Data Science Trends Relevant to Nursing Practice: A Rapid Review of the 2020 Literature

Brian J. Douthit[1]   Rachel L. Walden[2]   Kenrick Cato[3]   Cynthia P. Coviak[4]   Christopher Cruz[5]
Fabio D'Agostino[6]   Thompson Forbes[7]   Grace Gao[8]   Theresa A. Kapetanovic[9]   Mikyoung A. Lee[10]
Lisiane Pruinelli[11]   Mary A. Schultz[12]   Ann Wieben[13]   Alvin D. Jeffery[14]

[1] Tennessee Valley Healthcare System, U.S. Department of Veterans Affairs; Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, Tennessee, United States
[2] Annette and Irwin Eskind Family Biomedical Library, Vanderbilt University, Nashville, Tennessee, United States
[3] Department of Emergency Medicine, Columbia University School of Nursing, New York, New York, United States
[4] Professor Emerita of Nursing, Grand Valley State University, Allendale, Michigan, United States
[5] Global Health Technology and Informatics, Chevron, San Ramon, California, United States
[6] Department of Medicine and Surgery, Saint Camillus International University of Health Sciences, Rome, Italy
[7] College of Nursing, East Carolina University, Greenville, North California, United States
[8] Department of Nursing, St Catherine University, Saint Paul, Minnesota, United States

Address for correspondence   Alvin D. Jeffery, PhD, RN-BC, CCRN-K, FNP-BC, 461 21st Avenue South, Nashville, TN 37240, United States (e-mail: alvinjeffery@gmail.com).

[9] College of Nursing, East Carolina University, Greenville, North California, United States
[10] College of Nursing, Texas Woman's University, Denton, Texas, United States
[11] School of Nursing, University of Minnesota, Minneapolis, Minnesota, United States
[12] Department of Nursing, California State University, San Bernardino, California, United States
[13] School of Nursing, University of Wisconsin-Madison, Wisconsin, United States
[14] School of Nursing, Vanderbilt University; Tennessee Valley Healthcare System, U.S. Department of Veterans Affairs, Nashville, Tennessee, United States

Abstract

Background   The term "data science" encompasses several methods, many of which are considered cutting edge and are being used to influence care processes across the world. Nursing is an applied science and a key discipline in health care systems in both clinical and administrative areas, making the profession increasingly influenced by the latest advances in data science. The greater informatics community should be aware of current trends regarding the intersection of nursing and data science, as developments in nursing practice have cross-professional implications.

Objectives   This study aimed to summarize the latest (calendar year 2020) research and applications of nursing-relevant patient outcomes and clinical processes in the data science literature.

Methods   We conducted a rapid review of the literature to identify relevant research published during the year 2020. We explored the following 16 topics: (1) artificial intelligence/machine learning credibility and acceptance, (2) burnout, (3) complex care (outpatient), (4) emergency department visits, (5) falls, (6) health care–acquired infections, (7) health care utilization and costs, (8) hospitalization, (9) in-hospital mortality, (10) length of stay, (11) pain, (12) patient safety, (13) pressure injuries, (14) readmissions, (15) staffing, and (16) unit culture.

Results   Of 16,589 articles, 244 were included in the review. All topics were represented by literature published in 2020, ranging from 1 article to 59 articles.

Keywords
► data analytics
► artificial intelligence
► nursing research
► outcome and process assessment

Numerous contemporary data science methods were represented in the literature including the use of machine learning, neural networks, and natural language processing.

Conclusion   This review provides an overview of the data science trends that were relevant to nursing practice in 2020. Examinations of such literature are important to monitor the status of data science's influence in nursing practice.

Background and Significance

Data science has become a ubiquitous term in health care. As we have transitioned from the "big data problem" to embracing new opportunities to deploy large-scale analytics,[1–3] the use of data science methods has become an increasingly important part of all health professions. Nursing is no exception, as data science developments are creating new opportunities to leverage both clinician and patient-generated data to augment nursing practice. Nurses permeate health care across all specialties and clinical areas, from inpatient to community based, bedside to provider, and pediatrics to geriatrics. Given nurses have such varied roles, the influences of data science on nursing can be widespread and could have implications on how nurses make decisions, collaborate with other professions, and provide care to their patients. To assess the impacts of data science on nursing practice, data science trends in nursing-related topics should be periodically examined.

Although it is positive to see the proliferation of data science methods being used in the literature, it can be overwhelming for most to keep abreast of the latest evidence. The idea of conducting a "data science in nursing year in review" was conceived by the members of the Data Science Workgroup of the Nursing Knowledge: Big Data Science Conference[4] hosted annually by the University of Minnesota School of Nursing. To our knowledge, such a review is not available in the literature. By conducting a yearly review, we seek to establish a reliable summary of how data science trends impact and augment nursing practice. Through this effort, nurses and the informatics community can efficiently review relevant studies published in the last year, highlighting strengths as well as areas for improvement with further research.

Objectives

The goal of this work is to provide a succinct rapid review of the literature which reflects data science trends relevant to nursing from the past year (2020). By conducting this review, we aim to inform readers of how nursing is being influenced by data science methods and to reveal general trends of their use in selected topics from the past year.

Methods

To examine the intersections of nursing practice and data science in the past year, we opted to conduct a rapid review of the literature.[5] A rapid review differs from more classic scoping review methodology in that it prioritizes search and appraisal strategies which allows for the identification of general trends in research areas without needing to conduct exhaustive literature searches. As dissemination of yearly contributions must be timely, we opted to provide highlights of data science in tandem with key topics related to nursing practice. While the appraisal methods and search strategies do not produce an exhaustive review of the literature, we believe we have retained enough key contributions to facilitate high-quality discussion and recommendations for future research.

First, we selected 16 topics for review. These topics were selected based on nursing-sensitive indicators as identified in the literature which included patient-, setting-, and nursing-related outcomes.[6,7] The authors (consisting of nursing experts, leaders, and scientists) engaged in discussion regarding the coverage of these topics as they relate to nursing practice, resulting in the addition of the topics, "Artificial Intelligence/Machine Learning Credibility and Acceptance," "Complex Care (Outpatient)," and "Emergency Department Visit," which further represents the diversity of nursing and its presence as a key stakeholder in these areas.

We conducted the review using PubMed and CINAHL databases in March of 2021 for studies published during the past year (2020) in the English language. Other databases were considered (including Embase), but through careful review, we noted that the inclusion of other databases complicated the search by introducing articles that were not relevant to the topics. As we employed a rapid review of the literature, we did not find that the addition of other databases yielded high-impact studies that were not otherwise captured in PubMed and CINAHL. The inclusion of CINAHL allowed us to check for articles that were not included in PubMed and also ensured that nursing-relevant articles were included.

Studies were limited to human studies. We crafted one main search strategy to find studies discussing the use of data science with a combination of keywords and subject headings. We used the following terms to create that strategy: data science, data analytics, artificial intelligence (AI), machine learning (ML), risk assessment, decision-support techniques, clinical prediction rule, natural language processing (NLP), computer-assisted image processing, along with analytic, forecast, prediction, risk, and statistical models. Terms related to "nursing" were not included, as the associated medical subject headings (MeSH) terms did not return additional results. In addition, when used as a

title/abstract term, the corpus of articles increased significantly due to irrelevant articles related to breastfeeding.

We then combined an outcome-specific search strategy with the data science search terms for all 16 topics (►Supplementary Appendix S1 [available in the online version] which presents full search strategies). Articles covering multiple topics were included in each topic summary if all reviewers acknowledged it equally represented each topic. We used the Rayyan web application[8] to perform both abstract and full-text screening. We developed an inclusion and exclusion review form via group consensus with the intention of providing a representative sample of data science publications rather than an exhaustive review of all publications (►Supplementary Appendix S2, available in the online version). We opted for more conservative methods for inclusion to further emphasize the exemplars of each topic. We did not require articles to include nurse authors nor nurse participants. Rather, we focused on including articles of interest that used data science methods relevant to nursing practice. Such studies with nonnursing study populations are useful to highlight, as they could either be applied to nursing-specific practice or could be replicated in the nursing population in the future. Specific interpretations are noted in each topic subsection.

To begin the review process, one reviewer per topic conducted an initial title and abstract review to eliminate nonrelevant articles. Reviewers were selected based on their individual expertise with the topic. Next, the authors conducted a full-text review. If there were questions whether an article should be included, the reviewer would ask a second reviewer to verify. We included publications that were either primary studies, systematic reviews, or meta-analyses. Studies were required to use data science methods which we defined as ML, NLP, unsupervised learning, and image analysis and/or sensor analysis. Studies that primarily used regression were included if they were prediction-focused and used a novel data source or were used in conjunction with other more advanced methods. Studies that used basic statistical tests (e.g., *t*-tests), evaluated psychometric properties, or written as opinion pieces were not included.

After full-text screening, we extracted information related to each study's purpose, study design, data sources, samples, settings, populations, operational definitions of outcomes, predictor variables, and data science methods into a standardized form. We summarized this extracted information for each topic in the results.

Results

Search Results and Screening
Overall, we screened 16,589 abstracts with 244 unique studies being included in this review with 11 instances of study-topic overlap (►Supplementary Appendix S3 [available in the online version] for inclusion/exclusion numbers by outcome). The most represented topics were in-hospital mortality, pain, and length of stay. The least commonly represented topics included unit culture, burnout, and AI/ML credibility and acceptance. ►Fig. 1 illustrates the
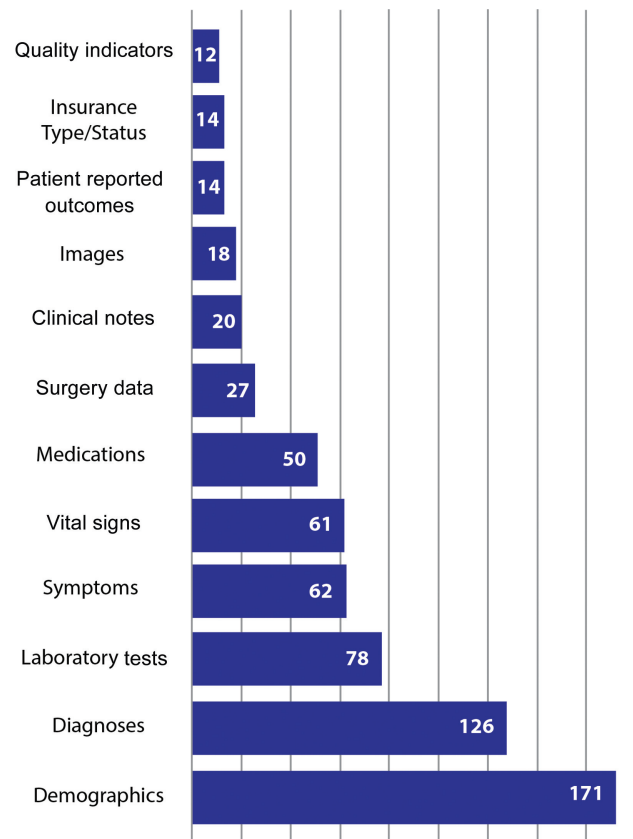

Fig. 1 Twelve most frequent predictor variables among the data science literature relating to nursing practice.

most common predictor variables, among the most common being demographics, diagnoses, and laboratory tests. Evaluating the most represented topics from year to year could help in making recommendations for future study areas.

Several data science methods were identified in the literature as outlined in ►Fig. 2. Generalized linear models were the most common and were only included in this review if they were used in tandem with more advanced methods or were used as prospective, clinical prediction models. More advanced data science methods were also used, including supervised ML ($n = 102$), unsupervised ML ($n = 42$), neural networks ($n = 28$), and NLP ($n = 19$).
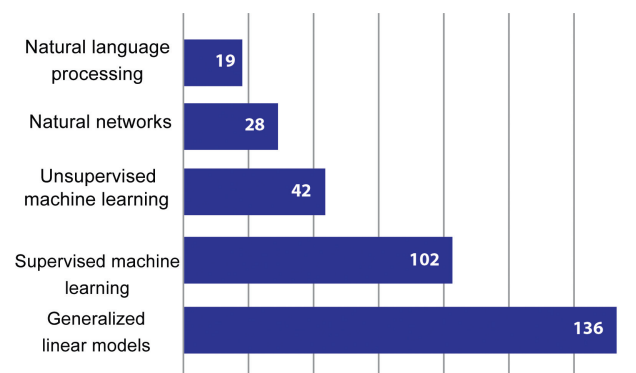

Fig. 2 Use of methods among the data science literature relating to nursing practice.

The data extracted from each study are listed in ►**Supplementary Appendix S4** (available in the online version). The following results are presented by topic, summarizing study designs, data science methods, and implications for nursing practice.

Artificial Intelligence/Machine Learning Credibility and Acceptance

We identified three studies for the topic of AI/ML credibility and acceptance. All three studies used a retrospective cohort design.[9–11] The data sources used by the studies varied among an Italian national injury database,[11] nine public computer-aided diagnostic datasets,[10] and endoscopic images.[9] Study populations included adults[9,10] and an adult/pediatric mixture.[11] One study was based in the United States,[10] one in Korea,[9] and one in Italy.[11] Sample sizes ranged from 13 to 76,911 observations.

Studies of AI/ML credibility and acceptance explored various outcomes. These included violent injury,[11] risk prediction, and colorectal cancer.[9] Studies also used a variety of predictor variables with two[9,10] using demographic and diagnostic data, one using case reports,[11] and one using endoscopic scans.[9] Studies used several different data science methods to address explainability for these two primary outcomes. First, a hybrid model using semantic frames and long short-term memory (LSTM) was used in NLP to extract concepts related to violent injury from notes.[11] Second, an adaptive-weighted method was used with a gradient-boosted classifier (adaBoost) to better understand feature contribution in diagnosis classification.[10] Third, Choi et al[9] combined class activation and neural network learning to display concerning regions that their computer-aided colorectal cancer diagnostic approach identified from colonoscopy images.

AI/ML-based risk stratification tools can support clinicians in decision-making. This is especially true for nurses who are expected to be the last check for most treatments and interventions that patients receive.[12,13] Increased explainability or interpretability may provide nurses with the necessary explanation that builds trust in AI/ML-based advice to support their expertise. However, only one of the studies includes nurses in the research[10] as domain experts. We found studies that approached explainability from a more quantitative analysis rather than a qualitative assessment. For example, these studies tried to improve explainability by increasing interpretation of contributing predictors[10,11] and visually highlighting concerning anatomical areas for human confirmation.[9]

Burnout

In all, only two articles met inclusion criteria for the topic of burnout. One article used applied system dynamics modeling[14] and the other employed an open trial design.[15] Although only two studies were able to be included for this topic, the data sources were somewhat novel: one used synthetic data with an unreported sample size[14] and the other used mobile devices and sensors to collect primary data from 83 medical students. Medical students[14] and nurses[15] were the subjects of the burnout studies. One study[15] was based in Portugal, and the other[14] used synthetic data not tied to any country (but authors were based in Canada).

The two studies approached burnout from different perspectives as follows: one focused on the stress level of nurses,[15] directly measuring physical responses to stress through a smartphone, while the other focused on burnout's effects on quality of care[14] using synthetic "clinician-generated" data. The methods were also very diverse, with the study focusing on system dynamics modeling[14] following a more established mathematical approach, while the study that collected sensor and smartphone data,[15] opted to predict stress levels with an ML model.

Even with the introduction of the quadruple aim in health care,[16] clinician well-being does not seem to be a primary focus in the data science literature. It could be speculated that data relating to clinician burnout is not readily accessible. Many of the study designs in this review are noted to be retrospective, meaning that data had, at one point, already been collected. In the health care field, including nursing, we do not see regular data collection about the clinician. Although a limited sample, this literature shows that it is possible to analyze burnout in clinicians using data science methods, but one of the challenges remains in how to facilitate consistent data capture that is clinician centric. Multiple studies were noted to address electronic health record (EHR) burden in this review but did not use data science methods. Future use of data science may be helpful in furthering our understanding of how to address burnout in health care professionals.

Complex Care (Outpatient)

We identified nine studies for the topic of complex care. Of these studies, five used a retrospective cohort design.[17–21] Other study designs included a combination of a retrospective cohort and a prognostic design built on an ML model,[22] a prognostic design,[23] a longitudinal analysis of a continuously recruited national cohort,[24] and a comparative design with a retrospectively identified cohort which was then matched to a referent cohort from the general population.[25] A majority of studies used administrative database.[17–19,21,22] Two studies used EHR data,[20,23] while the remaining two studies used either data warehouse/registry from the National Patient Register[25] or a questionnaire/survey.[24] Study populations included older adults with intellectual disability,[25] home health care,[19,20] complex care needs,[17,24] adults with cancer,[23] veterans with diabetes,[22] Medicare recipients with dementia,[18] and sepsis survivors.[21] Most studies were based in the United States, but other study locations included Sweden,[25] Canada,[17] and New Zealand.[24] Average sample sizes were large, ranging from 7,936 to 275,190 observations.

Studies of complex care explored various outcomes. These included hospitalizations,[17–19] emergency department (ED) use,[17,19] mortality,[21–23] hospice use,[21] health care utilization,[20,25] and falls risk.[24] Studies used a wide range of predictor or explanatory variables, including home health care agency characteristics,[19] continuity of primary and

specialty physician care,[17] prognostic indices based on patient demographics, comorbidities, procedure codes, laboratory values and anthropomorphic measurements, medication history, and previous health service utilization,[22] patients' demographic characteristics, comorbidities,[21,25] clinical characteristics,[21] racial/ethnic disparities,[20] dementia severity,[18] and urinary and fecal incontinence.[24] Regression was a popular method being used in eight studies with a variety of approaches. Six studies used either multivariable or multilevel regression model to find predictors of home health care agency characteristics for hospitalization and emergency room visits,[19] predictors of patients' demographic characteristics, comorbidities, and clinical characteristics for mortality,[21] predictors of racial/ethnic disparities for health care utilization,[20] predictors of dementia severity for hospitalization,[18] associations of urinary and fecal incontinence with fall risks,[24] and associations of different diagnoses and specialist psychiatric health care utilization.[25] One study used a Cox's regression model to explore continuity of primary and specialty physician care for hospitalization and emergency room visits.[17] A second study used a combination of regression and ML methods to select variables associated with mortality risk and create prognostic indices for 5- and 10-year mortality.[22] A third study used a gradient-boosted ML algorithm to predict 180-day mortality among outpatients with cancer.[23]

Outpatient complex care in reviewed studies often occurred in home health care in the community setting as a continuation or transition of care from hospital settings. Data science methods relied heavily on administrative databases and sometimes on EHR data. Management of complex care requires comprehensive data sources and inputs of health care teams, and it might obscure nursing specific data and render nursing specific data not easily distinguishable. Outcomes and measures for complex care used to build prediction models reflect the all-encompassing nature of addressing complex care management that involves the whole health care team. The reviewed studies demonstrated a lack of electronic data to represent nurses' presence and contributions in home health care. There also appeared to be a lack of method diversity in building predictive models or exploring associations of variables and outcomes related to outpatient complex care.

Emergency Department Visits

For the topic of ED visits, we identified 14 studies. These studies used a retrospective cohort design except for one[26] which used a prospective cohort design. A vast majority of studies used EHR data, while two studies used administrative and claims as the primary dataset.[27,28] Study populations included adults in the ED,[26–37] home care patients,[38] and a mixture of adult and pediatric ED patients.[39] Most studies were based in the United States, but other study locations included Hong Kong,[27] Germany,[32] Italy,[39] Portugal,[37] and South Korea.[34,35] Sample size ranged from 199 to 2,910,321 observations.

Outcomes addressed in studies included mortality,[29,30] future posttraumatic stress disorder (PTSD) sequelae,[26] the novel coronavirus disease 2019 (COVID-19) infection status,[31,35,39] ED wait time,[27] intensive care unit (ICU) admission from the ED,[37] need for head computed tomography (CT),[32] cardiac arrest,[34] stroke severity,[28] and ED utilization.[30,33,36,38] ML methods were used in several studies, including the use of logistic regression, generalized linear models, neural network, and decision tree–based models to apply statistical learning to the prediction of deterioration,[30,37] stroke severity,[28] COVID-19 diagnosis,[31,35,39] wait time,[27] and need for head CT,[32] while the other study used autoregressive integrated moving average to explore time-dependent patient flow.[33,36] NLP was an alternative method used in four studies. Two used NLP to predict patient deterioration,[37,38] one used NLP to extract concepts related to the need for a head CT,[32] and another identified stroke-related concepts from notes to aid in stroke scoring.[28]

The ED is a highly collaborative setting where the medical and nursing domains often overlap. Most of the ED-specific AI/ML studies were related to both nurses and physicians, except for one that predicted ICU admission[37] which is in the physician domain. Many studies have the potential to impact nurse's future clinical practice. First, the study by Schulte-braucks et al[26] prospectively creates risk prediction for the development of PTSD after ED visits. This work may influence the discharge teaching that patients at high risk for PTSD will receive from nurses. Second, three studies modeled ED utilization[33,36] and wait time,[27] including the innovative use of weather as a predictive factor. These studies promise to solve the intractable problem of nurse surge (short-term) staffing where it is difficult to understand who will be entering the ED for care. Third, the study by Topaz et al[38] in the home care setting helps to risk stratify prehospital patients at high risk for ED visits. This study may help EDs to forecast home care patients that will be visiting the hospital. Fourth, nurses are increasingly being asked to collect patient's socioeconomic status (SES) data in the hospital. Schuler et al[30] used SES data in their modeling to improve health care utilization prediction. Finally, 2020 was the year of the COVID-19 pandemic, with ED's being impacted significantly. There were fewer ED COVID-19 papers than expected, possibly because ED clinicians have been too burdened by work demands to publish. However, three studies used data science methods to help answer ED COVID-19 clinical questions: if computer vision could be used to aid in diagnosing COVID-19-related pneumonia,[35] if EHR data predict COVID-19 absent laboratory test confirmation,[39] and if COVID-19 predicts routine blood tests.[31]

Falls

For the topic of falls, we identified 24 studies that met inclusion criteria. Of these studies, eight used a retrospective cohort design[40–47]; seven used a prospective cohort design[48–54]; six were secondary analyses of research data obtained from prospective, retrospective, and cross-sectional studies[55–60]; one used mixed methods wherein data from a public dataset were used in conjunction with measurements collected from sensors[61]; and one was a meta-analysis of prospective cohort and observational studies.[62] Ten of the

studies used health records as a source of data but in two of these studies,[44,47] it was not clear whether the records were electronic when they were obtained. Several of the studies, including two of the secondary analyses, incorporated data from mobility and gait sensors.[48,49,51,53,55,60,61] Registries and administrative datasets were used in eight studies,[40–43,45,46,50,56] while questionnaires or surveys were a source of data for four studies.[49,51,57,60] With the exception of one study that employed sensor data from 17-year-old persons,[55] all study participants were community dwelling, inpatient, and outpatient adults. Adults with chronic diseases of all types were included, but three of the studies included adults with specific conditions. The conditions were postpolio syndrome,[51] neurology, neurosurgery, hematology, oncology,[52] and neurology.[53] Most studies were conducted in the United States, but studies were also completed in Italy,[49] England,[43] Japan,[44,51] Poland,[59] and South Korea.[52] The 11 studies included in the meta-analysis[62] were conducted in seven countries. Sample sizes ranged from 42 to 275,940 observations.

The outcome of falls was defined and measured in a variety of ways. In several studies, the fall was self-reported,[49,51,57,60,63] but if it occurred while the participant was in an inpatient setting or being tracked in an outpatient setting, it was often documented in medical records, registries, or administrative databases used to track adverse events.[40–48,50,52,56] These differences in measuring the outcome are important, in that predictive models may then be more or less accurate, merely because of the accuracy or inaccuracy of the outcome measurement. The types of predictors across studies were quite consistent, nevertheless. Age was a demographic predictor for all studies. Gender was tested but not always a significant predictor. Diagnoses and/or symptoms of the participants were tested as predictors in most of the studies.[40–47,49–52,54,56–60,62,63] Several categories of predictors were noteworthy, including strength, balance, and gait test scores[40,46–57,59–63] and nutritional status.[42,56,59] In 15 studies, prediction models were developed and evaluated with regression models.[40,42–47,51,52,56,57,59,60,62,63] Data science methods also leveraged supervised and unsupervised ML methods, including neural networks for developing risk prediction models, improving prediction of fall risk, and automating selection of data from electronic records for use in fall risk prediction algorithms.[41,45,46,48–50,52–55,58,61]

Data science studies included in this review appeared to reveal a step forward in methods for predicting fall risk. Various activity monitors and robotics technology are capable of creating large datasets of time series tracings that can be examined for patterns suggesting motor movements or muscle weaknesses that predispose a person to falls.[48,49,53,55,60,61] Preprocessing and analysis of such datasets present major challenges that are difficult to manage using traditional statistical techniques and programs, but it is now possible to use ML and other data science methods to determine the patterns in data that are associated with the devastating outcome of falls. From what is observed in the studies published in 2020, the use of devices and sensors is likely to increase in the future exploration of factors that predict falls.

Healthcare-associated Infections

We identified 11 studies for the topic of health care–associated infections (HAIs). Of these 11 studies (five used a retrospective cohort design,[64–68] three used an observational design,[69–71] two used a case-control design,[72,73] and one used a prospective cohort design) were included.[74] A vast majority of studies solely used EHR data, while two studies added to EHR data with breath sensor data[73] and National Database of Nursing Quality Indicator (NDNQI) and Catheter Associated Urinary Tract Infection (CAUTI) datasets.[71] One study used the National Institutes of Health (NIH) Gene Expression Omnibus data.[72] Most studies focused on adult inpatients, while three studies included adult surgical patients[64,65,67] and one study focused on pediatric cardiology surgery patients.[70] Most studies were based in the United States but other study locations included Taiwan,[73] Italy,[66] China,[70] France, and Switzerland.[74] Patients were the unit of analysis for most studies, while one study analyzed ICU admissions,[71] one examined operative events,[64] and one focused on hospitalizations.[68] Samples sizes varied widely from study to study, ranging from 20 to 897,344 observations.

Studies explored various outcomes and included candidemia infection,[74] cardiac implantable device infections,[67] CAUTIs,[71] *Clostridium difficile* infections (CDIs),[69] urinary tract infection (UTIs),[66,68] surgical site infections (SSIs),[64,65] and ventilator-associated pneumonia (VAP).[72,73] The majority of studies used some combination of demographic, diagnosis, laboratory, vital sign, and/or medication data as predictor variables. Several studies used additional unique predictors such as data on patient movement,[69] nurse staffing,[71] breath compounds,[73] and differentially expressed genes.[72] Several different data science methods were used. First, logistic regression was used to predict HAI outcomes in three studies.[67,70,74] Many studies compared the predictive performance of various supervised ML models including support vector machines,[66,71,73] neural networks,[66,68,73] decision trees,[68,71,73] and/or random forest models.[72,73] Two studies used multilayer perceptrons,[65,72] one study used naïve Bayes' classification[73] and one study conducted social network analysis.[69] NLP was used in one study to extract data from clinical notes and operative reports for surveillance of SSIs,[64] and another study used text mining of clinical notes to inform model development and case identification.[67]

Of the studies using data science techniques to examine or predict HAIs, two specifically addressed nursing implications and included nurse authors.[68,71] Park et al demonstrated a knowledge discovery and data mining approach and aimed to describe techniques that could be used to further nursing practice and guide nursing professionals in the use of data science methods. Zachariah et al[68] described the benefit of risk stratification systems in relieving the burden on nurses to complete and document traditional risk assessment forms. While Mancini et al[66] do not specifically name nurses as a

target audience, they do describe their data-science-as-a-service system as an online, user friendly platform that can help domain experts, such as clinicians and validate simple predictive models. From these studies, nursing administrators may gain valuable insights on the role of intrahospital transfers on HAIs to inform patient-placement strategies[69] and the use of predictive risk models in dressing type selection to prevent SSIs and the estimated cost savings.[65] For nurses interested in exemplars of data visualization techniques, the publication by Cai et al[72] showcases some impressive data visualizations.

Health Care Utilization and Costs

For the topic of health care utilization and costs, 24 articles were included in this review. Most were retrospective cohort studies,[75–87] while six used prospective cohort studies,[88–93] four used a cross-sectional design,[94–97] and one used a survey for primary data collection.[98] Most studies used the EHR and administrative databases to collect data but three used surveys,[87,96,98] two used public datasets,[78,95] one used mobile phone data,[97] one used images,[90] and one used data from social media.[92] All studies were adult based with the exception of one study examining families.[98] Most studies were based in the United States, with the exception of three from Singapore,[83,85,93] one from China,[80] one from Brazil,[77] one from Italy,[88] one from Canada,[90] and one from the United Kingdom.[87] Sample sizes ranged from 190 to 780,295 observations.

While most studies focused on cost and included some form of cost analysis, several studies examined behaviors related to costs such as predicting health care utilization,[83] quantifying reliance on health care services,[84] verifying complete surgical removal of tumors,[90] and predicting no-shows.[93] As expected, many studies incorporated costs and insurance status as predictor variables but patient-reported variables were also common. Most studies used supervised ML. Unsupervised learning and linear models were common too, and often, multiple models were compared in search for the most accurate. One study conducted geospatial analysis.[98]

As would be expected, many of the articles included in this review used data science methods to predict cost. Such information would be helpful to hospital administration, but this does not necessarily pertain directly to nursing practice. Instead, nursing may focus its efforts on developing interventions to increase adherence to care. One example may be following-up with patients who are predicted to have a high risk of missing an important magnetic resonance imaging (MRI). While this information would be important for executives to know and potentially avoid loss of revenue, nursing can use this as an opportunity to support continuity of care.

Hospitalization

We identified 21 studies for the topic of hospitalization. Of these studies, 13 used a retrospective cohort design,[68,86,99–109] 1 used an observational design,[110] 1 used a cross-sectional design,[111] 2 adopted a prognostic approach,[30,112] 2 performed a longitudinal analysis,[113,114]

and 2 used survey data.[115,116] A vast majority of studies used EHR data, while the remaining eight studies used administrative databases[100,106,107,109,111,114] or surveys as the primary collection tool.[115,116] Study populations included adults[30] with chronic illnesses,[100,105,107,114] pediatrics,[86,106] veterans,[116] COVID-19 patients,[99,101–103,108,110,113] hospice patients,[104] inpatients,[68,112] and Medicare recipients.[109,111,115] All studies were based in the United States. Sample sizes ranged from 207 to 3,100,000 observations.

Studies of hospitalization explored various areas. These included hospitalization,[101,107,108] hospital readmissions,[110] hospitalization rates,[30,105,106,114] hospitalization risks,[103,113,116] health care utilization,[86,109] level-of-care requirements,[102] medication orders,[112] risk of urinary tract infections (UTI) during hospitalization,[68] risk for critical COVID-19,[99] risk of live discharge,[104] ischemic strokes,[111] recovery of function following hospitalization,[115] and the Functional Independence Measure (FIM) instrument score.[100] Interestingly, most studies used a cluster of various characteristics as predictor variables including demographics,[30,100,103,105,106,108,110,111,113,116] sociodemographic,[106,110,113] or neighborhood SES[30] or neighborhood level characteristics,[104] patient level characteristics,[104] clinical data,[30,68,99–103,105,108,110,113,116] medication data,[112] social determinants of health (SDH) data,[86,111] administrative data,[100] claims data,[30] patient reported outcomes,[108,116] geriatric syndrome risk factors,[109] air quality,[106] and cost trajectories.[115] Two studies used a single variable, either body mass index[107] or food swamp severity,[114] as their predictor variable. Regression was applied in most of the studies. The remaining studies used ML,[30,68,86,99,102,112] NLP,[102] and geospatial coding.[110,114] Studies used ML to build a prediction model of clinical data for risk for critical COVID-19,[99] level-of-care requirements,[102] and risk of UTI during hospitalization,[68] SDH data for health care utilization,[86] and EHR data for medication orders.[112]

The reviewed studies demonstrated a broad range of foci, from unique patient populations and conditions to health care management and utilization. Data science methods employed in these studies incorporated mostly EHR data sources in addition to administrative databases and occasionally survey data. Study outcomes and variables were often a cluster of characteristics that branched into the administrative and clinical domains and occasionally neighborhood and community level of characteristics. Nursing specific data were embedded and not easily distinguished. There appears to be continued needs for nursing specific considerations in studies related to hospitalization using data science methods. However, many outcomes and variables have great implications for nursing care because nursing plays a critical part in health care teams.

In-Hospital Mortality

We identified 59 studies for the topic of in-hospital mortality.[117–175] While the majority of studies used a retrospective cohort design, 11 used a prospective approach,[120,122,129,140,143,166,169,170,172,174,175] and two were meta-analysis studies.[154,157] The majority of studies

used EHR data, while 16 studies used some kind of registry data,[117,118,121,127,128,131,132,138,140–142,150,153,155,156,161,174] four studies used questionaries/surveys,[166,169,170,172] and two studies used administrative data.[134,167] With the emergence of public databases containing COVID-19 data, since 2020, many studies used these databases for their studies.[120,123,128,131,140,142,143,146,169] Study populations primarily comprised adults (sometimes limited to subpopulations such as those with chronic illnesses [e.g., Takada et al[164] and Sukmark et al[163]]), but three studies included pediatric populations,[138,154,161] one study included newborns,[150] and two studies included elder patients aged over 65 years.[121,155] Sample sizes ranged from 15 to 9,000,000 observations.

To predict in-hospital mortality, studies used several methodological approaches. The majority of studies used a regression model including Cox's proportional hazards[131,167] and mixed effect models.[141,158] More contemporary techniques included neural networks,[117,118,125–127,134,139,142,165,171,173] random forests,[124,126,133,135,139,142,144] gradient boosting,[127–130,134,135,139,140,144,168] and NLP.[127] Four studies[137,149,159,173] leveraged unsupervised methods, with or without supervised methods. Almost all studies performed some level of validation, such as bootstrapping, cross-validation, or a hold-out approach. A variety of predictors were used as input for these models. Almost all studies included demographics and medical diagnosis. The majority of the studies also used medications and some sort of diagnostic techniques (e.g., laboratory values, images, vital signs, or surgical data). Some studies used COVID-19-specific data.[120,123,128,131,140,142,143,146,169] Some studies used clinical notes,[127,145,154,170,173] and two studies used socioeconomic data.[126,167] The inclusion of a variety of variables is possible as a result of a large sample size for the majority of the studies.

Notable aspects of the in-hospital mortality literature include the use of frailty as a predictor in two studies, either as a way of predicting mortality or as a better clinical measure for symptom representation,[160,162] as well as the use of portable lung ultrasound findings as predictors.[122] Although many studies included vital signs which are often collected by nurses, there were no studies evaluating how other aspects of nursing care delivery can predict in-hospital mortality. The use of publicly available datasets (e.g., Awad et al,[124] Baxter et al,[127] and Kong et al[144]) facilitates reproducibility and allows future investigators to explore additional data science methods, including the use of novel predictors, such as innovative features generated from text data. Notably, there were limited pediatric/neonatal population studies and limited inclusion of socioeconomic predictors which could be opportunities for future research.

Length of Stay

We identified 26 studies regarding the prediction of the hospital length of stay that used data science methods. Twenty-three studies used a retrospective cohort design,[126,133,159,173,176–194] while three were prospective cohort studies.[129,195,196] Data sources mostly used administrative databases[126,133,179,180,182,186,191,192,194] and

EHRs,[129,133,176,183,184,188,190,192,196] while other studies used publicly available datasets,[159,173,178,187,189] data warehouses and registries,[133,177,180,195] paper clinical notes,[193] paper patient records,[185] research electronic data capture systems,[188] trial datasets,[181] questionnaires,[196] and routine bedside monitors.[176] Sample sizes ranged from 143 to 2,997,249 patients. Study populations included surgical patients,[133,159,177,179,181–183,195,196] ICU patients,[173,176,178,187,189,190] medical-surgical patients,[126,129,180,191] patients presenting to the ED,[184,188,193,194] and psychiatric patients.[185,186,192] Most studies were conducted using U.S. patient data,[129,133,159,173,176–178,181–183,187,189,191] while other studies used patient data from Australia,[126,179,193,194] Brazil,[186,188] Canada,[195,196] China,[180] England,[190] Germany,[192] Switzerland,[185] and Taiwan.[184]

Studies about length of stay also investigated other outcomes, such as mortality,[126,129,133,159,173,178,181,187,188,196] clinical and functional complications (e.g., surgical, respiratory complications, or disability),[126,133,159,183,192,196] readmission,[129,173,182] discharge destination,[126,183,193] patient-reported outcome measures,[182] patient phenotyping,[178] and hospital admission.[188] Demographic data were used as a predictor variable in all studies, while another common predictor was medical diagnosis.[126,133,159,177,178,181,182,183,186,188,190–192,195] Other predictors used clinical data,[126,176–178,180,181,184,185,187,190,194,195] laboratory tests,[126,129,133,173,178,181,182,187,189,190] vital signs,[126,129,133,173,176,178,184,187,189] hospitalization data (e.g., admission/discharge data and hospital characteristics),[126,129,133,159,186,191,192,194] surgery data,[133,177,179,181,182,195,196] anthropometric data,[159,177,178,181,184,187,195] scales/instruments,[126,180,188,192,196] social data,[126,181,185,186,195] medications,[129,133,177,183] insurance status/type,[133,179,180,191] clinical notes,[173,184,193] services used,[183,186,194] and data collected by nurses using the International Classification of Functioning, Disability and Health (ICF).[180] Studies used supervised ML algorithms,[126,129,133,159,176,177,181,183,185,187,189,192–194] generalized linear models,[178,180,182,184,186,188,190–192,195,196] deep learning models,[126,173,178,179,181,187,189,193] as well as unsupervised ML algorithms,[176,186,187] and NLP.[184,193] Among the supervised ML methods, random forest was one of the most used classification algorithms.[126,133,159,176,177,181,183,189,193] Deep learning architectures, such as neural networks, were applied in studies with a large amount of data. Unsupervised algorithms used clustering methods to mine datasets and find patient data features to be used for predicting length of stay. NLP was used to extract data from clinical notes for predicting length of stay and discharge destinations. Interestingly, more than one data science method was used in some studies. For example, in one study,[187] supervised, unsupervised, and deep learning algorithms were applied to develop a predictive model for determining length of stay. In another study,[193] supervised, deep learning algorithms, and NLP were used to predict length of stay and discharge destination.

Future prospective studies are needed for external validation of the models developed. Unstructured data (e.g.,

clinical notes) and structured data (e.g., administrative data) have commonly been used in the studies. However, we did not find any study that used a combination of both. Further studies are required to incorporate these two types of data in the same prediction model because patient information is typically found in unstructured and structured data. Nursing-generated data were mentioned only in two studies with nursing notes and assessment data using a nonmedical classification (i.e., the ICF). Nurses represent the largest health care profession worldwide and the profession that generates the most data about the patient condition; therefore, failing to use these nursing-generated data could become a significant issue. Further studies should use both unstructured and structured nursing-generated data (e.g., standard nursing terminologies) jointly with the commonly used predictors to build prediction models.

Pain

Out of the total of 27 studies that were identified for the topic of pain, 14 used a prospective cohort design,[197–210] 11 used an observational design,[200,204,205,207,210–216] 6 used a retrospective cohort design,[211,214,215,217–219] 4 used a randomized control trial,[201,212,220,221] 1 used a cross-sectional design,[222] and 1 used mixed methods.[223] Most studies used questionnaire/survey data, but eight used administrative databases,[206–208,210,212,220–222] seven used mobile devices/sensors,[200,203–205,210,216,220] and four used a data warehouse or registry.[198,203,208,214] Study populations were mostly done with adults in the outpatient setting but four were inpatient[197,201,211,223] and one was done with a pediatric population.[205] Although many studies were conducted in the United States, others included China,[213–215,222] Australia,[207] Canada,[202] the Netherlands,[199,212] Germany,[208,210,211] Norway,[201] Finland,[204] South Korea,[203] Argentina,[219] Portugal,[197] Japan,[209] and Spain.[206] Sample sizes ranged from 10 to 6,316 observations.

Studies explored various outcomes including surgical applications such as determination of postsurgical measures based on residual pain,[197] predicting patellofemoral pain 1 year after intervention,[201] predicting neuropathic pain,[202] predicting chronic pain of 7 to 10 years into the future,[217] predicting complex regional pain syndrome,[207] predicting pain relief for knee osteoarthritis patients,[209] detection of pain,[210,214,216,222] and pain intensity estimation/classification.[205,213,215,220] Other outcomes focused on pain as a predictor of anxiety and depression, coronary heart disease,[199] health status,[218] non-cancer pain as predictor of brain aging,[208] and length of stay.[211] For patients with low back pain, societal cost,[212] and clinical and sociodemographic predictors of increased disability[221] have been studied. Some outcomes focused on the data science method as a clinical tool such as NLP of pain context from clinical notes.[219,223] There were several novel data sources included, such as the use of physiologic signals from electroencephalograms (EEG),[213] electromyography,[204,220] spectrogram,[205] electrodermal activity,[216] sensor data,[200] MRIs,[198,208,214,222] kinematics/motion data,[203,210,220] and medical images.[197,198,208,210,214,215,222]

Many of these studies have significant impact on nursing, most notably in situations where pain cannot be feasibly assessed (e.g., patients who are nonverbal). The ability to use data science methods for analyzing facial expressions, medical images, vital signs, and other biomechanical data could augment existing conventional methods in classifying and quantifying pain experience. Using EEG and electromyography data have high potential for improving pain assessment. Leveraging ML on geospatial and kinematic data can provide benefits not just for nursing assessment but also in other medical/health disciplines.

Patient Safety

We identified seven studies exploring patient safety. The majority of studies were retrospective cohort designs.[52,224–227] Two used cross-sectional designs.[228,229] Four studies used patient safety or incident reports as primary data for analysis,[224,227–229] two used EHR data,[52,226] and one used a publicly available dataset.[225] Study populations primarily consisted of adult inpatients who had an event or safety report submitted during their inpatient stay.[52,226–229] Studies were based in the United States,[224–226,229] China,[227] Korea,[52] and the United Kingdom.[228] Sample sizes ranged from 348 to 1,740,770 observations.

Studies explored various outcomes, including predicting allergic reactions,[229] classifying medication incidents,[227] identifying falls incidents from event reports,[52,226] identifying drug-to-drug interactions,[225] and classifying the contents of safety reports.[224,228] Data science methods included NLP,[226] deep neural networks,[227,229] support vector machines,[225,228] logistic regression,[52] and naïve Bayes' classification.[224]

Maintaining patient safety in the inpatient setting requires a high level of diligence and oversight by members of the health care team and primarily rests with nurses who provide the majority of care while patients are hospitalized. Patient safety studies using data science methods could advance the health care team's ability to intervene before events occur or improve the efficiency and accuracy in the classification of patient safety events, so that improvement activities are more focused. While studies of patient safety and the reporting of patient safety events are directly related to the daily work of nurses and their diligence at the bedside, only one study was led by a nurse.[52] Two other studies included one nurse in the study team.[224,226]

Pressure Injuries

We identified 13 studies for the topic of pressure injuries (PIs). Of these 13 studies, 7 studies used a retrospective cohort design,[230–236] 3 used a prospective cohort design,[237–239] 1 used a clinical trial,[240] 1 used a cross-sectional design,[241] and 1 used secondary data analysis.[242] A variety of data sources were used for the studies, including EHR data,[233,234,238,239] data warehouses,[230,231,235,236] a publicly available dataset,[232] sensor data,[240,242] and surveys as the primary collection tool.[237] The samples across studies were adult patients admitted in hospitals,[230–232,234,236,238,239]

adults in residential hospices,[237] elderly patients in nursing homes (NHs),[241] Medicare beneficiaries,[235] and adults (unspecified).[240,242] Six studies were based in the United States,[230–232,234–236] with other study locations including Brazil,[239] Canada,[240] France,[242] Indonesia,[241] Italy,[237] South Korea,[238] and Taiwan.[233] Sample sizes ranged from 12 to 2,091,058 observations.

Most studies used the incidence rate of PIs as the outcome, except for one study[234] that projected PI closure and two studies[240,242] that explored PI images. Various data science methods were used to detect or predict PIs including logistic regression,[230–232,234,239] generalized estimating equations,[237] multiple regression,[235] path analysis,[241] supervised ML,[233,236,238] and imaging processing.[240,242] The predominant predictor variables used across studies included demographics and diagnoses,[231–237,239] followed by clinical assessment data,[231,233,235–239,241] Braden's scale,[231,232,236,237,239] laboratory tests,[233,234,236,238,239] and medications.[232,236,239] Two studies used organizational factors such as nursing unit characteristics, nurse job satisfaction, facility types, or rural/urban hospital location.[230,235]

The prevention and management of PIs remains a challenge. The prediction models developed in these studies can help nurses screen high-risk groups and manage risk factors of PIs. The predictive models could create a monitoring system that provides real-time warnings of PI onset or worsening trajectory to nurses and other health care providers and prompt them to personalize PI prevention interventions. The use of bed sheet sensors through PI classification or prediction modeling could develop an automated feedback system with body pressure mapping and consequently, changing posture or redistributing pressure, which would allow remote monitoring.[240] Repositioning in bed could be rescheduled or individualized according to patient conditions. Also, the study by Baernholdt et al[230] on the predictive impact of organizational factors on PI rates suggests that hospitals should focus on organizational structures to improve nurses' work environments and workflow, so that nurses can enhance PI interventions. Although these predictive models are promising, the generalizability and overfitting possibility need to be carefully considered due to the high heterogeneity of samples across studies and the small sample sizes in some studies. Further validation studies of such risk prediction models are needed.

Readmissions

We identified nine studies for the topic of readmissions. Of these nine studies, seven used a retrospective cohort design[243–249] and two used a prospective cohort design.[244,246] Seven studies primarily used EHR data[243–248,250] stored in a data warehouse of the affiliated facility,[243,246,248–250] with one in combination with other data sources that included mobile device sensor data[244] and one with a governmental administrative database (Medicare).[245] Study populations included adults in hospital intensive care,[173,250] those hospitalized with medical conditions,[243] those having had cardiac surgery in a progressive care unit (PCU),[244] and those who utilized Medicare services.[245] Two studies focused on

Medicare patient data,[245,249] and one study of Medicare patients included encounter information from a nonhospital setting (i.e., inpatient rehabilitation, skilled nursing, and home health services).[245] Sample sizes ranged from 100 patients[244] to over 1 million patients.[249] Data in each study were collected from health care facilities in the United States.

Risk prediction outcomes in each study included acute care readmissions occurring at 7, 30, or 90 days of hospital discharge. One study looked at readmission back to the ICU.[250] In addition to acute care readmissions, some data were used to predict length of stay of postoperative cardiac patients,[244] hospital or 180 mortality,[246] and elective surgery mortality at 30 and 90 days.[249] Studies generally included predictor variables comprising length of stay, gender, number of recent admissions, age, surgical procedure, admission location (e.g., ED, clinic, and physician referral), insurance type, diagnosis, procedures, medication, vital signs, and comorbidities. Methods used to predict readmissions included ML,[173,243,244,246–250] NLP,[173,246,248–250] general linear regression,[173,243,247,248] a combination of statistical modeling and ML,[245] and combined structured and unstructured data neural network.[173] Interestingly, Saleh et al[248] used an existing 30-day prediction model to compare strengths of predictors in 7-day readmissions. Only one study focused on social determinants of health that may be predictors in readmissions.[247]

The importance of hospital patient readmissions in a 30-day (or less) time interval is viewed as a quality metric by the Medicare program and other insurers. Reimbursement changes are occurring in government programs that incentivize hospitals for quality and penalize hospitals if quality metrics are not maintained. Nurses have a role in assessment, planning, and implementation of an accurate discharge plan that can help identify patients most at risk for readmission due to health condition, comorbidities, or other risk factors. ML, NLP, and predictive modeling with EHR data can provide valuable information to assist in risk identification of importance to nursing care and discharge planning. As structured and unstructured data in the health record can be combined through the design of multimodal architecture to support understanding of risk reduction, nurses can use these data in the care of at-risk populations.

Staffing

We identified four studies for the topic of staffing. One study used a prospective cohort design[251] and the remaining three used a retrospective cohort design.[252–254] Three of these studies were conducted in the United States[252–254] with one study[251] using data from a single ICU in an Italian medical center. All studies used a combination of EHR data plus administrative or systems data.[251–254] Study populations varied with the adult medical-surgical population used in two studies,[251,253] a NH population in one[252] and one study used a pediatric population.[254] Scheduling or workload studies were not discovered in the search. Sample sizes ranged from 148 to 30,679 observations.

Operational outcomes comprised (pediatric) readmissions,[254] the prediction of adverse events,[251] leaving (ED)

without being seen,[253] and infection risk.[252] Unadjusted logistic regression was used to evaluate each response on the tool (insurance type, home medical equipment, home nursing, home therapy, and others) with weighted scores assigned to each category. In attempting to determine if a tool (the Patient Acuity and Complexity Score) developed for their study of the prediction of adverse events, the Sanson research group (2020) sought to discriminate between patients having experienced/not experienced a serious event in the discharged unit after intensive care was received. In a study of NH quality,[252] tree-based gradient-boosting algorithms were used to evaluate the risk of COVID-19 infection (the presence of at least one confirmed COVID-19 resident in the NH). A logistic regression model and two-layer feedforward neural network were also developed using the identified stable predictors (including the number of care personnel/1,000 feet) to serve as benchmark predictive models for comparison.

Interestingly, only one study reported a traditional measure of staffing,[252] the number of patients per nurse. A new variable, leaving without being seen,[253] could spark further interest in the layered relationships of systems/administrative data when coupled with what is traditionally termed "clinical data," particularly when clear administrative implications emerge as is the case in this study (administrative actions on ED process variables, e.g., wait times). The collection of data in 1-hour increments[253] could also prove a necessary improvement in studies with administrative variables (e.g., door-to-provider time), yet will demand further methodological scrutiny if the variability of certain hourly measures (number of persons in waiting room) outdistances that of nurse or other provider variables known to impact outcomes.

Unit Culture

Only one study explored a unit cultural element using data science methods. This study used the Hospital Survey of Patient Safety Culture to predict if a patient safety event would be voluntarily reported.[255] This study was conducted in the United States with a sample size of 526,645 survey responses.

The study included regression techniques to validate that many of elements of patient safety culture influence the possibility that a patient safety event would be voluntarily reported. Some examples of these elements include communication openness, teamwork, staffing, and hand-offs and transitions. Outcomes explored in this study included frequency of events reported, near-miss events, no potential for harm events, and potential for harm events.

The study included in this review explored how a culture of patient safety influenced voluntary reporting of patient safety events. While an argument could be made that this article may be better suited in the patient safety category, we included this as a unit cultural element due to the impact unit level dynamics have on creating a patient safety culture. More exploration is needed on unit culture using data science methods that could help explain and explore those behaviors from leaders and nurses that promote positive cultures on patient units.

Discussion

Applications of data science have a profound impact on nursing practice because our ability to meaningfully use data are expanding. Once such area that is apparent in this review is the use of predictive modeling and forecasting. As nursing shortages persist,[256] a global pandemic introduces complexities to care, and as patient populations are aging with increasing rates of comorbidities,[257] the need to accurately target interventions, resources, and clinician time is at an all-time high. While the ability of data science to augment nursing practice is not yet fully realized, this review has helped to highlight some specific use in cases and has identified some areas for future development. As noted in the results, the most represented topics were in-hospital mortality, pain, and length of stay and the least commonly represented topics included unit culture, burnout, and AI/ML credibility and acceptance. Currently, it is not clear why some topics were more represented in the literature aside from the possibility of data availability. For instance, unit culture most likely did not have the quantity of data readily available to analyze in comparison to in-hospital mortality. However, as single year cannot determine the coverage of science, future iterations of this review should note year-to-year trends.

This rapid review highlights the important intersections between nursing practice and data science. We were able to examine both patient-centered topics (such as pain) and clinician-centered topics (such as burnout), calling attention to the multifaceted approaches in which data science can support and study nursing practice. In this review, we opted to include studies that were conducted by nonnurse authors and studies that were conducted on nonnurse populations. While nursing is a unique discipline defined by a diverse set of roles in a variety of health care settings, the inclusion of works that are relevant to nursing practice but studied among nonnursing populations, provides useful information. Foremost, this informs us of data science trends that are transforming health care, giving us insight into how nursing may change, or how nursing may support the latest practice recommendations. Additionally, this provides the informatics community insights into work that needs to be validated in nurse populations. We have identified several areas for future work in the discussion, including gaps in nurse-focused research. The necessary inclusion of these multidisciplinary studies among a nursing-focused review also raises the question of whether or not the informatics community is utilizing nurses and nursing data frequently enough, especially when considering data science methods. As nurses make up the largest portion of clinicians in health care, the volume of data and availability of patients is not a barrier. Continual assessments of nursing presence among studies that examine nurse-relevant topics are important to ensure representation of nursing in data science.

Finally, the year 2020 presented unique challenges, as the COVID-19 pandemic impacted research and hospital operations,[258] publication times, and research foci. For this

reason, we expected that a large amount of literature to focus on addressing the pandemic. Somewhat surprisingly, only 16 of the 244 unique articles included in this review addressed COVID-19 (albeit we did not specifically search for COVID-19-related articles). There was also significant attention on racial justice in 2020, so we examined whether any articles addressed racial bias. Only one study[235] included an emphasis on race-based health inequities. These low counts could be due to publishing times and should be examined in reviews of the literature for 2021.

Limitations

This review had three notable limitations. First, the rapid search strategy and screening process was designed to be nonexhaustive. While the purpose of this review was to provide a high-level overview of nursing-relevant literature using data science methods, it is possible that not all of the relevant literature was captured for each topic. Second, the abstract review and literature inclusion process was conducted by single authors (i.e., validation of inclusion was not done by a second reviewer). While this was necessary to help expedite this rapid review, the omission of validation by a second reviewer may have yielded different results and introduced bias. Third, not all topics related to nursing practice were included in this review. While this is not possible, every effort was made to include a representative sample of topics as determined by the literature and in-depth discussion among the nurse authors.

Conclusion

The intersection of nursing and data science provides new opportunities to improve health care by augmenting care processes. This rapid literature review has revealed several areas that have been widely studied in the past year, and some that could benefit from more research. In particular, effort should be made in improving the availability of nursing generated data in an interoperable form. It is in the best interest of the informatics community to monitor the most current trends in data science across different disciplines, as the latest methods are only helpful if they can be applied to real-world practice. Nursing is especially rife with opportunity, as it permeates inpatient, outpatient, and community settings, with nurses generating data at exponential rates.

Clinical Relevance Statement

An understanding of how data science methods influence research regarding nursing-relevant patient outcomes and clinical processes is important for nurses and the health informatics community. Data science is shaping how nursing can be practiced and how care can be delivered, as is evidenced by the literature highlighted in this review. Examining such literature is crucial to monitoring the uptake of research into real-world practice.

Multiple Choice Questions

1. Of all the studies included in this review, which topics had the most representation?
   a. In-hospital mortality, pain, and length of stay
   b. Unit culture, burnout, and AI/ML credibility and acceptance
   c. Health care cost and utilization, complex care, and falls
   d. Pressure injuries, until culture, and falls

   Correct Answer: The correct answer is option a. While the other topics had representative literature, in-hospital mortality, pain, and length of stay were most frequently represented. By understanding what topics are most frequently represented (and those that are less frequently represented) in the literature, we may make informed decisions regarding our approach to future research.

2. What method was used to identify the literature in this review?
   a. Natural language processing
   b. Deep learning model
   c. Rapid literature review
   d. Systematic literature review

   Correct Answer: The correct answer is option c. This review followed a rapid literature review protocol. We examined data science methods such as natural language processing and deep learning, but these methods were not used to conduct this review. A systematic review follows a more robust protocol, but we opted to report our findings following rapid review to expedite the dissemination of this 2020 review.

Protection of Human and Animal Subject
This research does not involve human subjects.

Conflict of Interest
None declared.

References

1   Johnson KE, McMorris BJ, Raynor LA, Monsen KA. What big size you have! Using effect sizes to determine the impact of public health nursing interventions. Appl Clin Inform 2013;4(03):434–444
2   Roosan D, Del Fiol G, Butler J, et al. Feasibility of population health analytics and data visualization for decision support in the infectious diseases domain: a pilot study. Appl Clin Inform 2016;7(02):604–623
3   Zhou Y, Zhao L, Zhou N, et al. Predictive big data analytics using the UK biobank data. Sci Rep 2019;9(01):6012
4   Center for Nursing Informatics. 2019 Nursing Knowledge: Big Data Science Conference. Accessed December 24, 2021 at:

https://nursing.umn.edu/centers/center-nursing-informatics/previous-conferences/2019-nursing-knowledge-big-data-science

5 Grant MJ, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. Health Info Libr J 2009;26(02):91–108

6 Jones TL. Outcome measurement in nursing: imperatives, ideals, history, and challenges. Online J Issues Nurs 2016;21(02):1

7 Heslop L, Lu S, Xu X. Nursing-sensitive indicators: a concept analysis. J Adv Nurs 2014;70(11):2469–2482

8 Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016;5(01):210

9 Choi K, Choi SJ, Kim ES. Computer-Aided Diagonosis for Colorectal Cancer using Deep Learning with Visual Explanations. Annu Int Conf IEEE Eng Med Biol Soc 2020;2020:1156–1159

10 Hatwell J, Gaber MM, Atif Azad RM. Ada-WHIPS: explaining AdaBoost classification with applications in the health sciences. BMC Med Inform Decis Mak 2020;20(01):250

11 Mensa E, Colla D, Dalmasso M, et al. Violence detection explanation via semantic roles embeddings. BMC Med Inform Decis Mak 2020;20(01):263

12 Hanson A, Haddad LM. Nursing Rights of Medication Administration. In: StatPearls. Treasure Island, FL: StatPearls; 2021

13 Malley A, Kenner C, Kim T, Blakeney B. The role of the nurse and the preoperative assessment in patient transitions. AORN J 2015; 102(02):181.e1–181.e9

14 Farid M, Purdy N, Neumann WP. Using system dynamics modelling to show the effect of nurse workload on nurses' health and quality of care. Ergonomics 2020;63(08):952–964

15 Silva E, Aguiar J, Reis LP, Sá JOE, Gonçalves J, Carvalho V. Stress among Portuguese medical students: the EuStress solution. J Med Syst 2020;44(02):45

16 Bodenheimer T, Sinsky C. From triple to quadruple aim: care of the patient requires care of the provider. Ann Fam Med 2014;12 (06):573–576

17 Jones A, Bronskill SE, Seow H, Junek M, Feeny D, Costa AP. Associations between continuity of primary and specialty physician care and use of hospital-based care among community-dwelling older adults with complex care needs. PLoS One 2020; 15(06):e0234205

18 Knox S, Downer B, Haas A, Middleton A, Ottenbacher KJ. Dementia severity associated with increased risk of potentially preventable readmissions during home health care. J Am Med Dir Assoc 2020;21(04):519–524.e3

19 Osakwe ZT, Sosina OA, Agu N, Fleur-Calixte RS. Home health agency factors associated with acute care hospitalization and emergency department use. Home Healthc Now 2020;38(02): 92–97

20 Chase JD, Russell D, Huang L, Hanlon A, O'Connor M, Bowles KH. Relationships between race/ethnicity and health care utilization among older post-acute home health care patients. J Appl Gerontol 2020;39(02):201–213

21 Courtright KR, Jordan L, Murtaugh CM, et al. Risk factors for long-term mortality and patterns of end-of-life care among medicare sepsis survivors discharged to home health care. JAMA Netw Open 2020;3(02):e200038–e200038

22 Griffith KN, Prentice JC, Mohr DC, Conlin PR. Predicting 5- and 10-year mortality risk in older adults with diabetes. Diabetes Care 2020;43(08):1724–1731

23 Manz CR, Chen J, Liu M, et al. Validation of a machine learning algorithm to predict 180-day mortality for outpatients with cancer. JAMA Oncol 2020;6(11):1723–1730

24 Schluter PJ, Askew DA, Jamieson HA, Arnold EP. Urinary and fecal incontinence are independently associated with falls risk among older women and men with complex needs: A national population study. Neurourol Urodyn 2020;39(03):945–953

25 Ahlström G, Axmon A, Sandberg M, Hultqvist J. Specialist psychiatric health care utilization among older people with

intellectual disability - predictors and comparisons with the general population: a national register study. BMC Psychiatry 2020;20(01):70

26 Schultebraucks K, Shalev AY, Michopoulos V, et al. A validated predictive algorithm of post-traumatic stress course following emergency department admission after a traumatic stressor. Nat Med 2020;26(07):1084–1088

27 Kuo YH, Chan NB, Leung JMY, et al. An integrated approach of machine learning and systems thinking for waiting time prediction in an emergency department. Int J Med Inform 2020;139:104143

28 Kogan E, Twyman K, Heap J, Milentijevic D, Lin JH, Alberts M. Assessing stroke severity using electronic health record data: a machine learning approach. BMC Med Inform Decis Mak 2020; 20(01):8

29 Tsiklidis EJ, Sims C, Sinno T, Diamond SL. Using the National Trauma Data Bank (NTDB) and machine learning to predict trauma patient mortality at admission. PLoS One 2020;15(11): e0242166

30 Schuler A, O'Súilleabháin L, Rinetti-Vargas G, et al. Assessment of value of neighborhood socioeconomic status in models that use electronic health record data to predict health care use rates and mortality. JAMA Netw Open 2020;3(10):e2017109

31 Plante TB, Blau AM, Berg AN, et al. Development and external validation of a machine learning tool to rule out covid-19 among adults in the emergency department using routine blood tests: a large, multicenter, real-world study. J Med Internet Res 2020;22 (12):e24048

32 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(02):153–160

33 Li J, Li BY, Wei ZJ, Zhao YZ, Li TS. Application research on gated recurrent unit deep learning prediction and graded early warning of emergency department visits based on meteorological environmental data. Biomed Environ Sci 2020;33(10):817–820

34 Jang SB, Lee SH, Lee DE, et al. Deep-learning algorithms for the interpretation of chest radiographs to aid in the triage of COVID-19 patients: A multicenter retrospective study. PLoS One 2020; 15(11):e0242759

35 Jang DH, Kim J, Jo YH, et al. Developing neural network models for early detection of cardiac arrest in emergency department. Am J Emerg Med 2020;38(01):43–49

36 Huang Y, Xu C, Ji M, Xiang W, He D. Medical service demand forecasting using a hybrid model based on ARIMA and self-adaptive filtering method. BMC Med Inform Decis Mak 2020;20(01):237

37 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(03):e0229331

38 Topaz M, Woo K, Ryvicker M, Zolnoori M, Cato K. Home healthcare clinical notes predict patient hospitalization and emergency department visits. Nurs Res 2020;69(06):448–454

39 Langer T, Favarato M, Giudici R, et al. Development of machine learning models to predict RT-PCR results for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in patients with influenza-like symptoms using only basic clinical data. Scand J Trauma Resusc Emerg Med 2020;28(01):113

40 Lee YS, Choi EJ, Kim YH, Park HA. Factors influencing falls in high- and low-risk patients in a tertiary hospital in Korea. J Patient Saf 2020;16(04):e376–e382

41 Lindberg DS, Prosperi M, Bjarnadottir RI, et al. Identification of important factors in an inpatient fall risk prediction model to improve the quality of care using EHR and electronic administrative data: a machine-learning approach. Int J Med Inform 2020;143:104272

42 Moskowitz G, Egorova NN, Hazan A, Freeman R, Reich DL, Leipzig RM. Using electronic health records to enhance predictions of fall risk in inpatient settings. Jt Comm J Qual Patient Saf 2020;46 (04):199–206

43 Stubbs B, Perara G, Koyanagi A, et al. Risk of hospitalized falls and hip fractures in 22,103 older adults receiving mental health care vs 161,603 controls: a large cohort study. J Am Med Dir Assoc 2020;21(12):1893–1899

44 Tago M, Katsuki NE, Oda Y, Nakatani E, Sugioka T, Yamashita SI. New predictive models for falls among inpatients using public ADL scale in Japan: A retrospective observational study of 7,858 patients in acute care setting. PLoS One 2020;15(07):e0236130

45 Womack JA, Murphy TE, Bathulapalli H, et al. Serious falls in middle-aged veterans: development and validation of a predictive risk model. J Am Geriatr Soc 2020;68(12):2847–2854

46 Ye C, Li J, Hao S, et al. Identification of elders at higher risk for fall with statewide electronic health records and a machine learning algorithm. Int J Med Inform 2020;137:104105

47 Zhao M, Li S, Xu Y, Su X, Jiang H. Developing a scoring model to predict the risk of injurious falls in elderly patients: a retrospective case-control study in multicenter acute hospitals. Clin Interv Aging 2020;15:1767–1778

48 Buisseret F, Catinus L, Grenard R, et al. Timed up and go and six-minute walking tests with wearable inertial sensor: one step further for the prediction of the risk of fall in elderly nursing home people. Sensors (Basel) 2020;20(11):E3207

49 Cella A, De Luca A, Squeri V, et al. Development and validation of a robotic multifactorial fall-risk predictive model: a one-year prospective study in community-dwelling older adults. PLoS One 2020;15(06):e0234904

50 Hsu YC, Weng HH, Kuo CY, Chu TP, Tsai YH. Prediction of fall events during admission using eXtreme gradient boosting: a comparative validation study. Sci Rep 2020;10(01):16777

51 Imoto D, Sawada K, Horii M, et al. Factors associated with falls in Japanese polio survivors. Disabil Rehabil 2020;42(13):1814–1818

52 Jung H, Park HA, Hwang H. Improving prediction of fall risk using electronic health record data with various types and sources at multiple times. Comput Inform Nurs 2020;38(03):157–164

53 Rehman RZU, Zhou Y, Del Din S, et al. Gait analysis with wearables can accurately classify fallers from non-fallers: a step toward better management of neurological disorders. Sensors (Basel) 2020;20(23):E6992

54 Suzuki M, Yamamoto R, Ishiguro Y, Sasaki H, Kotaki H. Deep learning prediction of falls among nursing home residents with Alzheimer's disease. Geriatr Gerontol Int 2020;20(06):589–594

55 Casilari E, Lora-Rivera R, García-Lagos F. A study on the application of convolutional neural networks to fall detection evaluated with multiple public datasets. Sensors (Basel) 2020;20(05):E1466

56 Charles A, Buckinx F, Locquet M, et al. Prediction of adverse outcomes in nursing home residents according to intrinsic capacity proposed by the World Health Organization. J Gerontol A Biol Sci Med Sci 2020;75(08):1594–1599

57 Chen PL, Lin HY, Ong JR, Ma HP. Development of a fall-risk assessment profile for community-dwelling older adults by using the National Health Interview Survey in Taiwan. BMC Public Health 2020;20(01):234

58 Lee SK, Ahn J, Shin JH, Lee JY. Application of machine learning methods in nursing home research. Int J Environ Res Public Health 2020;17(17):E6234

59 Magnuszewski L, Swietek M, Kasiukiewicz A, Kuprjanowicz B, Baczek J, Beata Wojszel Z. Health, functional and nutritional determinants of falls experienced in the previous year-a cross-sectional study in a geriatric ward. Int J Environ Res Public Health 2020;17(13):E4768

60 Umegaki H, Makino T, Uemura K, et al. Falls in community-dwelling prefrail older adults. Health Soc Care Community 2020;28(01):110–115

61 Martinez M, De Leon PL. Falls risk classification of older adults using deep neural networks and transfer learning. IEEE J Biomed Health Inform 2020;24(01):144–150

62 Zhao J, Liang G, Huang H, et al. Identification of risk factors for falls in postmenopausal women: a systematic review and meta-analysis. Osteoporos Int 2020;31(10):1895–1904

63 Manis DR, McArthur C, Costa AP. Associations with rates of falls among home care clients in Ontario, Canada: a population-based, cross-sectional study. BMC Geriatr 2020;20(01):80

64 Bucher BT, Shi J, Ferraro JP, et al. Portable automated surveillance of surgical site infections using natural language processing: development and validation. Ann Surg 2020;272(04):629–636

65 Chang B, Sun Z, Peiris P, Huang ES, Benrashid E, Dillavou ED. Deep learning-based risk model for best management of closed groin incisions after vascular surgery. J Surg Res 2020;254:408–416

66 Mancini A, Vito L, Marcelli E, et al. Machine learning models predicting multidrug resistant urinary tract infections using "DsaaS". BMC Bioinformatics 2020;21(Suppl 10):347

67 Mull HJ, Stolzmann KL, Shin MH, Kalver E, Schweizer ML, Branch-Elliman W. Novel method to flag cardiac implantable device infections by integrating text mining with structured data in the veterans health administration's electronic medical record. JAMA Netw Open 2020;3(09):e2012264–e2012264

68 Zachariah P, Sanabria E, Liu J, Cohen B, Yao D, Larson E. Novel strategies for predicting healthcare-associated infections at admission: implications for nursing care. Nurs Res 2020;69(05):399–403

69 McHaney-Lindstrom M, Hebert C, Miller H, Moffatt-Bruce S, Root E. Network analysis of intra-hospital transfers and hospital onset clostridium difficile infection. Health Info Libr J 2020;37(01):26–34

70 Meng L, Li J, He Y, et al. The risk factors analysis and establishment of an early warning model for healthcare-associated infections after pediatric cardiac surgery: A STROBE-compliant observational study. Medicine (Baltimore) 2020;99(49):e23324

71 Park JI, Bliss DZ, Chi CL, Delaney CW, Westra BL. Knowledge discovery with machine learning for hospital-acquired catheter-associated urinary tract infections. Comput Inform Nurs 2020;38(01):28–35

72 Cai Y, Zhang W, Zhang R, Cui X, Fang J. Combined use of three machine learning modeling methods to develop a ten-gene signature for the diagnosis of ventilator-associated pneumonia. Med Sci Monit 2020;26:e919035

73 Chen CY, Lin WC, Yang HY. Diagnosis of ventilator-associated pneumonia using electronic nose sensor array signals: solutions to improve the application of machine learning in respiratory research. Respir Res 2020;21(01):45

74 Poissy J, Damonti L, Bignon A, et al; FUNGINOS Allfun French Study Groups. Risk factors for candidemia: a prospective matched case-control study. Crit Care 2020;24(01):109

75 Ames CP, Smith JS, Gum JL, et al; European Spine Study Group and International Spine Study Group. Utilization of predictive modeling to determine episode of care costs and to accurately identify catastrophic cost nonwarranty outlier patients in adult spinal deformity surgery: a step toward bundled payments and risk sharing. Spine 2020;45(05):E252–E265

76 Coleman BC, Fodeh S, Lisi AJ, et al. Exploring supervised machine learning approaches to predicting Veterans Health Administration chiropractic service utilization. Chiropr Man Therap 2020;28(01):47

77 de Carvalho LSF, Gioppato S, Fernandez MD, et al. Machine learning improves the identification of individuals with higher morbidity and avoidable health costs after acute coronary syndromes. Value Health 2020;23(12):1570–1579

78 Hyer JM, Paredes AZ, White S, Ejaz A, Pawlik TM. Assessment of utilization efficiency using machine learning techniques: A study of heterogeneity in preoperative healthcare utilization among super-utilizers. Am J Surg 2020;220(03):714–720

79 Lerner J, Ruppenkamp J, Etter K, et al. Preoperative behavioral health, opioid, and antidepressant utilization and 2-year costs after spinal fusion-revelations from cluster analysis. Spine 2020;45(02):E90–E98

80 Luo L, Li J, Lian S, et al. Using machine learning approaches to predict high-cost chronic obstructive pulmonary disease patients in China. Health Informatics J 2020;26(03):1577–1598

81 Mazumdar M, Lin JJ, Zhang W, et al. Comparison of statistical and machine learning models for healthcare cost data: a simulation study motivated by Oncology Care Model (OCM) data. BMC Health Serv Res 2020;20(01):350

82 Mortani Barbosa EJ Jr., Kelly K. Statistical modeling can determine what factors are predictive of appropriate follow-up in patients presenting with incidental pulmonary nodules on CT. Eur J Radiol 2020;128:109062

83 Ng SHX, Rahman N, Ang IYH, et al. Characterising and predicting persistent high-cost utilisers in healthcare: a retrospective cohort study in Singapore. BMJ Open 2020;10(01):e031622

84 Wong ES, Schuttner L, Reddy A. Does machine learning improve prediction of VA primary care reliance? Am J Manag Care 2020; 26(01):40–44

85 Xie Y, Nguyen QD, Hamzah H, et al. Artificial intelligence for teleophthalmology-based diabetic retinopathy screening in a national programme: an economic analysis modelling study. Lancet Digit Health 2020;2(05):e240–e249

86 Chen S, Bergman D, Miller K, Kavanagh A, Frownfelter J, Showalter J. Using applied machine learning to predict healthcare utilization based on socioeconomic determinants of care. Am J Manag Care 2020;26(01):26–31

87 Chien I, Enrique A, Palacios J, et al. A machine learning approach to understanding patterns of engagement with internet-delivered mental health interventions. JAMA Netw Open 2020;3(07): e2010791

88 Zonta G, Malagù C, Gherardi S, et al. Clinical validation results of an innovative non-invasive device for colorectal cancer preventive screening through fecal exhalation analysis. Cancers (Basel) 2020;12(06):1471

89 Hong JC, Eclov NCW, Dalal NH, et al. System for high-intensity evaluation during radiation therapy (SHIELD-RT): a prospective randomized study of machine learning-directed clinical evaluations during radiation and chemoradiation. J Clin Oncol 2020;38 (31):3652–3661

90 Santilli AML, Jamzad A, Janssen NNY, et al. Perioperative margin detection in basal cell carcinoma using a deep learning framework: a feasibility study. Int J CARS 2020;15(05): 887–896

91 Rozenblum R, Rodriguez-Monguio R, Volk LA, et al. Using a machine learning system to identify and prevent medication prescribing errors: a clinical and cost analysis evaluation. Jt Comm J Qual Patient Saf 2020;46(01):3–10

92 Guntuku SC, Schwartz HA, Kashyap A, et al. Variability in language used on social media prior to hospital visits. Sci Rep 2020;10(01):4346

93 Chong LR, Tsai KT, Lee LL, Foo SG, Chang PC. Artificial intelligence predictive analytics in the management of outpatient mri appointment no-shows. AJR Am J Roentgenol 2020;215(05): 1155–1162

94 Akbarzadeh Khorshidi H, Hassani-Mahmooei B, Haffari G. An interpretable algorithm on post-injury health service utilization patterns to predict injury outcomes. J Occup Rehabil 2020;30 (03):331–342

95 Lauffenburger JC, Mahesri M, Choudhry NK. Not there yet: using data-driven methods to predict who becomes costly among low-cost patients with type 2 diabetes. BMC Endocr Disord 2020;20 (01):125

96 Wilson FA, Zallman L, Pagán JA, et al. Comparison of use of health care services and spending for unauthorized immigrants vs authorized immigrants or US citizens using a machine learning model. JAMA Netw Open 2020;3(12):e2029230

97 Benitez AE, Musinguzi N, Bangsberg DR, et al. Super learner analysis of real-time electronically monitored adherence to antiretroviral therapy under constrained optimization and comparison to non-differentiated care approaches for persons living with HIV in rural Uganda. J Int AIDS Soc 2020;23(03):e25467

98 Yu AT, Shakya R, Adhikari B, et al. A cluster-based, spatial-sampling method for assessing household healthcare utilization patterns in resource-limited settings. Clin Infect Dis 2020;71 (Suppl 3):S239–S247

99 Assaf D, Gutman Y, Neuman Y, et al. Utilization of machine-learning models to accurately predict the risk for critical COVID-19. Intern Emerg Med 2020;15(08):1435–1443

100 Brown AW, Lee M, Lennon RJ, Niewczyk PM. Functional performance and discharge setting predict outcomes 3 months after rehabilitation hospitalization for stroke. J Stroke Cerebrovasc Dis 2020;29(05):104746

101 Cholankeril G, Podboy A, Aivaliotis VI, et al. Association of digestive symptoms and hospitalization in patients with SARS-CoV-2 infection. Am J Gastroenterol 2020;115(07):1129–1132

102 Hao B, Sotudian S, Wang T, et al. Early prediction of level-of-care requirements in patients with COVID-19. eLife 2020;9:9

103 Jehi L, Ji X, Milinovich A, et al. Development and validation of a model for individualized prediction of hospitalization risk in 4,536 patients with COVID-19. PLoS One 2020;15(08):e0237419

104 Russell D, Luth EA, Ryvicker M, Bowles KH, Prigerson HG. Live discharge from hospice due to acute hospitalization: the role of neighborhood socioeconomic characteristics and race/ethnicity. Med Care 2020;58(04):320–328

105 Schrauben SJ, Chen HY, Lin E, et al; CRIC Study Investigators. Hospitalizations among adults with chronic kidney disease in the United States: a cohort study. PLoS Med 2020;17(12): e1003470

106 Steurer MA, Costello J, Baer RJ, et al. Cohort study of respiratory hospital admissions, air quality and sociodemographic factors in preterm infants born in California. Paediatr Perinat Epidemiol 2020;34(02):130–138

107 Temkin-Greener H, Wang S, Caprio T, Mukamel DB, Cai S. Obesity among nursing home residents: association with potentially avoidable hospitalizations. J Am Med Dir Assoc 2020;21(09): 1331–1335

108 Williams BA, Geba D, Cordova JM, Shetty SS. A risk prediction model for heart failure hospitalization in type 2 diabetes mellitus. Clin Cardiol 2020;43(03):275–283

109 Wu SS, Bellantoni M, Weiner JP. Geriatric syndrome risk factors among hospitalized postacute Medicare patients. Am J Manag Care 2020;26(10):e319–e326

110 Azar KMJ, Shen Z, Romanelli RJ, et al. Disparities in outcomes among COVID-19 patients in a large health care system in California. Health Aff (Millwood) 2020;39(07):1253–1262

111 Wang A, Kho AN, French DD. Association of the Robert Wood Johnson Foundations' social determinants of health and Medicare hospitalisations for ischaemic strokes: a cross-sectional data analysis. Open Heart 2020;7(01):e001189

112 Rough K, Dai AM, Zhang K, et al. Predicting inpatient medication orders from electronic health record data. Clin Pharmacol Ther 2020;108(01):145–154

113 Ioannou GN, Locke E, Green P, et al. Risk factors for hospitalization, mechanical ventilation, or death among 10 131 US veterans with SARS-CoV-2 infection. JAMA Netw Open 2020;3(09): e2022310

114 Phillips AZ, Rodriguez HPUS. U.S. county "food swamp" severity and hospitalization rates among adults with diabetes: A nonlinear relationship. Soc Sci Med 2020;249:112858

115 Colón-Emeric CS, Huang J, Pieper CF, Bettger JP, Roth DL, Sheehan OC. Cost trajectories as a measure of functional resilience after hospitalization in older adults. Aging Clin Exp Res 2020;32(12): 2595–2601

116 Zulman DM, Maciejewski ML, Grubber JM, et al. Patient-reported social and behavioral determinants of health and estimated risk of hospitalization in high-risk veterans affairs patients. JAMA Netw Open 2020;3(10):e2021457

117 Abujaber A, Fadlalla A, Gammoh D, Abdelrahman H, Mollazehi M, El-Menyar A. Prediction of in-hospital mortality in patients on mechanical ventilation post traumatic brain injury: machine learning approach. BMC Med Inform Decis Mak 2020;20(01):336

118 Abujaber A, Fadlalla A, Gammoh D, Abdelrahman H, Mollazehi M, El-Menyar A. Prediction of in-hospital mortality in patients with post traumatic brain injury using National Trauma Registry and Machine Learning Approach. Scand J Trauma Resusc Emerg Med 2020;28(01):44

119 Adjei Boakye E, Osazuwa-Peters N, Chen B, et al. Multilevel associations between patient- and hospital-level factors and in-hospital mortality among hospitalized patients with head and neck cancer. JAMA Otolaryngol Head Neck Surg 2020;146 (05):444–454

120 Allenbach Y, Saadoun D, Maalouf G, et al; DIMICOVID. Development of a multivariate prediction model of intensive care unit transfer or death: A French prospective cohort study of hospitalized COVID-19 patients. PLoS One 2020;15(10):e0240711

121 Ahmed N, Greenberg P. Early risk stratification of in hospital mortality following a ground level fall in geriatric patients with normal physiological parameters. Am J Emerg Med 2020;38(12): 2531–2535

122 Araujo GN, Silveira AD, Scolari FL, et al. Admission bedside lung ultrasound reclassifies mortality prediction in patients with ST-segment-elevation myocardial infarction. Circ Cardiovasc Imaging 2020;13(06):e010269

123 Altschul DJ, Unda SR, Benton J, et al. A novel severity score to predict inpatient mortality in COVID-19 patients. Sci Rep 2020; 10(01):16726

124 Awad A, Bader-El-Den M, McNicholas J, Briggs J, El-Sonbaty Y. Predicting hospital mortality for intensive care unit patients: Time-series analysis. Health Informatics J 2020;26(02): 1043–1059

125 Ayers B, Wood K, Gosev I, Prasad S. Predicting survival after extracorporeal membrane oxygenation by using machine learning. Ann Thorac Surg 2020;110(04):1193–1200

126 Bacchi S, Oakden-Rayner L, Menon DK, Jannes J, Kleinig T, Koblar S. Stroke prognostication for discharge planning with machine learning: a derivation study. J Clin Neurosci 2020;79:100–103

127 Baxter SL, Klie AR, Saseendrakumar BR, Ye GY, Hogarth M, Nemati S. Predicting mortality in critical care patients with fungemia using structured and unstructured data. Annu Int Conf IEEE Eng Med Biol Soc 2020;2020:5459–5463

128 Bertsimas D, Lukin G, Mingardi L, et al; Hellenic COVID-19 Study Group. COVID-19 mortality risk assessment: an international multi-center study. PLoS One 2020;15(12):e0243262

129 Burdick H, Pino E, Gabel-Comeau D, et al. Effect of a sepsis prediction algorithm on patient mortality, length of stay and readmission: a prospective multicentre clinical outcomes evaluation of real-world patient data from US hospitals. BMJ Health Care Inform 2020;27(01):e100109

130 Brajer N, Cozzi B, Gao M, et al. Prospective and external evaluation of a machine learning model to predict in-hospital mortality of adults at time of admission. JAMA Netw Open 2020;3(02): e1920733

131 Castro VM, McCoy TH, Perlis RH. Laboratory findings associated with severe illness and mortality among hospitalized individuals with coronavirus disease 2019 in Eastern Massachusetts. JAMA Netw Open 2020;3(10):e2023934

132 Cutti S, Klersy C, Favalli V, et al. A multidimensional approach of surgical mortality assessment and stratification (SMATT score). Sci Rep 2020;10(01):10964

133 Datta S, Loftus TJ, Ruppert MM, et al. Added value of intra-operative data for predicting postoperative complications: the MySurgeryRisk PostOp Extension. J Surg Res 2020;254:350–363

134 Du X, Min J, Shah CP, Bishnoi R, Hogan WR, Lemas DJ. Predicting in-hospital mortality of patients with febrile neutropenia using machine learning models. Int J Med Inform 2020;139:104140

135 Faisal M, Scally A, Howes R, Beatson K, Richardson D, Mohammed MA. A comparison of logistic regression models with alternative machine learning methods to predict the risk of in-hospital mortality in emergency medical admissions via external validation. Health Informatics J 2020;26(01):34–44

136 Gao N, Qi X, Dang Y, et al. Establishment and validation of a risk model for prediction of in-hospital mortality in patients with acute ST-elevation myocardial infarction after primary PCI. BMC Cardiovasc Disord 2020;20(01):513

137 Haug N, Deischinger C, Gyimesi M, Kautzky-Willer A, Thurner S, Klimek P. High-risk multimorbidity patterns on the road to cardiovascular mortality. BMC Med 2020;18(01):44

138 Iriondo M, Thio M, Del Río R, Baucells BJ, Bosio M, Figueras-Aloy J. Prediction of mortality in very low birth weight neonates in Spain. PLoS One 2020;15(07):e0235794

139 Kang MW, Kim J, Kim DK, et al. Machine learning algorithm to predict mortality in patients undergoing continuous renal replacement therapy. Crit Care 2020;24(01):42

140 Knight SR, Ho A, Pius R, et al; ISARIC4C investigators. Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: development and validation of the 4C Mortality Score. BMJ 2020;370: m3339

141 Kerckhoffs MC, Brinkman S, de Keizer N, et al. The performance of acute versus antecedent patient characteristics for 1-year mortality prediction during intensive care unit admission: a national cohort study. Crit Care 2020;24(01):330

142 Ko H, Chung H, Kang WS, et al. An artificial intelligence model to predict the mortality of COVID-19 patients at hospital admission time using routine blood samples: development and validation of an ensemble model. J Med Internet Res 2020;22(12):e25442

143 Laosa O, Pedraza L, Álvarez-Bustos A, Carnicero JA, Rodriguez-Artalejo F, Rodriguez-Mañas L. Rapid assessment at hospital admission of mortality risk from COVID-19: the role of functional status. J Am Med Dir Assoc 2020;21(12):1798–1802

144 Kong G, Lin K, Hu Y. Using machine learning methods to predict in-hospital mortality of sepsis patients in the ICU. BMC Med Inform Decis Mak 2020;20(01):251

145 Lei Lei E, Ladha K, Mueller B, et al. Noncardiac determinants of death and intensive care morbidity in adult congenital heart disease surgery. J Thorac Cardiovasc Surg 2020;159(06): 2407–2415

146 Li J, Chen Y, Chen S, et al. Derivation and validation of a prognostic model for predicting in-hospital mortality in patients admitted with COVID-19 in Wuhan, China: the PLANS (platelet lymphocyte age neutrophil sex) model. BMC Infect Dis 2020;20 (01):959

147 Li C, Ren Q, Wang Z, Wang G. Early prediction of in-hospital mortality in acute pancreatitis: a retrospective observational cohort study based on a large multicentre critical care database. BMJ Open 2020;10(12):e041893

148 Lindenmeyer CC, Kim A, Sanghi V, et al. The EMALT score: an improved model for prediction of early mortality in liver transplant recipients. J Intensive Care Med 2020;35(08):781–788

149 Matsumoto K, Nohara Y, Soejima H, Yonehara T, Nakashima N, Kamouchi M. Stroke prognostic scores and data-driven prediction of clinical outcomes after acute ischemic stroke. Stroke 2020;51(05):1477–1483

150 Medvedev MM, Brotherton H, Gai A, et al. Development and validation of a simplified score to predict neonatal mortality risk among neonates weighing 2000 g or less (NMR-2000): an analysis using data from the UK and The Gambia. Lancet Child Adolesc Health 2020;4(04):299–311

151 Menezes RC, Ferreira IBB, Carmo TA, et al. Are prognostic tools losing accuracy? Development and performance of a novel age-calibrated severity scoring system for critically ill patients. PLoS One 2020;15(11):e0240793

152 Mohamadlou H, Panchavati S, Calvert J, et al. Multicenter validation of a machine-learning algorithm for 48-h all-cause mortality prediction. Health Informatics J 2020;26(03):1912–1925

153 Mohapatra A, Liang NL, Makaroun MS, Schermerhorn ML, Farber A, Eslami MH. Risk factors for mortality after endovascular repair for blunt thoracic aortic injury. J Vasc Surg 2020;71(03):768–773

154 Muttalib F, Clavel V, Yaeger LH, Shah V, Adhikari NKJ. Performance of pediatric mortality prediction models in low- and middle-income countries: a systematic review and meta-analysis. J Pediatr 2020;225:182–192.e2

155 Morris RS, Milia D, Glover J, et al. Predictors of elderly mortality after trauma: A novel outcome score. J Trauma Acute Care Surg 2020;88(03):416–424

156 Rac H, Gould AP, Bookstaver PB, Justo JA, Kohn J, Al-Hasan MN. Evaluation of early clinical failure criteria for gram-negative bloodstream infections. Clin Microbiol Infect 2020;26(01):73–77

157 Siddiqi TJ, Usman MS, Khan MS, et al. Systematic review and meta-analysis of current risk models in predicting short-term mortality after transcatheter aortic valve replacement. EuroIntervention 2020;15(17):1497–1505

158 Shin JH, Kunisawa S, Imanaka Y. New outcome-specific comorbidity scores excelled in predicting in-hospital mortality and healthcare charges in administrative databases. J Clin Epidemiol 2020;126:141–153

159 Sinha I, Aluthge DP, Chen ES, Sarkar IN, Ahn SH. Machine learning offers exciting potential for predicting postprocedural outcomes: a framework for developing random forest models in IR. J Vasc Interv Radiol 2020;31(06):1018–1024.e4

160 Soong JTY, Rolph G, Poots AJ, Bell D. Validating a methodology to measure frailty syndromes at hospital level utilising administrative data. Clin Med (Lond) 2020;20(02):183–188

161 St-Louis E, Hassamal R, Razek T, Baird R, Poenaru D, Hardcastle TC. Validation of the PRESTO score in injured children in a South-African quaternary trauma center. J Pediatr Surg 2020;55(07):1245–1248

162 Stolz E, Rásky É, Jagsch C. Frailty index predicts geriatric psychiatry inpatient mortality: a case-control study. Psychogeriatrics 2020;20(04):469–472

163 Sukmark T, Lumlertgul N, Praditpornsilpa K, Tungsanga K, Eiam-Ong S, Srisawat N. THAI-ICU score as a simplified severity score for critically ill patients in a resource limited setting: Result from SEA-AKI study group. J Crit Care 2020;55:56–63

164 Takada T, Hoogland J, Yano T, et al. Added value of inflammatory markers to vital signs to predict mortality in patients suspected of severe infection. Am J Emerg Med 2020;38(07):1389–1395

165 Thorsen-Meyer HC, Nielsen AB, Nielsen AP, et al. Dynamic and explainable machine learning prediction of mortality in patients in the intensive care unit: a retrospective study of high-frequency data in electronic patient records. Lancet Digit Health 2020;2(04):e179–e191

166 Tuxun M, Zhao Q, Xiang Y, et al. Predicting value of white cell count and total bilirubin on clinical outcomes in patients with ST-elevation myocardial infarction following percutaneous coronary intervention: a cohort study. BMJ Open 2020;10(02):e031227

167 Tsai KY, Hsieh KY, Ou SY, Chou FH, Chou YM. Comparison of elixhauser and charlson methods for discriminative performance in mortality risk in patients with schizophrenic disorders. Int J Environ Res Public Health 2020;17(07):E2450

168 Vaid A, Somani S, Russak AJ, et al. Machine learning to predict mortality and critical events in a cohort of patients with COVID-19 in New York City: model development and validation. J Med Internet Res 2020;22(11):e24018

169 Wang K, Zuo P, Liu Y, et al. Clinical and laboratory predictors of in-hospital mortality in patients with coronavirus disease-2019: a cohort study in Wuhan, China. Clin Infect Dis 2020;71(16):2079–2088

170 Wong DJN, Harris S, Sahni A, et al; SNAP-2: EPICCS collaborators. Developing and validating subjective and objective risk-assessment measures for predicting mortality after major surgery: an international prospective cohort study. PLoS Med 2020;17(10):e1003253

171 Yu R, Zheng Y, Zhang R, Jiang Y, Poon CCY. Using a multi-task recurrent neural network with attention mechanisms to predict hospital mortality of patients. IEEE J Biomed Health Inform 2020;24(02):486–492

172 Zelis N, Buijs J, de Leeuw PW, van Kuijk SMJ, Stassen PM. A new simplified model for predicting 30-day mortality in older medical emergency department patients: The rise up score. Eur J Intern Med 2020;77:36–43

173 Zhang D, Yin C, Zeng J, Yuan X, Zhang P. Combining structured and unstructured data for predictive models: a deep learning approach. BMC Med Inform Decis Mak 2020;20(01):280

174 Diktas H, Uysal S, Erdem H, et al. A novel id-iri score: development and internal validation of the multivariable community acquired sepsis clinical risk prediction model. Eur J Clin Microbiol Infect Dis 2020;39(04):689–701

175 Rahmatinejad Z, Tohidinezhad F, Reihani H, et al. Prognostic utilization of models based on the APACHE II, APACHE IV, and SAPS II scores for predicting in-hospital mortality in emergency department. Am J Emerg Med 2020;38(09):1841–1846

176 Castiñeira D, Schlosser KR, Geva A, et al. Adding continuous vital sign information to static clinical data improves the prediction of length of stay after intubation: a data-driven machine learning approach. Respir Care 2020;65(09):1367–1377

177 Biron DR, Sinha I, Kleiner JE, et al. A novel machine learning model developed to assist in patient selection for outpatient total shoulder arthroplasty. J Am Acad Orthop Surg 2020;28(13):e580–e585

178 Sheikhalishahi S, Balaraman V, Osmani V. Benchmarking machine learning models on multi-centre eICU critical care dataset. PLoS One 2020;15(07):e0235424

179 Schulz EB, Phillips F, Waterbright S. Case-mix adjusted postanaesthesia care unit length of stay and business intelligence dashboards for feedback to anaesthetists. Br J Anaesth 2020;125(06):1079–1087

180 Zhang X, Qiu H, Liu S, Li J, Zhou M. Prediction of prolonged length of stay for stroke patients on admission for inpatient rehabilitation based on the International Classification of Functioning, Disability, and Health (ICF) generic set: a study from 50 centers in China. Med Sci Monit 2020;26:e918811

181 Jalali A, Lonsdale H, Do N, et al. Deep learning for improved risk prediction in surgical outcomes. Sci Rep 2020;10(01):9289

182 Anis HK, Strnad GJ, Klika AK, et al; Cleveland Clinic OME Arthroplasty Group. Developing a personalized outcome prediction tool for knee arthroplasty. Bone Joint J 2020;102-B(09):1183–1193

183 Goldstein BA, Cerullo M, Krishnamoorthy V, et al. Development and performance of a clinical decision support tool to inform resource utilization for elective operations. JAMA Netw Open 2020;3(11):e2023547

184 Chen CH, Hsieh JG, Cheng SL, Lin YL, Lin PH, Jeng JH. Early short-term prediction of emergency department length of stay using natural language processing for low-acuity outpatients. Am J Emerg Med 2020;38(11):2368–2373

185 Kirchebner J, Günther MP, Sonnweber M, King A, Lau S. Factors and predictors of length of stay in offenders diagnosed with schizophrenia - a machine-learning-based approach. BMC Psychiatry 2020;20(01):201

186 Gholi Zadeh Kharrat F, Shydeo Brandão Miyoshi N, Cobre J, Mazzoncini De Azevedo-Marques J, Mazzoncini de Azevedo-Marques P, Cláudio Botazzo Delbem A. Feature sensitivity criterion-based sampling strategy from the optimization based on phylogram analysis (Fs-OPA) and Cox regression applied to mental disorder datasets. PLoS One 2020;15(07):e0235147

187 Ma X, Si Y, Wang Z, Wang Y. Length of stay prediction for ICU patients using individualized single classification algorithm. Comput Methods Programs Biomed 2020;186:105224

188 Curiati PK, Gil-Junior LA, Morinaga CV, Ganem F, Curiati JAE, Avelino-Silva TJ. Predicting hospital admission and prolonged length of stay in older adults in the emergency department: the PRO-AGE scoring system. Ann Emerg Med 2020;76(03):255–265

189 Alsinglawi B, Alnajjar F, Mubin O, et al. Predicting length of stay for cardiovascular hospitalizations in the intensive care unit: machine learning approach. Annu Int Conf IEEE Eng Med Biol Soc 2020;2020:5442–5445

190 Wilk M, Marsh DWR, De Freitas S, Prowle J. Predicting length of stay in hospital using electronic records available at the time of admission. Stud Health Technol Inform 2020;270:377–381

191 Doctoroff L, Herzig SJ. Predicting patients at risk for prolonged hospital stays. Med Care 2020;58(09):778–784

192 Wolff J, Gary A, Jung D, et al. Predicting patient outcomes in psychiatric hospitals with routine data: a machine learning approach. BMC Med Inform Decis Mak 2020;20(01):21

193 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(06):989–995

194 Rahman MA, Honan B, Glanville T, Hough P, Walker K. Using data mining to predict emergency department length of stay greater than 4 hours: Derivation and single-site validation of a decision tree algorithm. Emerg Med Australas 2020;32(03):416–421

195 Sun LY, Bader Eddeen A, Ruel M, MacPhee E, Mesana TG. Derivation and validation of a clinical model to predict intensive care unit length of stay after cardiac surgery. J Am Heart Assoc 2020;9(21):e017847

196 McIsaac DI, Harris EP, Hladkowicz E, et al. Prospective comparison of preoperative predictive performance between 3 leading frailty instruments. Anesth Analg 2020;131(01):263–272

197 Barroso J, Wakaizumi K, Reckziegel D, et al. Prognostics for pain in osteoarthritis: Do clinical measures predict pain after total joint replacement? PLoS One 2020;15(01):e0222370

198 Cheng R, Alexandridi NA, Smith RM, et al. Fully automated patellofemoral MRI segmentation using holistically nested networks: Implications for evaluating patellofemoral osteoarthritis, pain, injury, pathology, and adolescent development. Magn Reson Med 2020;83(01):139–153

199 de Heer EW, Palacios JE, Adèr HJ, van Marwijk HWJ, Tylee A, van der Feltz-Cornelis CM. Chest pain, depression and anxiety in coronary heart disease: Consequence or cause? A prospective clinical study in primary care. J Psychosom Res 2020;129:109891

200 Fritz RL, Wilson M, Dermody G, Schmitter-Edgecombe M, Cook DJ. Automated smart home assessment to support pain management: multiple methods analysis. J Med Internet Res 2020;22(11):e23943

201 Hott A, Brox JI, Pripp AH, Juel NG, Liavaag S. Predictors of pain, function, and change in patellofemoral pain. Am J Sports Med 2020;48(02):351–358

202 Juwara L, Arora N, Gornitsky M, Saha-Chaudhuri P, Velly AM. Identifying predictive factors for neuropathic pain after breast cancer surgery using machine learning. Int J Med Inform 2020;141:104170

203 Kwon SB, Ku Y, Han HU, Lee MC, Kim HC, Ro DH. A machine learning-based diagnostic model associated with knee osteoarthritis severity. Sci Rep 2020;10(01):15743

204 Mieronkoski R, Syrjälä E, Jiang M, et al. Developing a pain intensity prediction model using facial expression: a feasibility study with electromyography. PLoS One 2020;15(07):e0235545

205 Parga JJ, Lewin S, Lewis J, et al. Defining and distinguishing infant behavioral states using acoustic cry analysis: is colic painful? Pediatr Res 2020;87(03):576–580

206 Pérez-Benito FJ, Conejero JA, Sáez C, et al. Subgrouping factors influencing migraine intensity in women: a semi-automatic methodology based on machine learning and information geometry. Pain Pract 2020;20(03):297–309

207 Russo MA, Georgius P, Pires AS, et al. Novel immune biomarkers in complex regional pain syndrome. J Neuroimmunol 2020;347:577330

208 Sörös P, Bantel C. Chronic noncancer pain is not associated with accelerated brain aging as assessed by structural magnetic resonance imaging in patients treated in specialized outpatient clinics. Pain 2020;161(03):641–650

209 Tanaka R, Hirohama K, Kurashige Y, et al. Prediction models considering psychological factors to identify pain relief in conservative treatment of people with knee osteoarthritis: a multi-center, prospective cohort study. J Orthop Sci 2020;25(04):618–626

210 Thiam P, Kestler HA, Schwenker F. Two-stream attention network for pain recognition from video sequences. Sensors (Basel) 2020;20(03):E839

211 Deffland M, Spies C, Weiss B, et al. Effects of pain, sedation and delirium monitoring on clinical and economic outcome: a retrospective study. PLoS One 2020;15(09):e0234801

212 Mutubuki EN, Luitjens MA, Maas ET, et al. Predictive factors of high societal costs among chronic low back pain patients. Eur J Pain 2020;24(02):325–337

213 Wang J, Wei M, Zhang L, et al. An autoencoder-based approach to predict subjective pain perception from high-density evoked EEG potentials. Annu Int Conf IEEE Eng Med Biol Soc 2020;2020:1507–1511

214 Wang R, Xu K, Feng H, Chen W. Hybrid RNN-ANN based deep physiological network for pain recognition. Annu Int Conf IEEE Eng Med Biol Soc 2020;2020:5584–5587

215 Xin X, Lin X, Yang S, Zheng X. Pain intensity estimation based on a spatial transformation and attention CNN. PLoS One 2020;15(08):e0232412

216 Kong Y, Posada-Quintero HF, Chon KH. Pain detection using a smartphone in real time. Annu Int Conf IEEE Eng Med Biol Soc 2020;2020:4526–4529

217 Brown TT, Lee W. The FUTUREPAIN study: validating a questionnaire to predict the probability of having chronic pain 7-10 years into the future. PLoS One 2020;15(08):e0237508

218 Dewitt B, Jalal H, Hanmer J. Computing PROPr utility scores for PROMIS profile instruments. Value Health 2020;23(03):370–378

219 Diaz Maffini M, Aguirre Ojea F, Manzotti M. Automatic detection of vital signs in clinical notes of the outpatient settings. Stud Health Technol Inform 2020;270:1211–1212

220 Liew BXW, Rugamer D, Stocker A, De Nunzio AM. Classifying neck pain status using scalar and functional biomechanical variables - development of a method using functional data boosting. Gait Posture 2020;76:146–150

221 Petersen T, Juhl CB, Fournier GL. Patients with persistent low back pain and nerve root involvement: to operate, or not to operate, that is the question. Spine 2020;45(07):483–490

222 Tu Y, Zeng F, Lan L, et al. An fMRI-based neural marker for migraine without aura. Neurology 2020;94(07):e741–e751

223 Alambo A, Andrew R, Gollarahalli S, et al. Measuring pain in sickle cell disease using clinical text. Annu Int Conf IEEE Eng Med Biol Soc 2020;2020:5838–5841

224 Liang C, Zhou S, Yao B, Hood D, Gong Y. Toward systems-centered analysis of patient safety events: Improving root cause analysis by optimized incident classification and information presentation. Int J Med Inform 2020;135:104054

225 Liu N, Chen CB, Kumara S. Semi-supervised learning algorithm for identifying high-priority drug-drug interactions through adverse event reports. IEEE J Biomed Health Inform 2020;24(01):57–68

226 Shiner B, Neily J, Mills PD, Watts BV. Identification of inpatient falls using automated review of text-based medical records. J Patient Saf 2020;16(03):e174–e178

227 Wong ZS-Y, So HY, Kwok BSC, Lai MWS, Sun DTF. Medication-rights detection using incident reports: a natural language processing and deep neural network approach. Health Informatics J 2020;26(03):1777–1794

228 Evans HP, Anastasiou A, Edwards A, et al. Automated classification of primary care patient safety incident report content and severity using supervised machine learning (ML) approaches. Health Informatics J 2020;26(04):3123–3139

229 Yang J, Wang L, Phadke NA, et al. Development and validation of a deep learning model for detection of allergic reactions using safety event reports across hospitals. JAMA Netw Open 2020;3 (11):e2022836

230 Baernholdt M, Yan G, Hinton ID, Cramer E, Dunton N. Effect of preventive care interventions on pressure ulcer rates in a national sample of rural and urban nursing units: Longitudinal associations over 4 years. Int J Nurs Stud 2020;105:103455

231 Brindle T. Incidence and variables predictive of pressure injuries in patients undergoing ventricular assist device and total artificial heart surgeries: an 8-year retrospective cohort study. Adv Skin Wound Care 2020;33(12):651–660

232 Cox J, Schallom M, Jung C. Identifying risk factors for pressure injury in adult critical care patients. Am J Crit Care 2020;29(03): 204–213

233 Hu YH, Lee YL, Kang MF, Lee PJ. Constructing inpatient pressure injury prediction models using machine learning techniques. Comput Inform Nurs 2020;38(08):415–423

234 Kenneweg KA, Welch MC, Welch PJ. A 9-year retrospective evaluation of 102 pressure ulcer reconstructions. J Wound Care 2020;29(09):48–56

235 Seibert J, Barch D, Bernacet A, et al. Examining social risk factors in a pressure ulcer quality measure for three post-acute care settings. Adv Skin Wound Care 2020;33(03):156–163

236 Sprigle S, McNair D, Sonenblum S. Pressure ulcer risk factors in persons with mobility-related disabilities. Adv Skin Wound Care 2020;33(03):146–154

237 Artico M, Piredda M, D'Angelo D, et al. Prevalence, incidence and associated factors of pressure injuries in hospices: a multicentre prospective longitudinal study. Int J Nurs Stud 2020;111:103760

238 Choi BK, Kim MS, Kim SH. Risk prediction models for the development of oral-mucosal pressure injuries in intubated patients in intensive care units: A prospective observational study. J Tissue Viability 2020;29(04):252–257

239 Osis SL, Diccini S. Incidence and risk factors associated with pressure injury in patients with traumatic brain injury. Int J Nurs Pract 2020;26(03):e12821

240 Matar G, Lina JM, Kaddoum G. Artificial neural network for in-bed posture classification using bed-sheet pressure sensors. IEEE J Biomed Health Inform 2020;24(01):101–110

241 Santosa A, Puspitasari N, Isnaini NUR. A path analysis study of factors influencing decubitus in a geriatric nursing home: a preliminary study. Family Med Primary Care Rev 2020;22(01): 67–70

242 Macron A, Pillet H, Doridam J, et al. Is a simplified Finite Element model of the gluteus region able to capture the mechanical response of the internal soft tissues under compression? Clin Biomech (Bristol, Avon) 2020;71:92–100

243 Brom H, Brooks Carthon JM, Ikeaba U, Chittams J. Leveraging electronic health records and machine learning to tailor nursing care for patients at high risk for readmissions. J Nurs Care Qual 2020;35(01):27–33

244 Jeong IC, Healy R, Bao B, et al. Assessment of patient ambulation profiles to predict hospital readmission, discharge location, and length of stay in a cardiac surgery progressive care unit. JAMA Netw Open 2020;3(03):e201074

245 Li CY, Karmarkar A, Kuo YF, et al. A comparison of three methods in categorizing functional status to predict hospital readmission across post-acute care. PLoS One 2020;15(05):e0232017

246 Shah N, Konchak C, Chertok D, et al. Clinical Analytics Prediction Engine (CAPE): development, electronic health record integration and prospective validation of hospital mortality, 180-day mortality and 30-day readmission risk prediction models. PLoS One 2020;15(08):e0238065

247 Zhang Y, Zhang Y, Sholle E, et al. Assessing the impact of social determinants of health on predictive models for potentially avoidable 30-day readmission or death. PLoS One 2020;15 (06):e0235064

248 Saleh SN, Makam AN, Halm EA, Nguyen OK. Can we predict early 7-day readmissions using a standard 30-day hospital readmission risk prediction model? BMC Med Inform Decis Mak 2020;20 (01):227

249 Hyer JM, White S, Cloyd J, et al. Can we improve prediction of adverse surgical outcomes? development of a surgical complexity score using a novel machine learning technique. J Am Coll Surg 2020;230(01):43–52.e1

250 Barbieri S, Kemp J, Perez-Concha O, et al. Benchmarking deep learning architectures for predicting readmission to the ICU and describing patients-at-risk. Sci Rep 2020;10(01):1111

251 Sanson G, Marino C, Valenti A, Lucangelo U, Berlot G. Is my patient ready for a safe transfer to a lower-intensity care setting? Nursing complexity as an independent predictor of adverse events risk after ICU discharge. Heart Lung 2020;49(04): 407–414

252 Sun CLF, Zuccarelli E, Zerhouni EGA, et al. Predicting coronavirus disease 2019 infection risk and related risk drivers in nursing homes: a machine learning approach. J Am Med Dir Assoc 2020; 21(11):1533–1538.e6

253 Rathlev NK, Visintainer P, Schmidt J, Hettler J, Albert V, Li H. Patient characteristics and clinical process predictors of patients leaving without being seen from the emergency department. West J Emerg Med 2020;21(05):1218–1226

254 Bradshaw S, Buenning B, Powell A, Teasley S, Olney A, Lee B. Retrospective chart review: readmission prediction ability of the high acuity readmission risk pediatric screen (HARRPS) tool. J Pediatr Nurs 2020;51:49–56

255 Burlison JD, Quillivan RR, Kath LM, et al. A multilevel analysis of U.S. hospital patient safety culture relationships with perceptions of voluntary event reporting. J Patient Saf 2020;16(03): 187–193

256 Haddad LM, Annamaraju P, Toney-Butler TJ. Nursing Shortage. In: StatPearls. Treasure Island, FL: StatPearls; 2021

257 Divo MJ, Martinez CH, Mannino DM. Ageing and the epidemiology of multimorbidity. Eur Respir J 2014;44(04):1055–1068

258 Grange ES, Neil EJ, Stoffel M, et al. Responding to COVID-19: the UW medicine information technology services experience. Appl Clin Inform 2020;11(02):265–275