labs	Physician narrative	Insurance	Trauma
Genes, N. 2013.[162]	X									X	X	X													
Mosley, I. 2013.[141]	X	X	X	X	X	X	X	X	X				X					X				X			
Vallmuur, K. 2013.[52]	X	X	X	X						X	X		X	X		X									
Haas, SW. 2014.[98]	X				X	X						X									X				
Liljeqvist, HTG. 2014.[154]	X			X	X																X				
Rhea, SK. 2014.[91]	X	X	X	X	X																				
Handly, N. 2015.[131]	X	X	X		X	X	X	X		X	X														
Hides, L. 2015.[149]	X	X	X					X						X		X									
Mitchell, RJ. 2015.[50]	X	X	X			X	X	X	X					X				X							
Gray, SE. 2016.[48]	X																								
Luther, M. 2016.[110]	X	X	X	X						X	X					X									
Rahme, E. 2016.[111]	X	X	X	X		X	X						X		X			X							
Whitlam, G. 2016.[112]	X	X	X	X		X			X	X	X			X							X				
Berendsen Russell, S. 2017.[143]	X				X		X																		
DeYoung, K. 2017.[86]	X			X	X					X						X								X	
Horng, S. 2017.[26]	X	X	X	X		X	X	X							X			X		X					

Study ID	Narrative	Age	Sex[i]	Chief complaint code	Discharge status	Triage acuity score	Triage system	Arrival means	Arrival day	Arrival time	Vital signs	Diagnosis	Admission location	Alcohol or drug use	Diagnostic imaging use	Mortality	Ethnicity[ii]	Length of stay	Pain	Syndrome classification	Treatment details	Labs	Physician narrative	Insurance	Trauma
Kuramoto-Crawford, SJ. 2017.[155]	**X**	**X**	**X**	**X**	**X**					**X**	**X**														
Kondis, J. 2017.[113]	**X**	**X**		**X**		**X**								**X**			**X**					**X**		**X**	
Morano, LH. 2017.[114]	**X**			**X**	**X**																				
Zhang, X. 2017.[115]	**X**	**X**	**X**	**X**		**X**		**X**	**X**	**X**	**X**	**X**	**X**	**X**						**X**					**X**
Chu, KH. 2018.[116]	**X**	**X**	**X**	**X**																					
Gligorijevic, D. 2018.[117]	**X**						**X**	**X**																	
Goldman-Mellor, S. 2018.[118]	**X**	**X**	**X**	**X**	**X**	**X**									**X**						**X**				
Hargrove, J. 2018.[119]	**X**	**X**	**X**	**X**		**X**			**X**					**X**											**X**
Hendin, A. 2018.[142]	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**			**X**			**X**		**X**		**X**	**X**		**X**	**X**	**X**	
Nagabhushan, M. 2018.[120]	**X**			**X**		**X**									**X**		**X**	**X**						**X**	
Petruniak, L. 2018.[173]	**X**	**X**	**X**	**X**			**X**	**X**	**X**	**X**	**X**	**X**													
Rice, BT. 2018.[148]	**X**	**X**	**X**		**X**																				
Soufi, MD. 2018.[99]	**X**						**X**	**X**				**X**													
Chen, M. 2019.[121]	**X**	**X**	**X**	**X**		**X**			**X**					**X**					**X**			**X**			
Choi, SW. 2019.[104]	**X**	**X**	**X**		**X**		**X**		**X**	**X**	**X**	**X**								**X**					

Study ID	Narrative	Age	Sex[i]	Chief complaint code	Discharge status	Triage acuity score	Triage system	Arrival means	Arrival day	Arrival time	Vital signs	Diagnosis	Admission location	Alcohol or drug use	Diagnostic imaging use	Mortality	Ethnicity[ii]	Length of stay	Pain	Syndrome classification	Treatment details	Labs	Physician narrative	Insurance	Trauma
Eley, R. 2019.[122]	X	X		X	X		X	X																	
Greenbaum, NR. 2019.[100]	X	X	X		X		X	X				X													
Jones, R. 2019.[150]	X	X	X	X		X	X	X	X	X	X		X	X							X				
Lee, SH. 2019.[90]	X	X	X	X	X	X				X	X														
Marx, GE. 2019.[123]	X	X	X	X	X											X							X		
Nanda, G. 2019.[132]	X				X									X											
Sterling, NW. 2019.[25]	X					X	X																		
Trivedi, TK. 2019.[124]	X	X	X	X		X	X	X	X	X	X			X		X	X		X						
Xingyu Zhang, M. 2019.[126]	X	X	X		X				X	X	X	X	X	X		X	X			X					X
Zhang, X. 2019.[169]	X	X	X		X		X	X	X	X	X	X	X			X	X			X					X
Bacchi, S. 2020.[164]	X					X									X		X		X			X	X	X	
Fernandes, M [1]. 2020.[167]	X	X	X	X	X	X	X	X	X	X	X	X					X			X			X		
Fernandes, M [2]. 2020.[170]	X	X	X		X	X	X	X	X			X			X		X			X					
Jones, PG. 2020.[140]	X	X	X		X					X	X		X												

Study ID	Narrative	Age	Sex[i]	Chief complaint code	Discharge status	Triage acuity score	Triage system	Arrival means	Arrival day	Arrival time	Vital signs	Diagnosis	Admission location	Alcohol or drug use	Diagnostic imaging use	Mortality	Ethnicity[ii]	Length of stay	Pain	Syndrome classification	Treatment details	Labs	Physician narrative	Insurance	Trauma
Joseph, JW . 2020.[127]	X			X	X	X	X	X				X			X			X							
Klang, E. 2020.[101]	X	X	X	X		X	X	X			X	X				X									
Klug, M. 2020.[102]	X	X	X		X		X	X	X	X		X						X							
Mor, S. 2020.[103]	X	X	X	X			X	X					X		X			X	X			X			
Stapelberg, NJC. 2020.[156]	X	X	X	X	X	X	X	X		X	X		X												
Robinson, J. 2020.[157]	X	X	X	X	X	X				X	X		X	X					X						
Roquette, BP. 2020.[165]	X	X	X		X		X	X				X							X	X					X
Sterling, NW. 2020.[106]	X	X	X				X					X	X				X			X		X	X		
Sveticic, J. 2020.[158]	X	X	X	X	X		X						X	X								X			
Vernon, N. 2020.[87]	X	X	X			X	X		X	X							X		X						
Bouchouar, E. 2021.[147]	X			X	X		X																		
Cheung, KY. 2021.[128]	X	X	X			X		X	X																
Delany, C. 2021.[151]	X	X	X	X		X	X	X	X	X	X		X			X			X						
Eyal Klang, MA. 2021.[168]	X	X	X	X	X	X	X	X	X			X	X					X					X	X	

Study ID	Narrative	Age	Sex[i]	Chief complaint code	Discharge status	Triage acuity score	Triage system	Arrival means	Arrival day	Arrival time	Vital signs	Diagnosis	Admission location	Alcohol or drug use	Diagnostic imaging use	Mortality	Ethnicity[ii]	Length of stay	Pain	Syndrome classification	Treatment details	Labs	Physician narrative	Insurance	Trauma
Ivanov, O. 2021.[53]	**X**	**X**	**X**		**X**		**X**	**X**	**X**	**X**	**X**	**X**				**X**				**X**					
Klang, E. 2021.[171]	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**			**X**	**X**		**X**			**X**						**X**	
Lam, T. 2021.[152]	**X**	**X**	**X**	**X**		**X**	**X**	**X**					**X**			**X**									
Metzger, P. 2021.[174]	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**				**X**						**X**						**X**
Personnic, J. 2021.[160]	**X**	**X**	**X**	**X**	**X**										**X**		**X**						**X**		
Rahilly-Tierney C. 2021.[125]	**X**	**X**	**X**	**X**	**X**				**X**	**X**	**X**	**X**				**X**									
Rodríguez Vico, A. 2021.[144]	**X**	**X**	**X**	**X**																					
Rozova, V. 2021.[107]	**X**			**X**																					
Tahayori, B. 2021.[166]	**X**					**X**									**X**										
Total	96	63	60	57	53	43	41	35	32	29	26	25	22	19	17	16	14	14	12	11	10	10	10	9	7

i – Sex includes: sex and gender

ii – Ethnicity includes: Ethnicity, race, aboriginal status, and country of residence

Figure 11. Study characteristics by country

Country	Studies, n	Total EDs included, n[i] (studies[ii])	Total initial sample, n[i] (studies[ii])	Total included patients, n[i] (studies[ii])	Total included visits, n[i] (studies[ii])	Total nurses, n[i] (studies[ii])	Studies using ML methods, n
USA	43	2,008 (39)	36,528,693 (35)	916,955 (12)	4,986,560 (34)	3,781 (3)	22
AUS	31	404 (27)	23,110,878 (24)	1,996 (5)	4,784,753 (24)	2 (1)	10
CAN	5	7 (5)	6,450 (4)	573 (3)	19,727 (4)	20 (1)	1
ISR	3	3 (3)	1,586,760 (3)	648,294 (3)	1,361,455 (2)	- (0)	2
GBR	2	2 (2)	11,911 (2)	355 (2)	50 (1)	4 (1)	0
BRA	1	1 (1)	499,853 (1)	0 (0)	499,853 (1)	- (0)	1
CHE	1	1 (1)	- (0)	519 (1)	519 (1)	15 (1)	0
CHN	1	1 (1)	44,237 (1)	295 (1)	295 (1)	- (0)	0
ESP	1	1 (1)	2,080 (1)	- (0)	1,572 (1)	- (0)	0
FIN	1	2 (1)	52,032 (1)	- (0)	42,247 (1)	12 (1)	0
FRA	1	1 (1)	80,320 (1)	- (0)	806 (1)	- (0)	0
IRN	1	1 (1)	537 (1)	- (0)	537 (1)	10 (1)	1
KOR	1	1 (1)	142,972 (1)	- (0)	138,022 (1)	- (0)	1
NZL	1	1 (1)	1,000 (1)	- (0)	1,000 (1)	- (0)	0
PRT	1	1 (1)	599,276 (1)	235,826 (1)	0 (0.00)	0 (0)	1
PRT/USA	1	2 (1)	719,925 (1)	- (0)	356,475 (1)	- (0)	0
UGA	1	1 (1)	26,995 (1)	- (0)	26,995 (1)	- (0)	0
Total	96	2,438 (88)	63,413,919 (79)	1,804,813 (28)	12,220,866 (76)	3,844 (9)	39

i. The sums represent pooled data from all studies generated in that country.
ii. The number of studies represents how many studies the sum was distributed across.

Figure 12. Studies that used reporting guidelines and types of guidelines used

Study ID	Reporting Guideline	Guideline body
Chu, KH. 2018.[116]	The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement	Enhancing the QUAlity and Transparency Of health Research (EQUATOR) Network
Jones, R. 2019.[150]	The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies.	Enhancing the QUAlity and Transparency Of health Research (EQUATOR) Network
Trivedi, TK. 2019.[124]	The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies	Enhancing the QUAlity and Transparency Of health Research (EQUATOR) Network
Zhang, M. 2019.[126]	Guidelines for Developing and Reporting Machine Learning Predictive Models in Biomedical Research (GRRAS)	Enhancing the QUAlity and Transparency Of health Research (EQUATOR) Network
Joseph, JW. 2020.[127]	i. HIPAA SAFE HARBOR method, ii. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement	i. U.S. Department of Health & Human Services ii. Enhancing the QUAlity and Transparency Of health Research (EQUATOR) Network
Cheung, KY. 2021.[128]	Guidelines for Developing and Reporting Machine Learning Predictive Models in Biomedical Research (GRRAS)	Enhancing the QUAlity and Transparency Of health Research (EQUATOR) Network
Lam, T. 2021.[152]	The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement	Enhancing the QUAlity and Transparency Of health Research (EQUATOR) Network
Rozova, V. 2021[107]	Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement	Enhancing the QUAlity and Transparency Of health Research (EQUATOR) Network

Chapter 4: Conclusion and future work

In this thesis, we have examined ED data from a Canadian context, highlighting the significant value of ED nurses' data, and argued that there is a need to increase the amount of data collected nationally in Canada and to increase nurses' presence in data collection efforts. To demonstrate the utility of these data, we examined triage narratives and determined that these narratives, as well as other data are commonly collected across many regions. These narratives in specific are currently serving several functions to the clinical and academic communities, they are being used for: case identification, quality improvement, as a predictor variable for machine learning models. In this chapter, we will highlight the ongoing and potential uses of our research and discuss future investigations that can build on our findings. We discuss these uses and implications within the contexts of nursing education, nursing practice, nursing research, and nursing policy development.

Nursing education

The current state of ED nursing involvement in the creation and content of academic curriculum is poorly defined. Although there are a few studies from other countries seeking to articulate and standardize emergency nursing education,[175] the work is in the nascent stages and there is apparently none from a Canadian context. Our scoping review has offered insights into the utility of triage nurse narratives for epidemiology and informatics researchers and academics. Future research could move beyond the uses of narrative data for these purposes and could incorporate the narratives from patients to build prototypical presentations. Furthermore, the narratives as a whole could also offer educators insight into what complaints result in patients presenting to the ED, and if specific clusters of complaints are likely to be associated with higher acuity clinical presentations.

Nursing practice

The work involved in this thesis have been used for local quality improvement projects and have shown that triage nurse data, including narratives, can improve clinical care in Canadian EDs.[135] In my role as a clinical nurse educator I led a project which used locally housed triage data, both structured data and narratives to perform large scale audits of triage nurse practice. The structured vital signs data allowed us to identify patients who were assigned incorrect triage acuity scores. The narrative data allowed us to retrospectively extract information about patient pain scores using text parsing algorithms. These measures were used to tailor departmental education efforts, to personalize education feedback to staff, to level workload amongst nurses, and to reduce the frequency of incorrect triage category assignment. The quality improvement efforts resulted in a 68% reduction in the requirement for tailored feedback, 50% reduction in inequitable triage nurse assignment, 86% reduction in high-risk vital sign error rates and 78% reduction in nurses overriding the systems to assign lower acuity triage scores.[135] The findings from this quality improvement initiative have resulted in the creation of an Edmonton Zone triage data dashboard which facilitate large scale audit and feedback by clinical nurse educators. Future research is needed to determine what the effects of these improvements translate to improvements in clinical outcomes for patients. Other national triage systems, specifically the Emergency Severity Instrument, have developed machine learning models with the intentions of incorporating similar feedback to offer nurses real-time feedback as they perform triage,[53] and triage narratives for these purposes may become the next major use for the data source.

Nursing research

The creation of a local database for monitoring triage data has resulted in a robust clinical dataset that can be used to demonstrate that triage nurse data is valuable. The process of completing this scoping review have identified opportunities for adapting study designs used by others to answer questions that are relevant to local researchers and there are a number of projects that are underway that are using triage narratives. These include projects that are in the conception, protocol development, and data analysis phases that will use narrative for predictive modeling and case identification.

I have submitted a funding request to the National Emergency Nurses Association. If awarded, these funds will be used to analyze the structure of triage narratives and offer the first comprehensive description of Canadian triage narratives. The analysis will be performed using the locally developed triage database. The results from our scoping suggest that this project would also be the first comprehensive review of any triage narrative. We would use the gaps identified in our scoping review to set the objectives for this project and would detail the structure (letter and word counts) and form of the narratives and would help to determine the degree of reported and observed data that is included. During the previously mentioned quality improvement project we observed that there were several groups of patients who were likely to have triage narratives that where characteristic of specific clinical conditions who were also being assigned incongruent visit categories Namely, women with complaints characteristic of myocardial infarction who were assigned a category other than cardiac feature chest pain (e.g., abdominal pain, or anxiety). Bias in triage narratives will be explored in a project that is currently in the protocol development phase with the Universities of Calgary and Alberta. These projects would examine the triage narratives of patients who present to the ED and are admitted with myocardial infarctions or acute strokes to determine if there are systemic differences in how men and women are treated when they present with similar complaints to the emergency department. The narratives of these groups will be analyzed using sentiment analysis machine learning models and will be examined for systematically biased language to help determine if there is an association with mis-categorization and clinician bias.

I am also participating in ongoing work in the Edmonton area using triage narratives to identify patients injured using electronic scooters. This project is led by a group from the department of emergency medicine and I have become involved to facilitate the use of triage narratives to identify cases of injuries associated with the devices. Although data extraction has not been completed interim results have demonstrated that the triage narrative identified more cases in a retrospective manner than the combined totals of a retrospective review of all ED cases using ICD codes and prospective recruitment. These findings are comparable to other studies identified in our scoping review which demonstrated that triage narrative can aid or improve the sensitivity and specificity of case identification when compared to or used in conjunction with more traditional approaches to case identification and suggests that the increase in search identification seen in other jurisdictions can likely be extrapolated to the Canadian context.

The final point that was identified by our study that needs to be addressed is the issue of heterogenous reporting of the literature. Future research could address the lack of quality reporting in studies that use ED triage narratives by developing triage data reporting guidelines. Delphi studies that include the voices of emergency nursing associations, emergency nursing and triage scholars, and informatics officers and scholars could identify salient data points that should be included in work of this kind. This work would help to ensure that future research into nursing narratives not only reflects the vision of those who collect the data, but would ensure that future research can be meaningfully compared.

Nursing policy development

In our first chapter we identified that there is a need for unified action on behalf of the National Emergency Nurses Association to define a nursing informatics strategy, to determine the current expertise level of their members, and to highlight the importance of the triage narratives. This is not unique to Canada: in what appears to be the only review of emergency nursing clinical practice guidelines Jones et al., determined that none of the professional emergency nursing bodies examined had explicit guidelines for data collection or informatics strategies.[176] This 2015 international review examined the professional practice statements of ED nurses associations from Canada, the United States, Australia, New Zealand and the United Kingdom and listed where each association made recommendations. Troublingly there were no Canadian ED nursing guidelines for: research, quality improvement, or leadership and professional development. Based on the results of our scoping review we would assert that the absence of guidance and leadership from professional associations has resulted in the large-scale use of triage data with little in the way of acknowledgement that the data was nurse generated or drive to improve the systems for collecting these data.

Limitations

The results from our discussion paper, scoping review protocol and results paper have limitations. Discussion or commentary papers can elucidate problems and propose courses of action but are non-binding and do not necessarily change practice. The scoping review protocol was the first of its type. Although it offers a framework for future reviews, the screening and data extraction forms are specific to the clinical question and may not be relevant to more focused reviews. There was a noted absence of search terms or filters dedicated to triage narratives. Despite the large volumes we retrieved, there might have been relevant studies that were missed. Because of the large volumes of studies that were screened there might also have been errors in the screening process resulting in missed studies. There was also a notable need to examine large amounts of primary literature. The scoping review is limited to describing the data it reviewed, as such we cannot make claims on the usefulness of the data without performing additional analyses. Finally, any review study has the potential for bias in data analysis and interpretation. We minimized these by registering and following a published review protocol.

References

1. Picard C, Kleib M. Advancing Emergency Nurses’ Leadership and Practice through Informatics : The unharnessed power of nurses’ data. *Can J Emerg Nurs*. 2020;43(3):13-17. doi:10.29173/cjen37

2. Picard CT, Kleib M, O’Rourke HM, Norris CM, Douma MJ. Emergency nurses’ triage narrative data, their uses and structure: a scoping review protocol. *BMJ Open*. 2022;12(4):e055132. doi:10.1136/bmjopen-2021-055132

3. Grafstein E, Unger B, Bullard M, Innes G. Canadian Emergency Department Information System (CEDIS) Presenting Complaint List (Version 1.0). *CJEM*. 2003;5(1):27-34. doi:10.1017/s1481803500008071

4. Canadian Institute for Health Information. NACRS data elements 2018–2019. Published 2018. Accessed May 15, 2020. https://www.cihi.ca/sites/default/files/document/nacrs-dataelements-2018-2019-en.pdf

5. Canadian Institute for Health Research. Data Quality Documentation, National Ambulatory Care Reporting System — Current-Year Information, 2017–2018. Published online 2018. https://www.cihi.ca/sites/default/files/document/current-year-information-nacrs-2017- 2018-en-web.pdf

6. Innes G, Murray M, Grafstein E. A consensus-based process to define standard national data elements for a Canadian emergency department information system. *CJEM*. 2001;3(4):277-284. doi:10.1017/s1481803500005777

7. Canadian Institute for Health Information. NACRS Emergency Department Visits and Length of Stay by Province/Territory, 2018–2019. Published online 2019. https://www.cihi.ca/en/nacrs-emergency-department-visits-and-length-of-stay-2018-2019-xlxs

8. Grafstein E, Bullard MJ, Warren D, Unger B, CTAS National Working Group. Revision of the Canadian Emergency Department Information System (CEDIS) Presenting Complaint List version 1.1. *CJEM*. 2008;10(2):151-173. doi:10.1017/s1481803500009878

9. Rowe BH, Bond K, Ospina MB, et al. Data collection on patients in emergency departments in Canada. *CJEM*. 2006;8(6):417-424. doi:10.1017/s1481803500014226

10. Hinson JS, Martinez DA, Cabral S, et al. Triage Performance in Emergency Medicine: A Systematic Review. *Ann Emerg Med*. 2019;74(1):140-152. doi:10.1016/j.annemergmed.2018.09.022

11. Farrohknia N, Castrén M, Ehrenberg A, et al. Emergency department triage scales and their components: a systematic review of the scientific evidence. *Scand J Trauma Resusc Emerg Med*. 2011;19:42. doi:10.1186/1757-7241-19-42

12. Dong SL, Bullard MJ, Meurer DP, et al. Emergency triage: comparing a novel computer triage program with standard triage. *Acad Emerg Med Off J Soc Acad Emerg Med*. 2005;12(6):502-507. doi:10.1197/j.aem.2005.01.005

13. Dong SL, Bullard MJ, Meurer DP, et al. Predictive validity of a computerized emergency triage tool. *Acad Emerg Med*. 2007;14(1):16-21.

14. Canadian Nurses Association. CIHI report reveals optimistic signposts for future of the nursing profession. Published 2019. Accessed May 15, 2020. https://www.cna-aiic.ca/en/news-room/news-releases/2019/cihi-report-reveals-optimistic signposts-for-future-of-the-nursing-profession

15. Canadian Nurses Association. Advancing an essential clinical dataset in Canada (Infographic). Published 2019. Accessed May 15, 2020. https://www.cna-aiic.ca/-/media/cna/page-content/pdf-en/advancing-an-essential-clinical data-set-in-canada-infographic.pdf

16. Canadian Association of Emergency Physicians. CEDIS National Working Group. CEDIS National Working Group. Accessed May 15, 2020. https://caep.ca/resources/cedis/

17. Nagle L. Nursing Data Standards in Canada. In: *2018 National Nursing Data Standards Symposium.* https://cnia.ca/resources/Documents/NNDS%202018%20FINAL%20Proceedings.pdf; 2018:7-9. https://cnia.ca/resources/Documents/NNDS%202018%20FINAL%20Proceedings.pdf

18. International Medical Informatics Association. IMIA-NI definition of nursing informatics updated. Published 2009. Accessed May 15, 2020. https://imianews.wordpress.com/2009/08/24/imia-ni-definition-ofnursing-informatics-up dated/

19. Canadian Nurses Association, Canadian Nursing Informatics Association. Nursing Informatics Joint Position Statement. Published 2017. Accessed May 15, 2020. https://www.cna-aiic.ca/en/~/media/cna/page-content/pdf-fr/nursing-informatics-joint-position-statement

20. Canadian Nurses Association. E-nursing strategy. Published 2006. Accessed May 15, 2020. https://www.cna-aiic.ca/~/media/cna/page-content/pdf-en/e-nursing-strategy-for-canada.p df ?la=en

21. Government of Western Australia, Department of Health. Operational directive OD0205/09: Emergency department and emergency services patient-level data collection and reporting. Published 2009. Accessed May 15, 2020. https://ww2.health.wa.gov.au/About-us/Policy-frameworks/Information-Management/Mandatory-requirements/Collection/Emergency-Department-and-Emergency-Services-Patie nt-Level-Data-Collection-and-Reporting

22. Chapman WW, Dowling JN, Wagner MM. Fever detection from free-text clinical records for biosurveillance. *J Biomed Inform*. 2004;37(2):120-127. doi:10.1016/j.jbi.2004.03.002

23. McKenzie K, Scott DA, Campbell MA, McClure RJ. The use of narrative text for injury surveillance research: a systematic review. *Accid Anal Prev*. 2010;42(2):354-363. doi:10.1016/j.aap.2009.09.020

24. Indig D, Copeland J, Conigrave KM, Rotenko I. Why are alcohol-related emergency department presentations under-detected? An exploratory study using nursing triage text. *Drug Alcohol Rev*. 2008;27(6):584-590. doi:10.1080/09595230801935680

25. Sterling NW, Patzer RE, Di M, Schrager JD. Prediction of emergency department patient disposition based on natural language processing of triage notes. *Int J Med Inf*. 2019;129:184-188. doi:10.1016/j.ijmedinf.2019.06.008

26. Horng S, Sontag DA, Halpern Y, Jernite Y, Shapiro NI, Nathanson LA. Creating an automated trigger for sepsis clinical decision support at emergency department triage using machine learning. *PloS One*. 2017;12(4):e0174708. doi:10.1371/journal.pone.0174708

27. Mitchell R, Finch C, Boufous S, Browne G. Examination of triage nurse text narratives to identify sports injury cases in emergency department presentations. *Int J Inj Contr Saf Promot*. 2009;16(3):153-157. doi:10.1080/17457300903024178

28. Indig D, Copeland J, Conigrave KM. Comparing methods of detecting alcohol-related emergency department presentations. *Emerg Med J EMJ*. 2009;26(8):596-600. doi:10.1136/emj.2008.067348

29. Canadian Association of Schools of Nursing. Nursing informatics entry-to-practice competencies for registered nurses. Published 2012. Accessed May 15, 2020. https://www.casn.ca/2014/12/casn-entry-practice-nursing-informatics-competencies/

30. Canada Health Infoway. National survey of Canadian nurses: Use of digital health technology in practice. Published 2017. Accessed May 15, 2020. https://www.infoway-inforoute.ca/en/component/edocman/3320-2017-national-survey-of-canadian-nurses-use-of-digital-health-technology-in-practice/view-document

31. Kleib M, Nagle L. Factors Associated With Canadian Nurses' Informatics Competency. *Comput Inform Nurs CIN*. 2018;36(8):406-415. doi:10.1097/CIN.0000000000000434

32. National Emergency Nurses Association. Emergency nursing scope and standards of Canadian practice, sixth edition. Published 2018. Accessed May 15, 2020. http://nena.ca/w/wp-content/uploads/2018/11/Standards-of-ED-Nursing-Practice-2018.pdf

33. National Emergency Nurses Association. Emergency nursing core competencies. Published 2014. Accessed May 15, 2020. http://nena.ca/w/wp-content/uploads/2014/11/NENA-Core-Competencies.pdf

34. National Emergency Nurses Association. Minimum requirements for an ED nurse. Published 2014. Accessed May 15, 2020. http://nena.ca/w/wp-content/uploads/2014/11/Minimum-requirements-for-Ed-nurse.pdf

35. Yoon S, Yen PY, Bakken S. Psychometric properties of the self-assessment of nursing informatics competencies scale. *Stud Health Technol Inform*. 2009;146:546-550.

36. Hunter KM, McGonigle DM, Hebda TL. TIGER-based measurement of nursing informatics competencies: The development and implementation of an online tool for self-assessment. *J Nurs Educ Pract*. 2013;3(12):70. doi:10.5430/jnep.v3n12p70

37. Kleib M, Nagle L. Psychometric Properties of the Canadian Nurse Informatics Competency Assessment Scale. *Comput Inform Nurs CIN*. 2018;36(7):359-365. doi:10.1097/CIN.0000000000000437

38. Canadian Association of Schools of Nursing. Digital Health eResource. Published 2019. Accessed May 15, 2020. https://digitalhealth.casn.ca/

39. Remus S, Kennedy MA. Innovation in transformative nursing leadership: nursing informatics competencies and roles. *Nurs Leadersh Tor Ont*. 2012;25(4):14-26. doi:10.12927/cjnl.2012.23260

40. Nakao H, Ukai I, Kotani J. A review of the history of the origin of triage from a disaster medicine perspective. *Acute Med Surg*. 2017;4(4):379-384. doi:10.1002/ams2.293

41. Mayer TA. Triage: history and horizons. *Adv Emerg Nurs J*. 1997;19(2):1-11.

42. Thompson JD, Dains JE. *Comprehensive Triage*. Reston Pub Co; 1982.

43. Council of the Australasian College for Emeregency Medicine, Standards committee. National Triage Scale. *Emerg Med*. 1994;6:145-146.

44. Health Policy Priorities Principal Committee. Australian triage process review, 2011.

45. Beveridge R, Ducharme J, Janes L, Beaulieu S, Walter S. Reliability of the Canadian Emergency Department Triage and Acuity Scale: Interrater agreement. *Ann Emerg Med*. 1999;34(2):155-159. doi:10.1016/S0196-0644%2899%2970223-4

46. Australian Institute of Health and Welfare. Non-admitted patient emergency department care NMDS 2019–20. Accessed April 24, 2022. https://meteor.aihw.gov.au/content/index.phtml/itemId/699738

47. Chapman WW, Dowling JN, Wagner MM. Classification of emergency department chief complaints into 7 syndromes: a retrospective analysis of 527,228 patients. *Ann Emerg Med*. 2005;46(5):445-455. doi:10.1016/j.annemergmed.2005.04.012

48. Gray SE, Finch CF. Assessing the completeness of coded and narrative data from the Victorian Emergency Minimum Dataset using injuries sustained during fitness activities as a case study. *BMC Emerg Med*. 2016;16(1):24. doi:10.1186/s12873-016-0091-4

49. South BR, South BR, Chapman WW, et al. Optimizing A syndromic surveillance text classifier for influenza-like illness: Does document source matter? *AMIA Annu Symp Proc AMIA Symp*. Published online November 6, 2008:692-696.

50. Mitchell RJ, Bambach MR. Examination of narratives from emergency department presentations to identify road trauma, crash and injury risk factors for different age groups. *Health Inf Manag J Health Inf Manag Assoc Aust*. 2015;44(1):21-29. doi:10.1177/183335831504400103

51. Indig D, Copeland J, Conigrave KM, Arcuri A. Characteristics and comorbidity of drug and alcohol-related emergency department presentations detected by nursing triage text. *Addict Abingdon Engl*. 2010;105(5):897-906. doi:10.1111/j.1360-0443.2009.02857.x

52. Vallmuur K, Limbong J, Barker R, Hides L. A comparison of methods to identify alcohol involvement in youth injury-related emergency department presentation data. *Drug Alcohol Rev*. 2013;32(5):519-526. doi:10.1111/dar.12051

53. Ivanov O, Wolf L, Brecher D, et al. Improving ED Emergency Severity Index Acuity Assignment Using Machine Learning and Clinical Natural Language Processing. *JEN J Emerg Nurs*. 2021;47(2):265-265. doi:10.1016/j.jen.2020.11.001

54. Vallmuur K. Machine learning approaches to analysing textual injury surveillance data: a systematic review. *Accid Anal Prev*. 2015;79:41-49. doi:10.1016/j.aap.2015.03.018

55. Hiller KM, Stoneking L, Min A, Rhodes SM. Syndromic surveillance for influenza in the emergency department-A systematic review. *PloS One*. 2013;8(9):e73832. doi:10.1371/journal.pone.0073832

56. Kirubarajan A, Taher A, Khan S, Masood S. Artificial intelligence in emergency medicine: A scoping review. *J Am Coll Emerg Physicians Open*. 2020;1(6):1691-1702. doi:10.1002/emp2.12277

57. Fernandes M, Vieira SM, Leite F, Palos C, Finkelstein S, Sousa JMC. Clinical Decision Support Systems for Triage in the Emergency Department using Intelligent Systems: a Review. *Artif Intell Med*. 2020;102:N.PAG-N.PAG. doi:10.1016/j.artmed.2019.101762

58. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol*. 2005;8(1):19-32. doi:10.1080/1364557032000119616

59. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. *Implement Sci IS*. 2010;5:69. doi:10.1186/1748-5908-5-69

60. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Ann Intern Med*. 2018;169(7):467-473. doi:10.7326/M18-0850

61. Khangura JK, Flodgren G, Perera R, Rowe BH, Shepperd S. Primary care professionals providing non-urgent care in hospital emergency departments. *Cochrane Database Syst Rev*. 2012;11:CD002097. doi:10.1002/14651858.CD002097.pub3

62. Scheurwegs E, Luyckx K, Luyten L, Daelemans W, Van den Bulcke T. Data integration of structured and unstructured sources for assigning clinical codes to patient stays. *J Am Med Inform Assoc*. 2016;23(e1):e11-e19. doi:10.1093/jamia/ocv115

63. International Council of Nurses. Nursing definitions: definition of nursing. ICN - International Council of Nurses. Accessed April 24, 2022. https://www.icn.ch/nursing-policy/nursing-definitions

64. Bazyar J, Farrokhi M, Salari A, Khankeh HR. The Principles of Triage in Emergencies and Disasters: A Systematic Review. *Prehospital Disaster Med*. 2020;35(3):305-313. doi:10.1017/S1049023X20000291

65. Leeies M, Ffrench C, Strome T, Weldon E, Bullard M, Grierson R. Prehospital Application of the Canadian Triage and Acuity Scale by Emergency Medical Services. *CJEM*. 2017;19(1):26-31. doi:10.1017/cem.2016.345

66. Buschhorn HM, Strout TD, Sholl JM, Baumann MR. Emergency medical services triage using the emergency severity index: is it reliable and valid? *J Emerg Nurs*. 2013;39(5):e55-63. doi:10.1016/j.jen.2011.11.003

67. Ibrahim H, Liu X, Denniston AK. Reporting guidelines for artificial intelligence in healthcare research. *Clin Experiment Ophthalmol*. 2021;49(5):470-476. doi:10.1111/ceo.13943

68. Cairns C, Kang K. *National Hospital Ambulatory Medical Care Survey: 2019 Emergency Department Summary Tables*. National Center for Health Statistics (U.S.); 2022. doi:10.15620/cdc:115748

69. Wang Y, Wang L, Rastegar-Mojarad M, et al. Clinical information extraction applications: A literature review. *J Biomed Inform*. 2018;77:34-49. doi:10.1016/j.jbi.2017.11.011

70. Ben-Israel D, Jacobs WB, Casha S, et al. The impact of machine learning on patient care: A systematic review. *Artif Intell Med*. 2020;103:101785. doi:10.1016/j.artmed.2019.101785

71. Sanchez-Salmeron R, Gomez-Urquiza JL, Albendin-Garcia L, et al. Machine learning methods applied to triage in emergency services: A systematic review. *Int Emerg Nurs*. 2022;60(101472191):101109. doi:10.1016/j.ienj.2021.101109

72. Austin EE, Blakely B, Tufanaru C, Selwood A, Braithwaite J, Clay-Williams R. Strategies to measure and improve emergency department performance: a scoping review. *Scand J Trauma Resusc Emerg Med*. 2020;28(1):55. doi:10.1186/s13049-020-00749-2

73. Australasian College for Emergency Medicine. The Australasian Triage Scale. *Emerg Med Fremantle WA*. 2002;14(3):335-336. doi:10.1046/j.1442-2026.2002.00371.x

74. Harding KE, Taylor NF, Leggat SG. Do triage systems in healthcare improve patient flow? A systematic review of the literature. *Aust Health Rev Publ Aust Hosp Assoc*. 2011;35(3):371-383. doi:10.1071/AH10927

75. FitzGerald G, Jelinek GA, Scott D, Gerdtz MF. Emergency department triage revisited. *Emerg Med J EMJ*. 2010;27(2):86-92. doi:10.1136/emj.2009.077081

76. Conway M, Dowling JN, Chapman WW. Using chief complaints for syndromic surveillance: a review of chief complaint based classifiers in North America. *J Biomed Inform*. 2013;46(4):734-743. doi:10.1016/j.jbi.2013.04.003

77. Gandomi A, Haider M. Beyond the hype: Big data concepts, methods, and analytics. *Int J Inf Manag*. 2015;35(2):137-144. doi:10.1016/j.ijinfomgt.2014.10.007

78. Gilboy N, Travers D, Wuerz R. Re-evaluating triage in the new millennium: A comprehensive look at the need for standardization and quality. *J Emerg Nurs*. 1999;25(6):468-473. doi:10.1016/S0099-1767(99)70007-3

79. Barthell EN, Aronsky D, Cochrane DG, Cable G, Stair T, Frontlines Work Group. The Frontlines of Medicine Project progress report: standardized communication of emergency department triage data for syndromic surveillance. *Ann Emerg Med*. 2004;44(3):247-252. doi:10.1016/j.annemergmed.2004.01.020

80. Bradley V. Toward a common language: Emergency Nursing Uniform Data Set (ENUDS). *J Emerg Nurs*. 1995;21(3):248-250. doi:10.1016/s0099-1767(05)80171-0

81. Kung J. A Filter to Retrieve Studies related to Emergency Departments from the EMBASE Database. Published online September 20, 2021. https://docs.google.com/document/d/1sILOgAVTzMI192Mb5kLvVEoC7RO1UoiPm7dDfr-eYgU/edit

82. Campbell SM. A Filter to Retrieve Studies related to Emergency Departments from the OVID MEDLINE Database. Published online April 20, 2021. https://docs.google.com/document/d/1VH1Un8TzC3EXEKfytIF_W7I8-5LwRhU7JcIBDiPpaio/edit#

83. Campbell SM. A Filter to Retrieve Studies related to Emergency Departments from the EBSCO CINAHL Database. Published online June 19, 2021. https://docs.google.com/document/d/10OC5XozEekvZT4P-ymzHrvnXkBheYd5HzUI3-siuofc/edit#

84. Campbell SM. A Filter to Retrieve Studies related to Emergency Departments from the Cochrane Library Databases. Published online May 12, 2021. https://docs.google.com/document/d/1rVKs9q5H1SVu2k5LcCXUdHywPp2MtQ46iyI7UotEbsU/edit

85. Thompson DA, Eitel D, Fernandes CMB, Pines JM, Amsterdam J, Davidson SJ. Coded Chief Complaints-Automated Analysis of Free-text Complaints. *Acad Emerg Med*. 2006;13(7):774-782. doi:10.1197/j.aem.2006.02.013

86. DeYoung K, Chen Y, Beum R, Askenazi M, Zimmerman C, Davidson AJ. Validation of a Syndromic Case Definition for Detecting Emergency Department Visits Potentially Related to Marijuana. *Public Health Rep Wash DC 1974*. 2017;132(4):471-479. doi:10.1177/0033354917708987

87. Vernon N, Maddu K, Hanna TN, Chahine A, Johnson JO, Leonard CE. Emergency department visits resulting from electric scooter use in a major southeast metropolitan area. *Emerg Radiol*. 2020;27(5):469-475. doi:10.1007/s10140-020-01783-4

88. Aronsky D, Kendall D, Merkley K, James BC, Haug PJ. A comprehensive set of coded chief complaints for the emergency department. *Acad Emerg Med*. 2001;8(10):980-989.

89. Day FC, Schriger DL, La M, Day FC, Schriger DL, La M. Automated linking of free-text complaints to reason-for-visit categories and International Classification of Diseases diagnoses in emergency department patient record databases. *Ann Emerg Med*. 2004;43(3):401-409.

90. Lee SH, Levin D, Finley PD, Heilig CM. Chief complaint classification with recurrent neural networks. *J Biomed Inform*. 2019;93(100970413, d2m):103158. doi:10.1016/j.jbi.2019.103158

91. Rhea SK, Weber DJ, Poole C, Waller AE, Ising AI, Williams C. Use of statewide emergency department surveillance data to assess incidence of animal bite injuries among humans in North Carolina. *J Am Vet Med Assoc*. 2014;244(5):597-603. doi:10.2460/javma.244.5.597

92. Peters M, Godfrey C, McInerney P, Munn Z, Tricco A, Khalil H. Chapter 11: Scoping Reviews (2020 version). I. In: *JBI Manual for Evidence Synthesis*. 2020th ed. ; 2020. https://doi.org/10.46658/JBIMES-20-12

93. Kabir A, Hanson R, Mellis CM, van Asperen PP. Is asthma documentation improved by computer-facilitated data entry?. *J Qual Clin Pract*. 1998;18(3):187-193.

94. Goodacre SW, Gillett M, Harris RD, Houlihan KP. Consistency of retrospective triage decisions as a standardised instrument for audit. *J Accid Emerg Med*. 1999;16(5):322-324. doi:10.1136/emj.16.5.322

95. Wagholikar AS, Lawley MJ, Hansen DP, Chu K. Identifying symptom groups from Emergency Department presenting complaint free text using SNOMED CT. *AMIA Annu Symp Proc AMIA Symp*. 2011;2011:1446-1453.

96. Grossmann FF, Frauchiger A, Delport K, Bingisser R, Nickel CH, Zumbrunn T. At risk of undertriage? Testing the performance and accuracy of the emergency severity index in older emergency department patients. *Ann Emerg Med*. 2012;60(3):317-325. doi:10.1016/j.annemergmed.2011.12.013

97. Malmstrom T, Huuskonen O, Torkki P, Malmstrom R. Structured classification for ED presenting complaints - from free text field-based approach to ICPC-2 ED application. *Scand J Trauma Resusc Emerg Med*. 2012;20(101477511):76. doi:10.1186/1757-7241-20-76

98. Haas SW, Travers D, Waller A, et al. Emergency Medical Text Classifier: New system improves processing and classification of triage notes. *Online J Public Health Inform*. 2014;6(2):e178. doi:10.5210/ojphi.v6i2.5469

99. Dehghani Soufi M, Samad-Soltani T, Shams Vahdati S, Rezaei-Hachesu P. Decision support system for triage management: A hybrid approach using rule-based reasoning and fuzzy logic. *Int J Med Inf*. 2018;114((Dehghani Soufi, Samad-Soltani, Rezaei-Hachesu) Department of Health Information Technology, School of Management and Medical Informatics, Tabriz University of Medical Sciences, Tabriz, Iran, Islamic Republic of(Shams Vahdati) Road Traffic Injury Research):35-44. doi:10.1016/j.ijmedinf.2018.03.008

100. Greenbaum NR, Jernite Y, Halpern Y, et al. Improving documentation of presenting problems in the emergency department using a domain-specific ontology and machine learning-driven user interfaces. *Int J Med Inf*. 2019;132:N.PAG-N.PAG. doi:10.1016/j.ijmedinf.2019.103981

101. Klang E, Barash Y, Soffer S, et al. Promoting head CT exams in the emergency department triage using a machine learning model. *Neuroradiology*. 2020;62(2):153-160. doi:10.1007/s00234-019-02293-y

102. Klug M, Barash Y, Bechler S, et al. A Gradient Boosting Machine Learning Model for Predicting Early Mortality in the Emergency Department Triage: Devising a Nine-Point Triage Score. *JGIM J Gen Intern Med*. 2020;35(1):220-227. doi:10.1007/s11606-019-05512-7

103. Mor S, Lev-Rn Z, Tal S. Is family history of coronary artery disease important in the emergency department triage? *Int Emerg Nurs*. 2020;50((Mor) Department of Nursing, The Faculty of Health and Welfare Sciences, University of Haifa, Haifa, Israel; Rambam Health Care Campus, Haifa, Israel(Lev-Rn) Rambam Health Care Campus, Haifa, Israel(Tal) Azrieli Faculty of Medicine, Bar-Ilan University, I):100855. doi:10.1016/j.ienj.2020.100855

104. Choi SW, Ko T, Hong KJ, Kim KH. Machine Learning-Based Prediction of Korean Triage and Acuity Scale Level in Emergency Department Patients. *Healthc Inform Res*. 2019;25(4):305-312. doi:10.4258/hir.2019.25.4.305

105. Travers DA, Haas SW. Using nurses' natural language entries to build a concept-oriented terminology for patients' chief complaints in the emergency department. *J Biomed Inform*. 2003;36(4-5):260-270. doi:10.1016/j.jbi.2003.09.007

106. Sterling NW, Brann F, Patzer RE, et al. Prediction of emergency department resource requirements during triage: An application of current natural language processing techniques. *J Am Coll Emerg Physicians Open*. 2020;1(6):1676-1683. doi:10.1002/emp2.12253

107. Rozova V, Witt K, Robinson J, Li Y, Verspoor K. Detection of self-harm and suicidal ideation in emergency department triage notes. *J Am Med Inform Assoc JAMIA*. 2022;29(3):472-480. doi:10.1093/jamia/ocab261

108. Bregman B, Slavinski S. Using emergency department data to conduct dog and animal bite surveillance in New York City, 2003-2006. *Public Health Rep Wash DC 1974*. 2012;127(2):195-201. doi:10.1177/003335491212700208

109. McKenzie K, Campbell MA, Scott DA, Discoll TR, Harrison JE, McClure RJ. Identifying work related injuries: comparison of methods for interrogating text fields. *BMC Med Inform Decis Mak*. 2010;10((McKenzie) National Centre for Health Information Research and Training, Queensland University of Technology, Victoria Park Road, Kelvin Grove, Queensland, 4059, Australia.):19.

110. Luther M, Gardiner FW, Hansen C, Caldicott D. Hot of Not: Physiological versus Meteorological Heatwaves—Support for a Mean Temperature Threshold. *Int J Environ Res Public Health*. 2016;13(8):753. doi:10.3390/ijerph13080753

111. Rahme E PhD, Low NCP MD, MSc, Lamarre S MD, et al. Correlates of Attempted Suicide from the Emergency Room of 2 General Hospitals in Montreal, Canada/Corrélats des tentatives de suicide du service d'urgence de deux hôpitaux généraux de Montréal, Canada. *Can J Psychiatry*. 2016;61(7):382-393. doi:10.1177/0706743716639054

112. Whitlam G, McGuire R, Ryan T, et al. Diagnosis-based emergency department alcohol harm surveillance: What can it tell us about acute alcohol harms at the population level? *Drug Alcohol Rev*. 2016;35(6):693-701. doi:10.1111/dar.12458

113. Kondis JS, Luhmann JD, Muenzer J. Missed Fractures in Infants Presenting to the Emergency Department with Fussiness. *Pediatr Emerg Care*. 2017;33(8):538-543. doi:10.1097/PEC.0000000000001106

114. Harduar Morano L, Waller AE. Evaluation of the Components of the North Carolina Syndromic Surveillance System Heat Syndrome Case Definition. *Public Health Rep Wash DC 1974*. 2017;132(1 Supplement):40S-47S. doi:10.1177/0033354917710946

115. Zhang X, Kim J, Patzer RE, Pitts SR, Patzer A, Schrager JD. Prediction of Emergency Department Hospital Admission Based on Natural Language Processing and Neural Networks. *Methods Inf Med*. 2017;56(5):377-389. doi:10.3414/ME17-01-0024

116. Chu KH, Brown NJ, Brown AF, et al. Incidence and outcome of subarachnoid haemorrhage in the general and emergency department populations in queensland from 2010 to 2014. *EMA - Emerg Med Australas*. 2018;30(4):503-510. doi:10.1111/1742-6723.12936

117. Gligorijevic D, Stojanovic J, Satz W, et al. Deep attention model for triage of emergency department patients. *Soc Ind Appl Math*. 2018;In Proceedings of the 2018 SIAM International Conference on Data Mining.:297-305.

118. Goldman-Mellor S, Yusheng Jia, Kwan K, Rutledge J, Jia Y. Syndromic Surveillance of Mental and Substance Use Disorders: A Validation Study Using Emergency Department Chief Complaints. *Psychiatr Serv*. 2018;69(1):55-60. doi:10.1176/appi.ps.201700028

119. Hargrove J, Waller A. Motor Vehicle Crash Case Definitions and How They Impact Injury Surveillance. *N C Med J*. 2018;79(6):351-357. doi:10.18043/ncm.79.6.351

120. Nagabhushan M, Webley J. Ripping the Myth: Patients' Symptomatic Descriptions of Acute Thoracic Aortic Dissection. *Spartan Med Res J*. 2018;3(1):6783.

121. Chen M, Cundy P, Antoniou G, Williams N. Children bouncing to the emergency department: Changes in trampoline injury patterns. *J Paediatr Child Health*. 2019;55(2):175-180. doi:10.1111/jpc.14144

122. Eley R, Vallmuur K, Catchpoole J. Value of emergency department triage data to describe and understand patterns and mechanisms of cycling injuries. *Emerg Med Australas*. 2019;31(2):234-240. doi:10.1111/1742-6723.13124

123. Marx GE, Chen Y, Askenazi M, Albanese BA. Syndromic Surveillance of Emergency Department Visits for Acute Adverse Effects of Marijuana, Tri-County Health Department, Colorado, 2016-2017. *Public Health Rep Wash DC 1974*. 2019;134(2):132-140. doi:10.1177/0033354919826562

124. Trivedi HM, Chang P, Sohn JH, et al. Large Scale Semi-Automated Labeling of Routine Free-Text Clinical Records for Deep Learning. *J Digit Imaging*. 2019;32(1):30-37. doi:10.1007/s10278-018-0105-8

125. Rahilly-Tierney C, Altincatal A, Agan A, et al. Linking Ambulance Trip and Emergency Department Surveillance Data on Opioid-Related Overdose, Massachusetts, 2017. *Public Health Rep Wash DC 1974*. 2021;136(1_suppl):47S-53S. doi:10.1177/00333549211011626

126. Zhang X, Bellolio MF, Medrano-Gracia P, Werys K, Yang S, Mahajan P. Use of natural language processing to improve predictive models for imaging utilization in children presenting to the emergency department. *BMC Med Inform Decis Mak*. 2019;19(1):287. doi:10.1186/s12911-019-1006-6

127. Joseph JW, Leventhal EL, Grossestreuer AV, et al. Deep-learning approaches to identify critically Ill patients at emergency department triage using limited information. *J Am Coll Emerg Physicians Open*. 2020;1(5):773-781. doi:10.1002/emp2.12218

128. Cheung KY, Leung LP. Validity and reliability of the triage scale in older people in a regional emergency department in Hong Kong. *Hong Kong J Emerg Med*. 2021;28(2):65-71. doi:10.1177/1024907920971633

129. Mikosz CA, Silva J, Black S, Gibbs G, Cardenas I. Comparison of two major emergency department-based free-text chief-complaint coding systems. *MMWR Suppl*. 2004;53:101-105.

130. Chapman WW, Wagner MM, Ivanov O, et al. Classifying free-text triage chief complaints into syndromic categories with natural language processing. *Artif Intell Med*. 2005;33(1):31-40. doi:10.1016/j.artmed.2004.04.001

131. Handly N, Li J, Chuirazzi DM, Venkat A, Thompson DA. Evaluation of a hospital admission prediction model adding coded chief complaint data using neural network methodology. *Eur J Emerg Med*. 2015;22(2):87-91. doi:10.1097/MEJ.0000000000000126

132. Nanda G, Vallmuur K, Lehto M. Semi-automated text mining strategies for identifying rare causes of injuries from emergency room triage data. *IISE Trans Healthc Syst Eng*. 2019;9(2):157-171. doi:10.1080/24725579.2019.1567628

133. Simonsen L, Gog JR, Olson D, Viboud C. Infectious Disease Surveillance in the Big Data Era: Towards Faster and Locally Relevant Systems. *J Infect Dis*. 2016;214(suppl_4):S380-S385. doi:10.1093/infdis/jiw376

134. Wolf L, Delao A, Perhats C, Baker K, Olson CM. Development of Nurse-Sensitive, Emergency Department-Specific Quality Indicators Using a Modified Delphi Technique. *J Nurs Care Qual*. Published online April 11, 2022. doi:10.1097/NCQ.0000000000000627

135. Cotton R, Drew R, Douma M, et al. An analysis of individual and department triage variances to identify, quantify, and improve markers of triage nurse accuracy. *Can J Emerg Nurs*. 2021;44(2):19-20. doi:10.29173/cjen130

136. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med*. 2007;147(8):573-577. doi:10.7326/0003-4819-147-8-200710160-00010

137. Benchimol EI, Smeeth L, Guttmann A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *PLoS Med*. 2015;12(10):e1001885. doi:10.1371/journal.pmed.1001885

138. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ*. 2015;350:g7594. doi:10.1136/bmj.g7594

139. Baumann MR, Strout TD. Evaluation of the Emergency Severity Index (version 3) triage algorithm in pediatric patients. *Acad Emerg Med*. 2005;12(3):219-224. doi:10.1197/j.aem.2004.09.023

140. Jones PG, Dance A, Shah N. Accuracy of SNOMED-CT Chief Presenting Complaint Codes: an audit of 1,000 cases. *N Z Med J*. 2020;133(1526):67-75.

141. Mosley I, Morphet J, Innes K, Braitberg G. Triage assessments and the activation of rapid care protocols for acute stroke patients. *Australas Emerg Nurs J AENJ*. 2013;16(1):4-9. doi:10.1016/j.aenj.2012.12.002

142. Hendin A, Eagles D, Myers V, Stiell IG. Characteristics and outcomes of older emergency department patients assigned a low acuity triage score. *CJEM*. 2018;20(5):762-769. doi:10.1017/cem.2018.17

143. Berendsen Russell S, Dinh MM, Bell N. Triage, damned triage... and statistics: Sorting out redundancy and duplication within an Emergency Department Presenting Problem Code Set to enhance research capacity. *Australas Emerg Nurs J AENJ*. 2017;20(1):48-52. doi:10.1016/j.aenj.2016.09.004

144. Vico AR, Hernández FS. Nursing triage in acute stroke. *Enferm Glob*. 2021;20(4):120-130. doi:10.6018/eglobal.465261

145. Begier EM, Sockwell D, Branch LM, et al. The National Capitol Region's Emergency Department syndromic surveillance system: do chief complaint and discharge diagnosis yield different results? *Emerg Infect Dis*. 2003;9(3):393-396. doi:10.3201/eid0903.020363

146. Muscatello DJ, Churches T, Kaldor J, et al. An automated, broad-based, near real-time public health surveillance system using presentations to hospital Emergency Departments in New South Wales, Australia. *BMC Public Health*. 2005;5((Muscatello, Churches, Kaldor, Zheng, Chiu, Jorm) Centre for Epidemiology and Research, New South Wales Department of Health, 73 Miller Street, North Sydney, NSW 2059, Australia(Correll) Australian Centre for Asthma Monitoring, PO Box M77, Camperdown, NSW):141. doi:10.1186/1471-2458-5-141

147. Bouchouar E, Hetman BM, Hanley B. Development and validation of an automated emergency department-based syndromic surveillance system to enhance public health surveillance in Yukon: a lower-resourced and remote setting. *BMC Public Health*. 2021;21(1):1247. doi:10.1186/s12889-021-11132-w

148. Rice BT, Bisanzo M, Maling S, Joseph R, Mowafi H, Global Emergency Care Investigators Group (Study Group). Derivation and validation of a chief complaint shortlist for unscheduled acute and emergency care in Uganda. *BMJ Open*. 2018;8(6):e020188. doi:10.1136/bmjopen-2017-020188

149. Hides L, Limbong J, Vallmuur K, Barker R, Daglish M, Young RM cD. Alcohol-related emergency department injury presentations in Queensland adolescents and young adults over a 13-year period. *Drug Alcohol Rev*. 2015;34(2):177-184. doi:10.1111/dar.12218

150. Jones R, Woods C, Usher K, Barker R. Patterns and features of methamphetamine-related presentations to emergency departments in QLD from 2005 to 2017. *Int J Ment Health Nurs*. 2019;28(4):833-844. doi:10.1111/inm.12618

151. Delany C, Crilly J, Ranse J. Drug- and alcohol-related emergency department patient presentations during the 2018 Commonwealth Games: A multi-site retrospective analysis. *EMA - Emerg Med Australas*. 2021;33(5):826-833. doi:10.1111/1742-6723.13746

152. Lam T, Hayman J, Berecki-Gisolf J, Sanfilippo P, Lubman DI, Nielsen S. Pharmaceutical opioid poisonings in Victoria, Australia: Rates and characteristics of a decade of emergency department presentations among nine pharmaceutical opioids. *Addict Abingdon Engl*. 2021;(bm3, 9304118). doi:10.1111/add.15653

153. Burt CW, Overpeck MD. Emergency visits for sports-related injuries. *Ann Emerg Med*. 2001;37(3):301-308. doi:10.1067/mem.2001.111707

154. Liljeqvist HT, Muscatello D, Sara G, Dinh M, Lawrence GL, Liljeqvist HTG. Accuracy of automatic syndromic classification of coded emergency department diagnoses in identifying mental health-related presentations for public health surveillance. *BMC Med Inform Decis Mak*. 2014;14(1):84-84. doi:10.1186/1472-6947-14-84

155. Kuramoto-Crawford SJ, Spies EL, Davies-Cole J. Detecting Suicide-Related Emergency Department Visits Among Adults Using the District of Columbia Syndromic Surveillance System. *Public Health Rep Wash DC 1974*. 2017;132(1_suppl):88S-94S. doi:10.1177/0033354917706933

156. Stapelberg NJC, Randall M, Sveticic J, Fugelli P, Dave H, Turner K. Data mining of hospital suicidal and self-harm presentation records using a tailored evolutionary algorithm. *Mach Learn Appl*. 2021;3:100012. doi:10.1016/j.mlwa.2020.100012

157. Robinson J, Witt K, Lamblin M, et al. Development of a self-harm monitoring system for victoria. *Int J Environ Res Public Health*. 2020;17(24):1-12. doi:10.3390/ijerph17249385

158. Sveticic J, Stapelberg NC, Turner K. Suicidal and self-harm presentations to Emergency Departments: The challenges of identification through diagnostic codes and presenting complaints. *Health Inf Manag J Health Inf Manag Assoc Aust*. 2020;49(1):38-46. doi:10.1177/1833358319857188

159. Howe A, Crilly M. Identification and characteristics of victims of violence identified by emergency physicians, triage nurses, and the police. *Inj Prev J Int Soc Child Adolesc Inj Prev*. 2002;8(4):321-323.

160. Personnic J, Titomanlio L, Auvin S, Dozieres-Puyravel B. Neurological disorders encountered in a pediatric emergency department: Pediatric neurology in emergency room. *Eur J Paediatr Neurol*. 2021;32((Personnic, Titomanlio, Auvin, Dozieres-Puyravel) Pediatric Emergency Department, Assistance Publique des Hopitaux de Paris, Robert Debre Hospital, Boulevard Serrurier, Paris 75019, France):86-92. doi:10.1016/j.ejpn.2021.03.017

161. Irvine AK, Haas SW, Sullivan T. TN-TIES: A system for extracting temporal information from emergency department triage notes. *AMIA Annu Symp Proc AMIA Symp AMIA Symp*. 2008;((Irvine, Haas, Sullivan) University of North Carolina, Chapel Hill, NC, USA.):328-332.

162. Genes N, Chandra D, Ellis S, Baumlin K. Validating emergency department vital signs using a data quality engine for data warehouse. *Open Med Inform J*. 2013;7(101480483):34-39. doi:10.2174/1874431101307010034

163. Gillam C, Versluis L, Hendrie D, Sprivulis P, Meuleners L. Electronic injury surveillance in Perth emergency departments: Validity of the data. *EMA - Emerg Med Australas*. 2007;19(4):309-314. doi:10.1111/j.1742-6723.2007.00942.x

164. Bacchi S, Gluck S, Tan Y, et al. Prediction of general medical admission length of stay with natural language processing and deep learning: a pilot study. *Intern Emerg Med*. 2020;15(6):989-995. doi:10.1007/s11739-019-02265-3

165. Roquette BP, Maiorano AC, Nagano H, Marujo EC. Prediction of admission in pediatric emergency department with deep neural networks and triage textual data. *Neural Netw*. 2020;126((Roquette, Maiorano) Data Science Team, Itau Unibanco, Praca Alfredo Egydio de Souza Aranha, 100, torre WMS, 10th floor, Sao Paulo-SP 04344-902, Brazil(Nagano) Fundacao Getulio Vargas (FGV/EAESP), Rua Itapeva, 474/9th floor, Sao Paulo-SP 01332-000, Br):170-177. doi:10.1016/j.neunet.2020.03.012

166. Tahayori B, Akhlaghi H, Chini-Foroush N. Advanced natural language processing technique to predict patient disposition based on emergency triage notes. *EMA - Emerg Med Australas*. 2021;33(3):480-484. doi:10.1111/1742-6723.13656

167. Fernandes M, Mendes R, Vieira SM, et al. Risk of mortality and cardiopulmonary arrest in critical patients presenting to the emergency department using machine learning and natural language processing. *PLoS ONE*. 2020;15(4):e0230876. doi:10.1371/journal.pone.0230876

168. Klang E, Levin MA, Soffer S, et al. A Simple Free-Text-like Method for Extracting Semi-Structured Data from Electronic Health Records: Exemplified in Prediction of In-Hospital Mortality. *Big Data Cogn Comput*. 2021;5(3):40. doi:10.3390/bdcc5030040

169. Zhang X, Kim J, Schrager JD, Patzer RE, Pitts SR, Chokshi FH. Advanced diagnostic imaging utilization during emergency department visits in the United States: A predictive modeling study for emergency department triage. *PLoS ONE*. 2019;14(4):e0214905. doi:10.1371/journal.pone.0214905

170. Fernandes M, Mendes R, Vieira SM, et al. Predicting intensive care unit admission among patients presenting to the emergency department using machine learning and natural language processing. *PLoS ONE*. 2020;15(3):e0229331. doi:10.1371/journal.pone.0229331

171. Klang E, Kia MA, Timsina P, et al. Predicting adult neuroscience intensive care unit admission from emergency department triage using a retrospective, tabular-free text machine learning approach. *Sci Rep*. 2021;11(1):1381. doi:10.1038/s41598-021-80985-3

172. Rhea S, Ising A, Fleischauer AT, Deyneka L, Vaughan-Batten H, Waller A. Using near real-time morbidity data to identify heat-related illness prevention strategies in North Carolina. *J Community Health*. 2012;37(2):495-500. doi:10.1007/s10900-011-9469-0

173. Petruniak L, El-Masri M, Fox-Wasylyshyn S. Exploring the Predictors of Emergency Department Triage Acuity Assignment in Patients With Sepsis. *Can J Nurs Res Rev Can Rech En Sci Infirm*. 2018;50(2):81-88. doi:10.1177/0844562118766178

174. Metzger P, Allum L, Sullivan E, Onchiri F, Jones M. Racial and Language Disparities in Pediatric Emergency Department Triage. *Pediatr Emerg Care*. 2021;((Metzger) From the Department of Pediatrics, University of Washington Heart Center, Seattle Children’s Hospital Seattle Children’s Research Institute Division of Emergency Medicine, Department of Pediatrics, University of Washington, Seattle, WA). doi:10.1097/PEC.0000000000002439

175. Theobald KA, Coyer FM, Henderson AJ, Fox R, Thomson BF, McCarthy AL. Developing a postgraduate professional education framework for emergency nursing: a co-design approach. *BMC Nurs*. 2021;20(1):43. doi:10.1186/s12912-021-00560-z

176. Jones T, Shaban RZ, Creedy DK. Practice standards for emergency nursing: An international review. *Australas Emerg Nurs J*. 2015;18(4):190-203. doi:10.1016/j.aenj.2015.08.002