2021

Data Science Methods for Nursing-Relevant Patient Outcomes and Clinical Processes The 2019 Literature Year in Review

Grace Gao

Data Science Methods for Nursing-Relevant Patient Outcomes and Clinical Processes

The 2019 Literature Year in Review

Mary Anne Schultz, BSN, MSN, MBA, PhD, Rachel Lane Walden, MLIS, Kenrick Cato, RN, PhD, CPHIMS, FAAN, Cynthia Peltier Coviak, PhD, RN, FNAP, Christopher Cruz, MSHI, RN-BC, CPHIMS, Fabio D'Agostino, PhD, MSN, RN, Brian J. Douthit, MSN, RN-BC, Thompson Forbes, PhD, RN, Grace Gao, PhD, DNP, RN-BC, Mikyoung Angela Lee, PhD, RN, Deborah Lekan, PhD, RN-BC, Ann Wieben, MS, BSN, RN-BC, Alvin D. Jeffery, PhD, RN-BC, CCRN-K, FNP-BC

Data science continues to be recognized and used within healthcare due to the increased availability of large data sets and advanced analytics. It can be challenging for nurse leaders to remain apprised of this rapidly changing landscape. In this article, we describe our findings from a scoping literature review of papers published in 2019 that use data science to explore, explain, and/or predict 15 phenomena of interest to nurses. Fourteen of the 15 phenomena were associated with at least one paper published in 2019. We identified the use of many contemporary data science methods (eg, natural language processing, neural networks) for many of the outcomes. We found many studies exploring *Readmissions* and *Pressure Injuries*. The topics of *Artificial Intelligence/Machine Learning Acceptance*, *Burnout*, *Patient Safety*, and *Unit Culture* were poorly represented. We hope that the studies described in this article help readers: (1) understand the breadth and depth of data science's ability to improve clinical processes and patient outcomes that are relevant to nurses and (2) identify gaps in the literature that are in need of exploration.

KEY WORDS: Artificial intelligence, Data analytics, Nursing research, Outcome and process assessment

Author Affiliations: California State University (Dr Schultz); Annette and Irwin Eskind Family Biomedical Library, Vanderbilt University (Ms Walden); Department of Emergency Medicine, Columbia University School of Nursing (Dr Cato); Grand Valley State University (Dr Coviak); Global Health Technology & Informatics, Chevron, San Ramon, CA (Mr Cruz); Saint Camillus International University of Health Sciences, Rome, Italy (Dr D'Agostino); Duke University School of Nursing (Mr Douthit); East Carolina University College of Nursing (Dr Forbes); St Catherine University Department of Nursing (Dr Gao); Texas Woman's University College of Nursing (Dr Lee); Assistant Professor, University of North Carolina at Greensboro School of Nursing (Dr Lekan); University of Wisconsin School of Nursing (Ms Wieben); and Vanderbilt University School of Nursing, and Tennessee Valley Healthcare System, US Department of Veterans Affairs (Dr Jeffery).

Dr Jeffery received support for this work from the Agency for Healthcare Research and Quality (AHRQ) and the Patient-Centered Outcomes Research Institute (PCORI) under Award Number K12 HS026395 as well as the resources and use of facilities at the Department of Veterans Affairs, Tennessee Valley Healthcare System. The content is solely the responsibility of the authors and does not necessarily represent the official views of AHRQ, PCORI, the Department of Veterans Affairs, or the US Government. The remaining authors have no conflicts of interest to declare.

Corresponding author: Alvin D. Jeffery, PhD, RN-BC, CCRN-K, FNP-BC, 461 21st Ave South, Nashville, TN 37240 (alvinjeffery@gmail.com).

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.cinjournal.com).

DOI: 10.1097/CIN.0000000000000705

The field of *data science* (inclusive of concepts such as *artificial intelligence* [AI], *predictive analytics*, and *machine learning*) is increasingly used not only in lay news and media but also in biomedical and nursing literature. There is hope that leveraging large data sets and advanced analytics is associated with improvements in clinical care delivery and patient outcomes. Unfortunately, the ever-expanding corpus of publications and the plethora of potential clinical applications can leave many nurse leaders struggling to remain apprised of the most contemporary methods being used in the literature. In this article, we describe a representative selection of papers published in 2019 that use data science to explore, explain, and/or predict phenomena of interest to nurses.

This project was based on interest from members of the Data Science Workgroup of the Nursing Knowledge: Big Data Science Conference[1] hosted annually by the University of Minnesota School of Nursing. Using a concept analysis paper[2] and group consensus, we identified 15 nursing-relevant patient outcomes and clinical process measures where data science techniques could lead to new insights or advance knowledge. The outcomes selected for review comprise (in alphabetical order) the following: AI/Machine Learning (ML) Acceptance, Burnout, Emergency Department (ED) Visits, Falls, Healthcare-Acquired Infections (HAIs), Healthcare Utilization and Costs, Hospitalization, In-Hospital Mortality, Length of Stay, Pain, Patient Safety, Pressure Injuries (PIs), Readmissions, Staffing/Scheduling/Workload, and Unit Culture.

METHODS

A scoping literature review was conducted using PubMed and CINAHL databases in December 2019 for English-language studies published during the past year. The species filter was also used to restrict to human studies. There was one main search strategy which used a combination of keywords and subject headings to find studies discussing the use of data science. The following terms were used to create that strategy: data science, data analytics, artificial intelligence, machine learning, 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. This main strategy was combined with an outcome-specific strategy for all 15 outcomes (see Supplemental Digital Content 1, http://links.lww.com/CIN/A81, which presents full search strategies). Each outcome was reviewed by an individual author who is an expert in the outcome reviewed. Abstract and full-text screening was done using the Raayan[3] Web application. Inclusion/exclusion criteria were developed via group consensus with the intention of providing a representative sample of data science publications rather than an exhaustive review of all publications. Overall, 8682 abstracts were screened, and 159 studies were included in this review (see Supplemental Digital Content 2, http://links.lww.com/CIN/A82, which breaks down inclusion/exclusion numbers by outcome). These studies each were analyzed to identify their aims, study designs, data sources, samples, settings, populations, operational definitions of outcomes, list of variables, and data science methods.

RESULTS

Artificial Intelligence/Machine Learning Acceptance

Key Findings

We identified relatively few topics of AI/ML acceptance or credibility by measuring different outcomes (n = 4).[4–7] The selected papers investigated acceptance,[4] satisfaction,[5] trust,[6] and use of AI.[7] Methodologically, two of the studies were quantitative,[4,7] one was qualitative,[6] and one was a mixed-method approach.[5] Finally, most of the selected research focused on specific AI-based products, such as a smartphone app,[6] self-driving cars,[6] and home assistants like Amazon's Echo.[7]

Participant sample sizes ranged from 76 to 724, with one study occurring in China,[4] South Korea,[6] United Arab Emirates, and United Kingdom.[7] The majority of studies used traditional statistical analytical methods (ie, structural equation modeling [SEM] and analysis of variance).

Discussion

Shin et al's study provided a fascinating examination of the thorny topics that arise in AI/ML-driven applications and products, such as fairness, accountability, and transparency (FAT). They point out that the user perspective, how users see, perceive, and feel, is paramount, with a need for "human-centered algorithms."[5] Notable gaps in the existing literature include how to consistently measure the FAT concepts during algorithm development and implementation. More work is needed to investigate feedback loops for AI/ML to increase FAT, like how Web-based systems like Amazon and Facebook use user beliefs as a system behavior component.[5]

Burnout

Key Findings

All burnout-related reports were cohort studies using survey-based data to predict some component of burnout. Three studies used logistic regression, and one study used SEM with path analysis. Sample sizes ranged from 228 to 49 158. Notenbomer et al[8] explored absenteeism as a component of burnout, where both number of days and length of absence were reported. Bosman et al[9] similarly explored risk of sick leave, predicting this outcome as a binary variable within a timeframe. Oliver et al[10] explored subjective well-being in staff who care for those with intellectual disabilities, using the Satisfaction With Life Scale as a measure of well-being. Dutra et al[11] explored burnout using the Maslach Burnout Inventory, with data collected from nurses and nursing technicians to formulate a predictive model.

Discussion

It is worth noting that a limited number of studies in 2019 discussed prediction around burnout and their variant approaches to measuring this phenomenon. Only one article approached burnout directly, using an established scale as the predictor. This variance might be due to the fact that "burnout" as a term is not clearly delineated and has many aspects that could be partially defined. Here, we included both caregiver and healthcare professional burnout, but conceivably, burnout outside of the healthcare space could be examined, as it is a factor that influences individual health.

Of interest, the data contained in these studies were all collected using surveys and questionnaires. Often, when considering data science methods, either real-time or historical data are collected using a standardized method, for example, the electronic health record (EHR) or wearables are used, and primary data collection is infrequent. It is possible that the data required to predict burnout are not readily available, leading to a lack of more advanced data science methodology. In this light, we should promote the regular collection of data on staff and caregiver well-being, as we would then be able to develop decision-support tools to aid in minimizing the acquisition and effects of burnout. This is especially important as we approach increasing pressures associated with staffing shortages and costs associated with job attrition.

Emergency Department Visits

Key Findings

We identified 17 publications related to the ED,[12–28] with all but 4 taking place in the adult ED setting.[13,18,21,22] The majority of studies used a retrospective observational design leveraging EHR data. Other commonly used data sources comprised of geospatial data. Sample sizes ranged from 268 to 1 721 298 patients within the ED. Most of the studies were single site (n = 12), with other used data sources including the Pediatric Emergency Care Applied Research Network, National Hospital and Ambulatory Medical Care Survey, the German government, and the Centers for Medicare and Medicaid Services.[13,15,22–24] Two studies used unstructured data (eg, triage notes and radiology reports).[26,27] Another study used modeling for electrocardiographic signal data.[17]

Most of the papers applied machine learning algorithms (eg, random forest, weighted decision trees, support vector machines [SVMs], gradient boosting, k-nearest neighbors, least absolute shrinkage and selection operator, neural networks, and SVM)[12–27] and naïve Bayes.[21] A number of the studies also use NLP to analyze free text notes.[26,27] In one study, social network analysis was used to understand relationships between the primary diagnoses of women who reported a history of gender-based violence during ED examination.[28]

The most studied primary outcomes of interest in the ED were readmission and disposition,[18,20,21,24,25,27] although there were a number of other endpoints, including traumatic brain injury,[13] sepsis, need for revascularization,[17] transfer to the intensive care unit or death,[22] subdural hematoma,[26] ST-elevation myocardial infarction,[14] falls risk,[16] overdose risk,[15] gender-based violence,[28] diagnosis prediction,[12,23] and number of ED visits.[19] For most of the studies interested in readmission, the criterion used was return to the hospital within 72 hours[18,21]; however, one study did investigate 2-, 7-, 14-, or 30-day readmission to ED.[20,24,25,27] Variables that served as explanatory variables included clinical and demographic information, with two exceptions that also used geospatial predictor variables.[20,21]

Discussion

While the goals of the ED publications address many relevant questions in the setting, including prediction of diagnosis, return to the ED, and correct discharge disposition, there was one glaring gap in the body of the literature. All of the studies reviewed were research studies using retrospective data for model building to support future decision support interventions. None of the studies involved the evaluation of machine learning–based implementation. Without rigorous implementation evaluation, much of the work described in these publications will stay fixed in the academic realm, never being applied to clinical settings. Hopefully, this gap in the literature will be addressed in the near future.

Of note is the use of nursing-generated data in data science analyses, especially NLP methods applied to nursing documentation. A number of the studies highlighted the importance of nursing-collected information by exclusively or mostly using nursing triage data for prediction.[12,18,22,24,25,27]

Falls

Key Findings

For the outcome of fall prevention, relatively few publications were located that reported the use of data science methods. In all but one investigation a retrospective design was used. In the remaining study,[29] the researchers used a 10-month observational case control design that matched records of patients who fell to individuals who did not fall and whose hospital stays overlapped the stays of those patients. Electronic health record data from hospitals, subacute, and home healthcare were used in five studies,[16,29–32] while one used data from a rehabilitation database,[32] and one used data from a patient safety organization and public dataset released by Yelp.[33] Sample sizes ranged from 814 to 90 441 patients.

Some researchers used or attempted to ascertain the most relevant items of subscales of standardized fall risk scales (eg, the Morse Fall Scale,[29] Falls Risk Assessment Scoring System,[32] and the Missouri Alliance for Home Care fall risk assessment[30]) to predict fall episodes. However, one research team[33] used Agency for Healthcare Research and Quality rubrics that can improve the quality of actual reports of fall incidents in real time. Overall, four studies used logistic and multiple regression methods,[16,29,31,32] three used machine learning methods that included examination of random forests,[16,30,33] and one employed neural networks[33] for their analytic strategies.

Discussion

The sample of studies obtained for this review revealed that most of the information entered in data science analyses for falls prediction was culled from EHRs in acute care settings. An emphasis on safety of patients while hospitalized is of great importance. However, further attention to use of data science methods with data from client homes, community settings (eg, primary care), and rehabilitation settings could improve fall prevention efforts overall.

Healthcare-Acquired Infections

Key Findings

We identified relatively few publications for predicting HAIs. The majority of studies were retrospective designs leveraging EHR data.[34–39] One study used a prospective case control design analyzing breath gas data.[40] Sample sizes ranged from 24 to 124 068 patients. Most studies were of hospitalized, adult patients in a variety of countries including the United States,[34,38,39] Korea,[35] Norway,[37] and Taiwan.[40]

One study focused on community-based nursing home residents in the United States.[36]

Each study aimed to predict different hospital-acquired infection outcome variables: unit-level *Clostridium difficile* infection,[34] catheter-associated urinary tract infection,[35] methicillin-resistant *Staphylococcus aureus* transmission,[36] surgical site infections,[37] ventilator-associated pneumonia,[40] drug-resistant gram-negative infections,[38] and septic shock.[39] Variables that served as predictors most commonly included clinical data,[34–40] demographic data,[34–36,38,39] and administrative data,[34,35,37–39] while one study used geospatial data.[38] One study used unstructured data (searched notes for word "infection") to differentiate cases and controls but not to identify explanatory features.[37] The majority of studies used a variety of regression techniques.[34–39] Network analysis,[34] XGBoost decision trees,[37] Bayesian networks,[39] as well as neural networks and SVMs[40] were additional machine learning methods utilized to predict HAIs.

Discussion

Researchers are leveraging novel explanatory variables in models to predict HAIs, such as patient transfers[34] and gas sensor data.[40] Notable gaps in the literature reviewed include the lack of inclusion of pediatric patients and the underutilization of unstructured data. Models that predict the risk of HAIs among adults may not be effective in predicting risk of HAIs for pediatric patients. There may also be untapped opportunities to utilize unstructured data, such as nursing notes, for more timely and effective prediction of risk of HAIs in order to develop early warning systems in healthcare settings.

Healthcare Utilization and Costs

Key Findings

All studies explored different facets of costs and utilization, using several methods, data sources, and variables to do so. Each study aimed to evaluate varying outcomes related to cost and utilization, including medication adherence, cost of services due to modifiable behaviors, cost prediction based on diagnosis, early detection and prevention of deleterious events, and patient ability to acquire services. A majority of studies used cohort or retrospective designs, but some used prospective designs such as feasibility and exploratory (quasi-experimental) designs.[41–47] Sample sizes ranged from only 13 to greater than 1 million. Many data sources included in these studies consisted of administrative data, which is often the primary source used to determine service cost. However, other studies collected data from patient surveys,[46,48] EHR data (clinical observations, clinical notes, and imaging),[42,43,46,48–54] and smartphone location data (geospatial and audio data).[44,51,55]

Samples vary for each study depending on their purpose. Many studies sample patients with specific diagnoses to compare cost utilization (both as secondary EHR data and as prospective data from participants). Other studies included clinical notes,[45,47,56,57] forum posts,[58] video/audio recordings,[42,44] and medical images[42,43] as their corpus of data. Study settings largely included patients and data from academic medical centers but also included community health clinics,[54] health Web sites,[58] and state Medicaid databases.[49] Outcomes and associated variables differed but often included cost of a service or clinical event (ie, readmission). Other outcomes included medication adherence,[49,53,58] fall prevention,[42] 30-day morbidity,[51,52,59] service utilization,[55] and risk prediction.[41,48,49,52,59] Variables to assess these outcomes included patient demographics,[48–50,52,54] use of medications,[41,53] clinical diagnoses,[45,46,50,53,54,56,59] distance from health services,[45] lab results,[54] modifiable risk factors,[46,50,51] images,[42,43] self-reported outcomes,[46,48] sound detection,[44] and various flags in unstructured data.[45,47,56–58]

Studies used various data science methods to evaluate their outcomes. Several used some form of regression,[48,49,51,52,57,59] but the use of more advanced methods was prevalent as well. Of the 20 articles that addressed cost and utilization in this review, 5 used NLP,[45,47,56–58] 5 used neural networks,[41,43,44,53,54] and 6 used other types of machine learning.[42,46,48,50,54,55] Of note, nontraditional data source examination was reported: video data with AI,[42] deep learning with images,[43] and deep learning with audio data.[44]

Discussion

The subject of "healthcare costs and utilization" covers a wide variety of topics, and this is clearly reflected in the sample of papers included in this review. It is promising to see that several nontraditional data types (audio, image, text, geospatial, and video)[42–44,51,55] are being used to the benefit of patient outcomes, reducing costs and increasing healthcare access beyond that of which traditional data are capable. A mix of direct and indirect economic-based outcomes was noted as well, including the use of deep learning–based image analysis to increase the diagnostic quality of lower-dose positron emission tomography images,[43] both reducing costs and advocating for patient safety. Of the articles reviewed, most were done to develop a prediction model or estimate a cost. Future studies should focus on the implementation of such models in real-world practice, possibly through feasibility studies or pragmatic clinical trials.

Hospitalization

Key Findings

We identified 10 publications for predicting/describing/exploring hospitalization-related outcomes. Data sources generally originated from existing administrative, commercial claim, and hospital data. Retrospective studies were a commonly adopted study design. Predictive and associative modeling dominates the data science methods employed in these studies. Several data modeling methods include risk prediction algorithm development,[60] linear regression,[61] multivariate statistical analysis using SEM,[62] multivariable logistic regression,[60,61,63–66] negative binomial-logit hurdle regression,[67] geospatial analytic methods,[55] and a network analysis approach.[34] Studies included sample sizes ranging from 4822 to more than 1 million and comprised a variety of ages, genders, and disease conditions. Financial impacts and implications appeared to be a common interest of study.

Discussion

The relative number and variety of data science methods to build predictive associations and relationships among different factors and variables pertaining to hospitalization are notable. The research in this space is showing promising results in mining predictive factors and associations to improve disease prevention and management, health promotion, and detecting gaps in geographical regions that relate to the impacts associated with hospitalization. One notable trend of employing geospatial analytic methods to detect gaps in geographical regions pertaining to hospitalization-associated impact points to a great potential with strong implications for future data science research.

In-Hospital Mortality

Key Findings

A number of predictive models exist for identifying patients at high risk for dying in the hospital. The vast majority of studies were retrospective cohort studies that leveraged EHR data. A few studies[68–72] were prospective cohorts, and one relevant article was a systematic review and meta-analysis.[73] Sample sizes ranged from 51 to 281 522. Study populations included hospitalized adults from the following countries: Australia,[72] Brazil,[74] China,[71,75,76] Israel,[77] Ireland,[72] Italy,[70] Korea,[78,79] Singapore,[12] Spain,[68] Switzerland,[69] and the United States.[24,80–85]

Several studies focused on specific admission diagnoses or surgical procedures, which resulted in a trend toward better model performance compared to models including all-cause hospitalizations. Variables serving as predictors primarily comprised demographic information, vital signs, laboratory values, and diagnoses/comorbidities/procedures. Less commonly included but notable predictor variables comprised physical assessments,[72] physiological status scores,[68,74,77,81,84] and medication exposures.[77,82,84] One study included a nutrition score,[74] one study included census-tract-level socioeconomic status,[83] and one study included nursing diagnoses.[70] The majority of the works used regression (with or without additional methods) for making predictions.[24,68–76,78–86] The regression models primarily leveraged logistic regression; however, two papers applied Cox proportional hazards regression.[76,86] Ten papers noted the use of more contemporary methods for prediction: random forests,[12,24,78,81,82,87] gradient boosting,[12,24,77,81,82,86] naïve Bayes,[82] SVMs,[12,87] and neural networks.[24,73,78,87] Interestingly, one paper conducted a network analysis of healthcare providers and used the network characteristics to serve as predictors.[84] Another paper used regular expressions to extract features for a prediction model.[78]

Discussion

All papers were limited to adult populations. There might be a need for pediatric-focused in-hospital mortality prediction models. From a nursing perspective, it was beneficial to see one paper include nursing diagnoses[70] and another paper include socioeconomic status.[83] These voids suggest promising areas for the nurse-investigator who possesses data science methods expertise or who works on the appropriately prepared interprofessional research team.

Length of Stay

Key Findings

We identified six studies for predicting the hospital length of stay[52,88–92] and one for describing the association with the hospital length of stay.[93] All studies used retrospective designs. Data sources comprised EHR, administrative data, and patient-reported data. Sample sizes ranged from 186 to 132 095 patients within hospital settings in different patient populations: (1) surgical patients undergoing orthopedic and neurosurgical operations[52,88,89,93]; (2) patients who underwent surgeries as first-case in a day[90]; (3) critical care patients[91]; and (4) children with psychiatric complaints.[92]

Variables that served as predictors mostly included demographic and administrative data,[52,89,90,92] while few studies included clinical data[91,92] or unstructured text data.[88] Different data science methods were applied to predict length of stay: three studies[90–92] used more than one method such as supervised[90,91] and unsupervised[90] machine learning techniques, neural network,[90,91] and linear regression,[92] while three studies[52,88,89] used only one method such as supervised machine learning[52,89] and a neural network.[88] Finally, in two studies, NLP was used to characterize variables to be used in a predictive model[92] and to study the association with length of stay,[93] respectively.

Discussion

Interestingly, a wide range of different data science methods were applied in the predictive models developed by the investigators. The most used methods were supervised machine learning techniques and neural networks. These methods have become popular in the healthcare field and are promising in demonstrating the ability to synthesize available data to predict hospital length of stay.

Unfortunately, only in one study[92] were nursing-generated data (specifically, clinical notes written by triage or bedside nurses) used to predict length of stay, even though several studies have shown the predictive power of nursing-generated data (ie, nursing diagnoses) on this outcome.[94] Nursing-generated data (eg, nursing diagnoses, nursing interventions) can complement routinely collected administrative data (eg, coded medical diagnoses), contributing to explaining the patient's complexity. Nursing-generated data should be easily accessed in EHRs for analysis since their use can be extremely useful in developing predictive modeling for hospital length of stay with the aim of improving the hospital management.

Pain

Key Findings

We identified many publications for predicting causes or contributors to pain.[95–113] Most studies were prospective designs leveraging EHR, registry, and survey data. Other commonly used study designs comprised of retrospective analysis using pain questionnaires and EHR data. Most of the studies were within the inpatient/hospital settings (ie, surgical, ED, and pediatrics), while a substantial portion were within the clinic/outpatient care, and a few were within community care,[99,103] dental,[112] and sports medicine[109,113] settings. The sample sizes varied greatly from 30 (mostly survey type) to 12 329 (EHR and registry database).

The specific areas explored by the studies varied from identifying the causative factors of pain to predicting the quality of life (given the severity of pain) to establishing the role of novel data types in detecting pain such as facial recognition algorithm, biomarkers, and physiological signals. Other applications focused on the surgical/postsurgical pain, specifically looking at the link between medications and postdischarge pain control. Most of the outcomes explored by the studies sought to identify the predictors of pain such as psychological/emotional well-being, physiologic factors, opioid use, amount of sleep, and duration of surgery. Some of the most common variables used as predictors of pain were medications use,[108] cognitive function,[99,111] pain perception,[109] gray matter volume,[97] patient age, brain imaging,[97,98,104] heart rate variability,[104] electromyography signal,[101] vital signs, chronic pain intensity, and sleep duration.[113] Most studies used multiple regression and artificial neural networks, while NLP, SVMs, and other machine learning models were also employed.

Discussion

There are some interesting applications of data science within pain assessment/management, including the exploration of brain imaging (resting-state blood oxygenation-level-dependent and arterial-spin labeling functional imaging[104]) and autonomic activity (heart rate variability) to predict pain intensity. One study covered the use of artificial neural networks to measure and predict pain intensity using multiple physiologic data such as facial surface electromyogram signal, Galvanic skin response, heart rate, and blood pressure.[101] Another study used machine learning to predict the variance of intensity of menstrual pain by analyzing the gray matter volume and using machine learning on magnetic resonance images of the brain.[97] These studies have developed a foundation for the use of complex data sources within data science methods for pain assessment and pain management.

Patient Safety—Additional Measures and Outcomes

Key Findings

We identified relatively few publications for exploring patient safety topics not elsewhere described in this paper. The majority of studies were retrospective designs using EHR data, narrative notes, and various surveys on hospital unit metrics. The primary outcomes were the identification and classification of falls and fall incident reports,[114–116] safety, and predicting perspectives of patient safety on the Hospital Survey on Patient Safety Culture.[117] Studies using data science techniques primarily used NLP,[33,114,118] neural networks,[33,119] and machine learning techniques such as random forests[33,115] and ranged in size from 252 to over 3000.

Discussion

It is interesting to note that only one patient safety study using data science techniques included nursing-generated data.[117] No other studies explicitly used nursing-generated data or were published in nursing journals. The limited exposure of nurses to data science techniques that investigate patient safety might be due to the lack of nurse researchers with expertise in patient safety and the use of data science techniques to create understanding.

Pressure Injuries

Key Findings

We identified eight publications for predicting PIs, including one systematic review article[120] and seven empirical studies.[121–127] Shi et al[120] identified 22 prognostic models for predicting PI risk published between 1996 and 2017; half of these models were developed using prospective longitudinal data, and the other half used retrospective data. Most of the models were built by logistic regression and Cox regression.

The majority of the seven empirical studies published in 2019 were retrospective designs leveraging clinical data from hospitals, except for Duvall et al's[126] feasibility pilot study. The sample sizes ranged from 10 to 2062 patients[121–126] and 396 images of PI.[127] Various data science methods have been used to detect or predict PIs, including logistic regression,[121,122] nonlinear regression,[124] a univariate Cox regression,[125] and three data mining algorithms (decision trees, neural networks, and SVMs).[121] The significant predictive factors comprised PI history,[121] without cancer,[121] excretion,[121] activity/mobility,[121] skin condition/circulation,[121] estimated surgery time,[122] serum albumin level,[122,125] multiple ulcers,[125] and presence of a single caregiver.[125]

Logistic regression analysis was also used to determine the utility of three different PI risk assessment scales (ie, the Spinal Cord Injury Pressure Ulcer Scale, Braden Scale, and Functional Independence Measure]) for identifying individuals at risk for developing PI.[123] Duvall et al[126] used a threshold-based detection algorithm and a K-nearest neighbor classification approach to investigate the feasibility of a sensor technology (ie, the E-scale system) for detecting and classifying movements in bed (ie, roll, turn in place, extremity movements, and assisted turn), which are relevant for PI risk assessment. Ohura et al[127] explored different architectures of the convolutional neural network (CNN) in image segmentation to detect and discriminate ulcer regions of PI during assessment via telemedicine.

Discussion

Data science methods facilitate the prediction, detection, and management of PIs via optimized assessments. Consideration of the best prognostic factors derived from the studies, such as blood albumin level, mobility, skin conditions, and single caregiver, can be used to develop and improve nutrition programs or home care nursing programs. Notably, nurses can improve their real-time monitoring of high-pressure areas in the bed and assessing PI risk with the use of sensor technologies (eg, E-scale system). Also, as explored by Ohura et al,[127] the use of CNN architectures could support the eHealth wound assessment system to significantly change the management of PIs or chronic wounds. For future research, it is recommended to include vital signs and nursing interventions in PI predictive modeling.[122]

Readmissions

Key Findings

We identified 37 publications for hospital readmission. A majority of papers used a retrospective observational study design and secondary data analysis. Six studies used a prospective study design,[128–133] and one study was a randomized clinical trial.[134] Sample sizes ranged from 42 to 452 277 patients in the hospital setting. Almost half of the studies (n = 17) were single center, while other commonly used data sources were multisite hospitals, health systems, or large datasets such as the National Readmission Dataset, Healthcare Cost and Utilization Project, Veterans Health Administration system, and the Centers for Medicare and Medicaid Services.

Most of the papers (n = 30) applied logistic regression in prediction models; of these, 22 papers used development and validation cohorts, boot-strapped internal validation, or cross-validation.[21,77,89,130,132–149] One paper applied logistic regression to structured data, whereas unstructured data were analyzed using term frequency-inverse document frequency statistic.[135] Another study applied principal components analysis, multiple correspondence analysis, and multiple factor analysis.[146] Eleven papers applied machine learning algorithms (eg, random forest, weighted decision trees, SVMs, gradient boosting, neural networks, decision curve analysis, and synthetic minority oversample technique),[77,89,130,135,137,144–146,150,151] and naïve Bayes.[21] In one nurse-authored paper by Kwon et al,[144] a case study was used to illustrate different statistical and ML risk models and hospital readmission outcomes of patients with diabetes mellitus.

Most papers defined readmission as an unplanned readmission to a hospital within 30 days of discharge,[21,77,89,128–132,134–147,150–158] although some studies used 90-day,[159,160] 180-day,[148] within 1 year[161] readmission rates; three or more readmissions over 1 year[133]; and "instantaneous hospital readmission risk over time."[149] Readmission was also defined by urgency[138,154] and etiology (eg, disease specific[133–135,145,146,148,161]). A provocative paper by Brittan et al[136] calculated three definitions of readmission with differing inclusion/exclusion criteria for index admissions and readmissions.

The populations studied were primarily adult, although, three papers focused on inpatients younger than 18 years[21,136,151] or special populations such as surgical[89,129,131,139,142–144,152,153,158–160] and trauma,[161] medical conditions such as heart failure,[132,133,135,145,146] heart failure or myocardial infarction,[148] chronic obstructive pulmonary disease,[128] diabetes,[144] human immunodeficiency virus,[134] falls,[161] antimicrobial therapy,[139] and skilled nursing facility discharge.[137,143]

The most common predictor variables were sociodemographic (eg, age, gender, and race/ethnicity), comorbidity (eg, medical diagnoses or comorbidity index), and hospital utilization (eg, length of stay, number of prior hospital admissions, and ED visits) data. Novel predictors used in risk models with relevance to nursing included physical function assessments,[128,129,131,132,137,153,154,157,158,160,161] symptoms,[128,129,133,137,148,160] psychosocial factors,[132–134,157,159,161] vital signs, pulse oximetry, and body mass index,[77,128,133,135,141,145, [146,151–153,156,158–161] and frailty.[128,129, 131,132,153,154,158,160,161] Other clinical predictors that are infrequently applied in prevailing risk models include laboratory and/or imaging tests[77,131–135,137,141,142,144,146–148,151,152,156,159] and medications.[77,133,134,140,141,144,147,151,155,157,159,161] One anomaly is the paper by McConachie et al[155] that included no patient demographic data. Nijhawan et al[134] incorporated novel sociobehavioral predictors such as health literacy, medication adherence, substance abuse, patient-provider relationship satisfaction, perceived health status, and housing and food insecurity. Only one study included unstructured data (eg, physician and social worker clinical notes) in a risk model.[135]

Discussion

Multiple papers demonstrated that risk factors such as older age, poor health, certain medical diagnoses, multimorbidity, frailty, and healthcare utilization confer high risk for hospital readmission. While these risk factors potentially improve the predictive ability of models, nurses can make important contributions to model development by filling data gaps with nursing-relevant data pertaining to patients' biopsychosocial health and function and social determinants of health.[144] Identifying and applying common data elements relevant to nursing across EHR systems in predictive models and including standard nursing terminology (eg, "International Classification of Nursing Practice" codes in the EHR as suggested by Kwon et al[144]) would capture some nuances that provide contextual information about patient health status and thereby improve the relevance and performance of the models. Notably, only one study used unstructured EHR text data; further research is needed in this area since unstructured data could be a rich source of information for novel readmission risk factors. Most papers focused on adult populations since adult readmission rates are considerably higher than pediatric readmission rates (17% vs 3%–5%, respectively)[21]; however, given implications for cost and consequences for health systems and children and their families, there is a need for more research in pediatric populations.

Staffing/Scheduling/Workload

Key Findings

We identified three publications[29,162,163] for staffing but none for scheduling or workload. Data science methods were reportedly used in these studies aimed at estimating the antecedents or consequences of nurse staffing. Two studies[29,162] reported predictive models formed from ML methods, and one study[163] was a report of NLP used to transform clinical notes into assessment forecasting. Each had retrospective designs using secondary data, two used a single tertiary care setting[29,162] (tertiary care, maternal care), and one[163] used two matched psychiatric settings. Diverse samples were used, with one study[162] using over 2500 maternal inpatients, another[29] using over 800 medical/surgical inpatients, and the remaining study[163] sampling the admission encounters of over 5000 psychiatric patients.

The Nadkarni[162] research group used a stepwise, iterative, object-oriented program written with workflow and treatment processes in mind in a sample of 343 patients with potentially life-threatening complications and 2285 uncomplicated mothers in a Tanzanian hospital. Aimed at providing decision makers with a tool to analyze the impact of resource limitations on maternal inpatient complications, key variables included treatment efficacy, severity distribution, number and frequency of nurse visits, nurse staffing at the shift level, deterioration rate, and maternal near-misses.

Similarly, the Lucero[29] group elucidated a data-driven and practice-based approach to identify factors associated with inpatient falls in a sample of 272 patients who fell and 542 who did not while hospitalized in 14 medical-surgical units of a Florida tertiary-care hospital. Manual, semiautomated, and automated procedures deploying theoretically or practice-derived risk factors yielded a meaningful and parsimonious set of predictors for this adverse event. Skill mix, rates of nurse certification, and nurse-educational levels were among the relevant staffing variables in this observational case-controlled study.

In the remaining Menger[163] study, NLP was used to transform clinical notes from the patient's EHR to develop and validate a multivariable prediction model for the assessment of inpatient violence risk. In this prognostic study, the authors used clinician notes from the admission encounters of over 5000 patients in one of two different psychiatric settings in the Netherlands. The model training and estimation of predictive validity were done in a nested cross-validation setup in which the outcome of interest—the manifestation of violent behavior within 4 weeks of admission—was successfully predicted from inpatient violence risk assessment derived from the documentation in this manner. Although a staffing variable was not explicitly or operationally stated, the availability of a nurse (or psychiatrist) to conduct the admission assessment is inferred in this initial encounter from which language within the nursing (and medical) domain is derived.

Discussion

In these studies, staffing variables were of two types: nurse hours relative to either all staffing or patient load as well as nurse characteristics such as education and certification. Further, studies of the impact nurse staffing may have on patient outcomes should include characteristics of the nurse that are known or hypothesized to have an impact, such as their education, training, and mentoring needs. From a systems perspective,

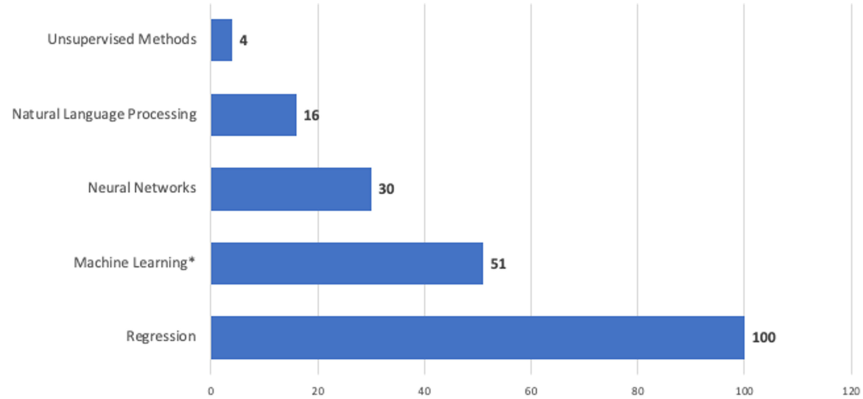

FIGURE 1. Frequency of use of data science methods among reviewed studies. *Note: While several methods on this figure could be considered "machine learning," this count does not include studies counted in a different category.

measures of the human capital resources, for example, nurse hours/patient day or skill mix, should be explicitly stated and for the relevant time partitions up to and including the time of injury, adversity, or other measurement.

Unit Culture

Of the 589 papers yielded in the initial literature search, none of the studies satisfied criteria for being included in the final analysis.

DISCUSSION

Through our literature review, we have identified and described a representative sample of publications focused on the use of data science methods relevant to nursing practice. All but one of the outcomes for which we searched were associated with at least one paper published in 2019. From a methodological perspective, in the reviewed studies, we noted the use of many contemporary data science methods (eg, NLP, neural networks, and social network analysis) and heterogenous predictors (see Figures 1 and 2).

We found a large number of studies exploring *Readmissions* and *PIs*. Observations of the contents of relevant journals suggest that risk prediction modeling for hospital readmission has increased in recent years due to the Affordable Care Act of 2010 and the subsequent Hospital Readmission Reduction Program, which has tied financial reimbursement penalties to potentially avoidable hospital readmissions. The high number of PI studies could be attributable to either of the following: (1) PI risk scores have existed for many years, so there is ample opportunity for including validated predictors within new analysis frameworks; and/or (2) PIs are regulatory quality indicators associated with malpractice litigation and excess costs.

Conversely, several topics (ie, AI/ML Acceptance, Burnout, Patient Safety, and Unit Culture) were poorly represented and could be areas where there is an opportunity to

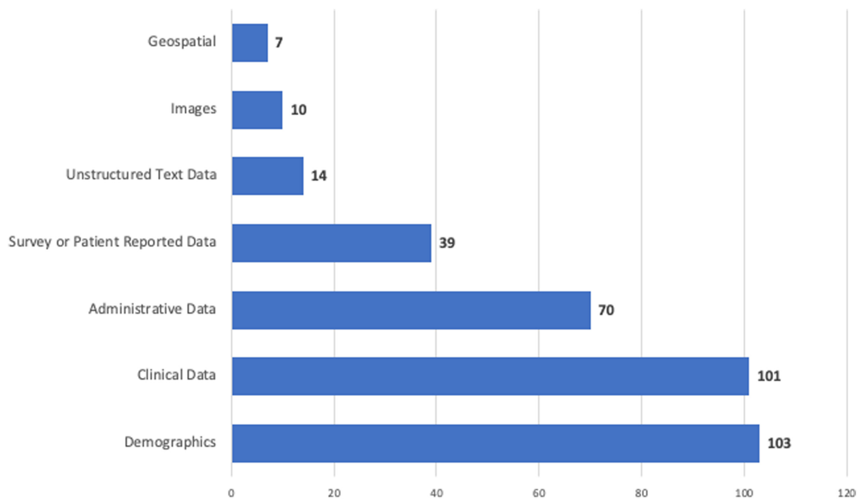

FIGURE 2. Frequency of use of independent variable categories among reviewed studies.

leverage data science methods in research on these nursing topics. In fact, our *Unit Culture* search did not reveal any results that met our inclusion criteria. The sparse results could be a limitation of our search strategy or the ambiguity in these concepts' descriptive terms (which could also be said of *Healthcare Utilization and Costs*); however, it is worth considering that more studies could be performed in these areas. Given nurses' long-standing attention to these latter areas of *Burnout*, *Patient Safety*, and *Unit Culture*, we are hopeful that the nursing research and nursing informatics communities will apply data science methods to these problems in the coming years. It is important to note that most outcomes in our review can be associated with Patient Safety, an essential aspect of healthcare quality.

Moving forward, we suggest data science efforts are best undertaken when data scientists can integrate their computer science knowledge with the clinical knowledge of healthcare providers to promote better tools for analysis (eg, pattern identification) or prediction (machine learning models to predict future patient outcomes) to improve healthcare quality. We believe identifying and applying common data elements relevant to nursing practice across EHR systems in predictive models and including standard nursing terminology codes in the EHR would capture some nuances that provide contextual information about patient health status that are not currently captured. The inclusion of these codes, along with the underrepresented unstructured text notes (eg, comments on flowsheets, progress notes) and geospatial data, could be worth pursuing in future research efforts. In sum, studies using machine learning techniques should include a variety of nursing-generated data, in addition to including nurses on the project team to help understand nuances of that data, in order to improve predictions for patient and process outcomes.

Limitations of our report include the nonexhaustive nature of the literature search and the single-person review process. Given that the intent of the article was to provide readers with a broad overview of nursing-relevant data science activities, an exhaustive literature search was beyond our purpose. For interested readers, we have published search strategies so that others can reproduce our findings and/or perform an exhaustive literature review. The use of a single-person review helped expedite the process of a year-in-review paper. Additionally, because we are focused on high-level description rather than inferential comparisons, the use of a second reviewer would not have significantly changed our findings.

CONCLUSION

Data science has significant potential to assist healthcare providers in improving the nursing environment, clinical processes, and patient outcomes. By using data science techniques to identify care environment improvement opportunities and/or individual patient risk factors, we create new opportunities to design and implement interventions best able to mitigate risk and improve patient care. The use of data science to understand problems related to nursing and nursing care must include modern methods of investigation and understanding. We hope that the studies we have identified and described in this article will help readers understand the breadth and depth of data science's ability to improve clinical processes and patient outcomes that are relevant to nurses.

References

1. Center for Nursing I. Data Science Workgroup. Nursing Knowledge: Big Data Science Conference | School of Nursing - University of Minnesota. z.umn.edu/bigdata. Accessed February 15, 2021..
2. Heslop L, Lu S, Xu X. Nursing-sensitive indicators: a concept analysis. *Journal of Advanced Nursing*. 2014;70(11): 2469–2482.
3. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a Web and mobile app for systematic reviews. *Systematic Reviews*. 2016;5(1): 210.
4. Liu P, Yang R, Xu Z. Public acceptance of fully automated driving: effects of social trust and risk/benefit perceptions. *Risk Analysis*. 2019;39(2): 326–341.
5. Shin D, Park YJ. Role of fairness, accountability, and transparency in algorithmic affordance. *Computers in Human Behavior*. 2019;98: 277–284.
6. Lee Y, Ha M, Kwon S, Shim Y, Kim J. Egoistic and altruistic motivation: how to induce users' willingness to help for imperfect AI. *Computers in Human Behavior*. 2019;101: 180–196.
7. McLean G, Osei-Frimpong K. Hey Alexa…examine the variables influencing the use of artificial intelligent in-home voice assistants. *Computers in Human Behavior*. 2019;99: 28–37.
8. Notenbomer A, van Rhenen W, Groothoff JW, Roelen CAM. Predicting long-term sickness absence among employees with frequent sickness absence. *International Archives of Occupational and Environmental Health*. 2019;92(4): 501–511.
9. Bosman LC, Roelen CAM, Twisk JWR, Eekhout I, Heymans MW. Development of prediction models for sick leave due to musculoskeletal disorders. *Journal of Occupational Rehabilitation*. 2019;29(3): 617–624.
10. Oliver A, Revuelto L, Fernandez I, Simo-Algado S, Galiana L. An integrative model of the subjective well-being of staff working in intellectual disability services. *Research in Developmental Disabilities*. 2019;87: 1–8.
11. Dutra HS, Guirardello EB, Li Y, Cimiotti JP. Nurse burnout revisited: a comparison of computational methods. *Journal of Nursing Measurement*. 2019;27(1): E17–E33.
12. Chiew CJ, Liu N, Tagami T, Wong TH, Koh ZX, Ong MEH. Heart rate variability based machine learning models for risk prediction of suspected sepsis patients in the emergency department. *Medicine (Baltimore)*. 2019;98(6): e14197.
13. Bertsimas D, Dunn J, Steele DW, Trikalinos TA, Wang Y. Comparison of machine learning optimal classification trees with the Pediatric Emergency Care Applied Research Network head trauma decision rules. *JAMA Pediatrics*. 2019;173(7): 648–656.
14. Wu CC, Hsu WD, Islam MM, et al. An artificial intelligence approach to early predict non-ST-elevation myocardial infarction patients with chest pain. *Computer Methods and Programs in Biomedicine*. 2019;173: 109–117.
15. Lo-Ciganic WH, Huang JL, Zhang HH, et al. Evaluation of machine-learning algorithms for predicting opioid overdose risk among Medicare beneficiaries with opioid prescriptions. *JAMA Network Open*. 2019;2(3): e190968.
16. Patterson BW, Engstrom CJ, Sah V, et al. Training and interpreting machine learning algorithms to evaluate fall risk after emergency department visits. *Medical Care*. 2019;57(7): 560–566.
17. Goto S, Kimura M, Katsumata Y, et al. Artificial intelligence to predict needs for urgent revascularization from 12-leads electrocardiography in emergency patients. *PLoS One*. 2019;14(1): e0210103.

18. Bergese I, Frigerio S, Clari M, et al. An innovative model to predict Pediatric emergency department return visits. *Pediatric Emergency Care*. 2019; 35(3): 231–236.
19. Howell P, Elkin PL. Can solo practitioners survive in value-based healthcare? Validating a predicative model for ED utilization. *Studies in Health Technology and Informatics*. 2019;264: 1682–1683.
20. Vest JR, Ben-Assuli O. Prediction of emergency department revisits using area-level social determinants of health measures and health information exchange information. *International Journal of Medical Informatics*. 2019;129: 205–210.
21. Wolff P, Grana M, Rios SA, Yarza MB. Machine learning readmission risk modeling: a pediatric case study. *BioMed Research International*. 2019;2019: 8532892.
22. Goto T, Camargo CA Jr., Faridi MK, Freishtat RJ, Hasegawa K. Machine learning–based prediction of clinical outcomes for children during emergency department triage. *JAMA Network Open*. 2019;2(1): e186937.
23. Kramer J, Schreyogg J, Busse R. Classification of hospital admissions into emergency and elective care: a machine learning approach. *Health Care Management Science*. 2019;22(1): 85–105.
24. Raita Y, Goto T, Faridi MK, Brown DFM, Camargo CA Jr., Hasegawa K. Emergency department triage prediction of clinical outcomes using machine learning models. *Critical care (London, England)*. 2019;23(1): 64.
25. Rendell K, Koprinska I, Kyme A, Ebker-White AA, Dinh MM. The Sydney Triage to Admission Risk Tool (START2) using machine learning techniques to support disposition decision-making. *Emergency Medicine Australasia*. 2019;31(3): 429–435.
26. Pruitt P, Naidech A, Van Ornam J, Borczuk P, Thompson W. A natural language processing algorithm to extract characteristics of subdural hematoma from head CT reports. *Emergency Radiology*. 2019;26(3): 301–306.
27. Sterling NW, Patzer RE, Di M, Schrager JD. Prediction of emergency department patient disposition based on natural language processing of triage notes. *International Journal of Medical Informatics*. 2019;129: 184–188.
28. Leone M, Lapucci E, De Sario M, Davoli M, Farchi S, Michelozzi P. Social network analysis to characterize women victims of violence. *BMC Public Health*. 2019;19(1): 494.
29. Lucero RJ, Lindberg DS, Fehlberg EA, et al. A data-driven and practice-based approach to identify risk factors associated with hospital-acquired falls: applying manual and semi- and fully-automated methods. *International Journal of Medical Informatics*. 2019;122: 63–69.
30. Lo Y, Lynch SF, Urbanowicz RJ, et al. Using machine learning on home health care assessments to predict fall risk. *Studies in Health Technology & Informatics*. 2019;264: 684–688.
31. Oshiro CES, Frankland TB, Rosales AG, et al. Fall ascertainment and development of a risk prediction model using electronic medical records. *Journal of the American Geriatrics Society*. 2019;67(7): 1417–1422.
32. Than S, Crabtree A, Moran C. Examination of risk scores to better predict hospital-related harms. *Internal Medicine Journal*. 2019;49(9): 1125–1131.
33. Klock M, Kang H, Gong Y. Scoring patient fall reports using quality rubric and machine learning. *Studies in Health Technology and Informatics*. 2019;264: 639–643.
34. Bush K, Barbosa H, Farooq S, et al. Predicting hospital-onset *Clostridium difficile* using patient mobility data: a network approach. *Infection Control & Hospital Epidemiology*. 2019;40(12): 1380–1386.
35. Hur EY, Jin Y, Jin T, Lee SM. Development and evaluation of the automated risk assessment system for catheter-associated urinary tract infection. *Computers, Informatics, Nursing: CIN*. 2019;37(9): 463–472.
36. Jackson SS, Lydecker AD, Magder LS, Roghmann MC. Development and validation of a clinical prediction rule to predict transmission of methicillin-resistant *Staphylococcus aureus* in nursing homes. *American Journal of Epidemiology*. 2019;188(1): 214–221.
37. Kocbek P, Fijacko N, Soguero-Ruiz C, et al. Maximizing interpretability and cost-effectiveness of surgical site infection (SSI) predictive models using feature-specific regularized logistic regression on preoperative temporal data. *Computational and Mathematical Methods in Medicine*. 2019;2019: 2059851.
38. Lodise TP, Bonine NG, Ye JM, Folse HJ, Gillard P. Development of a bedside tool to predict the probability of drug-resistant pathogens among hospitalized adult patients with gram-negative infections. *BMC Infectious Diseases*. 2019;19(1): 718.
39. Yee CR, Narain NR, Akmaev VR, Vemulapalli V. A data-driven approach to predicting septic shock in the intensive care unit. *Biomed Inform Insights*. 2019;11: 1178222619885147.
40. Liao YH, Wang ZC, Zhang FG, Abbod MF, Shih CH, Shieh JS. Machine learning methods applied to predict ventilator-associated pneumonia with *Pseudomonas aeruginosa* infection via sensor array of electronic nose in intensive care unit. *Sensors (Basel)*. 2019;19(8): 1866.
41. Lin ED, Hefner JL, Zeng X, et al. A deep learning model for pediatric patient risk stratification. *The American Journal of Managed Care*. 2019;25(10): e310–e315.
42. Xiong GL, Bayen E, Nickels S, et al. Real-time video detection of falls in dementia care facility and reduced emergency care. *The American Journal of Managed Care*. 2019;25(7): 314–315.
43. Kaplan S, Zhu Y-M. Full-dose PET image estimation from low-dose PET image using deep learning: a pilot study. *Journal of Digital Imaging*. 2019;32(5): 773–778.
44. Monge-Alvarez J, Hoyos-Barcelo C, Lesso P, Casaseca-de-la-Higuera P. Robust detection of audio-cough events using local Hu moments. *IEEE Journal of Biomedical and Health Informatics*. 2019;23(1): 184–196.
45. Balyan R, Crossley SA, Brown W 3rd, et al. Using natural language processing and machine learning to classify health literacy from secure messages: the ECLIPPSE study. *PLoS One*. 2019;14(2): e0212488.
46. Forman EM, Goldstein SP, Zhang F, et al. OnTrack: development and feasibility of a smartphone app designed to predict and prevent dietary lapses. *Translational Behavioral Medicine*. 2019;9(2): 236–245.
47. Tang Y, Yang J, Ang PS, et al. Detecting adverse drug reactions in discharge summaries of electronic medical records using Readpeer. *International Journal of Medical Informatics*. 2019;128: 62–70.
48. Liu L, Ni Y, Zhang N, Nick Pratap J. Mining patient-specific and contextual data with machine learning technologies to predict cancellation of children's surgery. *International Journal of Medical Informatics*. 2019;129: 234–241.
49. Desai R, Jo A, Marlow NM. Risk for medication nonadherence among Medicaid enrollees with fibromyalgia: development of a validated risk prediction tool. *Pain Practice*. 2019;19(3): 295–302.
50. Sun C, Ippel L, van Soest J, et al. A privacy-preserving infrastructure for analyzing personal health data in a vertically partitioned scenario. *Studies in Health Technology and Informatics*. 2019;264: 373–377.
51. Ramirez AG, Schneider EB, Mehaffey JH, Zeiger MA, Hanks JB, Smith PW. Effect of travel time for thyroid surgery on treatment cost and morbidity. *The American Surgeon*. 2019;85(9): 949–955.
52. Ramkumar PN, Navarro SM, Haeberle HS, et al. Development and validation of a machine learning algorithm after primary total hip arthroplasty: applications to length of stay and payment models. *The Journal of Arthroplasty*. 2019;34(4): 632–637.
53. Taranik M, Kopanitsa G. Using machine learning for personalized patient adherence level determination. *Studies in Health Technology and Informatics*. 2019;261: 174–178.
54. Hussain OA, Junejo KN. Predicting treatment outcome of drug-susceptible tuberculosis patients using machine-learning models. *Informatics for Health & Social Care*. 2019;44(2): 135–151.
55. Delamater PL, Shortridge AM, Kilcoyne RC. Using floating catchment area (FCA) metrics to predict health care utilization patterns. *BMC Health Services Research*. 2019;19(1): 144.
56. Masters ET, Ramaprasan A, Mardekian J, et al. Natural language processing-identified problem opioid use and its associated health care costs. *Journal of Pain & Palliative Care Pharmacotherapy*. 2018;32(2–3): 106–115.
57. Skaljic M, Patel IH, Pellegrini AM, Castro VM, Perlis RH, Gordon DD. Prevalence of financial considerations documented in primary care encounters as identified by natural language processing methods. *JAMA Network Open*. 2019;2(8): –e1910399.

58. Bigeard E, Thiessard F, Grabar N. Detecting drug non-compliance in Internet Fora using information retrieval and machine learning approaches. *Studies in Health Technology and Informatics*. 2019;264: 30–34.

59. Kan HJ, Kharrazi H, Chang HY, Bodycombe D, Lemke K, Weiner JP. Exploring the use of machine learning for risk adjustment: a comparison of standard and penalized linear regression models in predicting health care costs in older adults. *PLoS One*. 2019;14(3): e0213258.

60. Chang ET, Piegari R, Wong ES, et al. Which patients are persistently high-risk for hospitalization? *American Journal of Managed Care*. 2019; 25(9): e274–e281.

61. Arora S, Stouffer GA, Kucharska-Newton AM, et al. Twenty year trends and sex differences in young adults hospitalized with acute myocardial infarction. *Circulation*. 2019;139(8): 1047–1056.

62. Kattan WM, Abduljawad AA. Predicting different factors that affect hospital utilization and outcomes among diabetic patients admitted with hypoglycemia using structural equation modeling. *Diabetes Research and Clinical Practice*. 2019;153: 55–65.

63. Lim S, Gangoli G, Adams E, et al. Increased clinical and economic burden associated with peripheral intravenous catheter-related complications: analysis of a US Hospital discharge database. *Inquiry*. 2019;56: 46958019875562.

64. Bhavsar GP, Probst JC, Bennett KJ, Hardin JW, Qureshi Z. Community-level electronic prescribing and adverse drug event hospitalizations among older adults. *Health Informatics Journal*. 2019;25(3): 661–675.

65. Hsu DY, Smith B, Silverberg JI. Atopic dermatitis and hospitalization for mental health disorders in the United States. *Dermatitis*. 2019;30(1): 54–61.

66. Vanderlaan J, Rochat R, Williams B, Dunlop A, Shapiro SE. Associations between hospital maternal service level and delivery outcomes. *Women's Health Issues*. 2019;29(3): 252–258.

67. Stockbridge EL, Chhetri S, Polcar LE, Loethen AD, Carney CP. Behavioral health conditions and potentially preventable diabetes-related hospitalizations in the United States: findings from a national sample of commercial claims data. *PLoS One*. 2019;14(2): e0212955.

68. Herran-Monge R, Muriel-Bombin A, Garcia-Garcia MM, et al. Epidemiology and changes in mortality of Sepsis after the implementation of surviving sepsis campaign guidelines. *Journal of Intensive Care Medicine*. 2019; 34(9): 740–750.

69. Malinovska A, Pitasch L, Geigy N, Nickel CH, Bingisser R. Modification of the emergency severity index improves mortality prediction in older patients. *The Western Journal of Emergency Medicine*. 2019;20(4): 633–640.

70. Sanson G, Welton J, Vellone E, et al. Enhancing the performance of predictive models for hospital mortality by adding nursing data. *International Journal of Medical Informatics*. 2019;125: 79–85.

71. Zhang J, Cheng B, Yang M, Pan J, Feng J, Cheng Z. Predicting in-hospital death in patients with type B acute aortic dissection. *Medicine (Baltimore)*. 2019; 98(32): e16462.

72. Cardona M, Lewis E, Shanmugam S, et al. Dissonance on perceptions of end-of-life needs between health-care providers and members of the public: quantitative cross-sectional surveys. *Australasian Journal on Ageing*. 2019;38(3): e75–e84.

73. Hassanipour S, Ghaem H, Arab-Zozani M, et al. Comparison of artificial neural network and logistic regression models for prediction of outcomes in trauma patients: a systematic review and meta-analysis. *Injury*. 2019; 50(2): 244–250.

74. de Lima Junior JD, Matias JEF, Stahlke Junior HJ. Risk factors associated with hospital mortality in mitral valve reoperation. *Revista do Colégio Brasileiro de Cirurgiões*. 2019;46(3): e20192176.

75. Wang L, Yang F, Wang X, et al. Predicting mortality in patients undergoing VA-ECMO after coronary artery bypass grafting: the REMEMBER score. *Critical Care (London, England)*. 2019;23(1): 11.

76. Wu L, Zhang Z, Wang Y, et al. A model to predict in-hospital mortality in HIV/AIDS patients with Pneumocystis pneumonia in China: the clinical practice in real world. *BioMed Research International*. 2019;2019: 6057028.

77. Zeltzer D, Balicer RD, Shir T, Flaks-Manov N, Einav L, Shadmi E. Prediction accuracy with electronic medical records versus administrative claims. *Medical Care*. 2019;57(7): 551–559.

78. Kwon JM, Kim KH, Jeon KH, Park J. Deep learning for predicting in-hospital mortality among heart disease patients based on echocardiography. *Echocardiography*. 2019;36(2): 213–218.

79. Kim Y, Kym D, Hur J, et al. Development of a risk prediction model (Hangang) and comparison with clinical severity scores in burn patients. *PLoS One*. 2019;14(2): e0211075.

80. Haider A, Con J, Prabhakaran K, et al. Developing a simple clinical score for predicting mortality and need for ICU in trauma patients. *The American Surgeon*. 2019;85(7): 733–737.

81. Hill BL, Brown R, Gabel E, et al. An automated machine learning–based model predicts postoperative mortality using readily-extractable preoperative electronic health record data. *British Journal of Anaesthesia*. 2019;123(6): 877–886.

82. Lei VJ, Kennedy EH, ThaiBinh L, et al. Model performance metrics in assessing the value of adding intraoperative data for death prediction: applications to noncardiac surgery. *Studies in Health Technology & Informatics*. 2019;264: 223–227.

83. McCarthy FH, Zhang L, Tam V, et al. Geographically derived socioeconomic factors to improve risk prediction in patients having aortic valve replacement. *The American Journal of Cardiology*. 2019;123(1): 116–122.

84. Park Y, Karampourniotis PD, Sylla I, Das AK. Hierarchical patient-centric caregiver network method for clinical outcomes study. *PLoS One*. 2019; 14(2): e0211218.

85. Pickens RC, Bloomer AK, Sulzer JK, et al. Modifying interhospital hepatopancreatobiliary transfers based on predictive analytics: moving from a Center of Excellence to a health-care system of excellence. *The American Surgeon*. 2019;85(9): 1033–1039.

86. Pieszko K, Hiczkiewicz J, Budzianowski P, et al. Predicting long-term mortality after acute coronary syndrome using machine learning techniques and Hematological markers. *Disease Markers*. 2019;2019: 9056402.

87. Lin K, Hu Y, Kong G. Predicting in-hospital mortality of patients with acute kidney injury in the ICU using random forest model. *International Journal of Medical Informatics*. 2019;125: 55–61.

88. Danilov G, Kotik K, Shifrin M, Strunina U, Pronkina T, Potapov A. Prediction of postoperative hospital stay with deep learning based on 101 654 operative reports in neurosurgery. *Studies in Health Technology and Informatics*. 2019;258: 125–129.

89. Merrill RK, Ferrandino RM, Hoffman R, Shaffer GW, Ndu A. Machine learning accurately predicts short-term outcomes following open reduction and internal fixation of ankle fractures. *The Journal of Foot and Ankle Surgery*. 2019;58(3): 410–416.

90. Tafti AP, Dong Y, Habermann E, Liu H, Herasevich V. Relationship between very cold outside weather and surgical outcome: integrating shallow and deep artificial neural nets. *Studies in Health Technology and Informatics*. 2019;264: 1783–1784.

91. Harutyunyan H, Khachatrian H, Kale DC, Ver Steeg G, Galstyan A. Multitask learning and benchmarking with clinical time series data. *Sci Data*. 2019; 6(1): 96.

92. McCoy TH Jr., Wiste AK, Doyle AE, Pellegrini AM, Perlis RH. Association between child psychiatric emergency room outcomes and dimensions of psychopathology. *General Hospital Psychiatry*. 2019;59: 1–6.

93. Menendez ME, Shaker J, Lawler SM, Ring D, Jawa A. Negative patient-experience comments after total shoulder arthroplasty. *The Journal of Bone and Joint Surgery American Volume*. 2019;101(4): 330–337.

94. Sanson G, Vellone E, Kangasniemi M, Alvaro R, D'Agostino F. Impact of nursing diagnoses on patient and organisational outcomes: a systematic literature review. *Journal of Clinical Nursing*. 2017;26(23–24): 3764–3783.

95. Erin Browne M, Hadjistavropoulos T, Prkachin K, Ashraf A, Taati B. Pain expressions in dementia: validity of observers' pain judgments as a function of angle of observation. *Journal of Nonverbal Behavior*. 2019;43(3): 309–327.

96. Cashin AG, Traeger AC, Hubscher M, et al. Persistent pain after wrist or hand fracture: development and validation of a prognostic model. *The Journal of Orthopaedic and Sports Physical Therapy*. 2019;49(1): 28–35.

97. Chen T, Mu J, Xue Q, et al. Whole-brain structural magnetic resonance imaging-based classification of primary dysmenorrhea in pain-free phase: a machine learning study. *Pain*. 2019;160(3): 734–741.

98. Corradi-Dell'Acqua C, Foerster M, Sharvit G, et al. Pain management decisions in emergency hospitals are predicted by brain activity during empathy and error monitoring. *British Journal of Anaesthesia*. 2019;123(2): e284–e292.

99. Gilpin HR, Stahl DR, McCracken LM. A theoretically guided approach to identifying predictors of treatment outcome in contextual cognitive behavioural therapy for chronic pain. *European Journal of Pain*. 2019;23(2): 354–366.

100. Hah JM, Cramer E, Hilmoe H, et al. Factors associated with acute pain estimation, postoperative pain resolution, opioid cessation, and recovery: secondary analysis of a randomized clinical trial. *JAMA Network Open*. 2019;2(3): e190168.

101. Jiang M, Mieronkoski R, Syrjala E, et al. Acute pain intensity monitoring with the classification of multiple physiological parameters. *Journal of Clinical Monitoring and Computing*. 2019;33(3): 493–507.

102. Kanbayashi Y, Sakaguchi K, Nakatsukasa K, et al. Predictive factors for taxane acute pain syndrome determined by ordered logistic regression analysis. *Supportive Care in Cancer: Official Journal of the Multinational Association of Supportive Care in Cancer*. 2019;27(7): 2673–2677.

103. Larsson B, Dragioti E, Gerdle B, Bjork J. Positive psychological well-being predicts lower severe pain in the general population: a 2-year follow-up study of the SwePain cohort. *Annals of General Psychiatry*. 2019; 18(1): 8.

104. Lee J, Mawla I, Kim J, et al. Machine learning–based prediction of clinical pain using multimodal neuroimaging and autonomic metrics. *Pain*. 2019;160(3): 550–560.

105. Lim H, Kim B, Noh GJ, Yoo SK. A deep neural network–based pain classifier using a photoplethysmography signal. *Sensors (Basel)*. 2019;19(2): 384.

106. Oosterhaven J, Wittink H, Dekker J, Kruitwagen C, Deville W. Pain catastrophizing predicts dropout of patients from an interdisciplinary chronic pain management programme: a prospective cohort study. *Journal of Rehabilitation Medicine*. 2019;51(10): 761–769.

107. Page MG, Boyd K, Ware MA. Examination of the course of low back pain intensity based on baseline predictors and health care utilization among patients treated in multidisciplinary pain clinics: a Quebec Pain Registry Study. *Pain Medicine (Malden, Mass)*. 2019;20(3): 564–573.

108. Parthipan A, Banerjee I, Humphreys K, et al. Predicting inadequate postoperative pain management in depressed patients: a machine learning approach. *PLoS One*. 2019;14(2): e0210575.

109. Priego Quesada JI, Kerr ZY, Bertucci WM, Carpes FP. A retrospective international study on factors associated with injury, discomfort and pain perception among cyclists. *PLoS One*. 2019;14(1): e0211197.

110. Pua YH, Poon CL, Seah FJ, et al. Predicting individual knee range of motion, knee pain, and walking limitation outcomes following total knee arthroplasty. *Acta Orthopaedica*. 2019;90(2): 179–186.

111. Varni JW, Nutakki K, Swigonski NL. Pain, skin sensations symptoms, and cognitive functioning predictors of health-related quality of life in pediatric patients with Neurofibromatosis type 1. *Quality of Life Research*. 2019;28(4): 1047–1052.

112. Wang M, Li Y, Li J, Fan L, Yu H. The risk of moderate-to-severe post-operative pain following the placement of dental implants. *Journal of Oral Rehabilitation*. 2019;46(9): 836–844.

113. Yamada K, Kubota Y, Shimizu Y, et al. Sleep shortage is associated with postherpetic neuralgia development through hyperesthesia and acute pain intensity: a community-based prospective cohort study. *Pain Practice*. 2019;19(5): 476–483.

114. Chapman AB, Peterson KS, Alba PR, DuVall SL, Patterson OV. Detecting adverse drug events with rapidly trained classification models. *Drug Safety*. 2019;42(1): 147–156.

115. Gupta J, Patrick J, Poon S. Clinical safety incident taxonomy performance on C4.5 decision tree and random Forest. *Studies in Health Technology and Informatics*. 2019;266: 83–88.

116. Shiima Y, Wong ZS. Classification scheme for incident reports of medication errors. *Studies in Health Technology and Informatics*. 2019;265: 113–118.

117. Stimpfel AW, Djukic M, Brewer CS, Kovner CT. Common predictors of nurse-reported quality of care and patient safety. *Health Care Management Review*. 2019;44(1): 57–66.

118. Liu J, Wong ZS, Tsui KL, So HY, Kwok A. Exploring hidden in-hospital fall clusters from incident reports using text analytics. *Studies in Health Technology and Informatics*. 2019;264: 1526–1527.

119. Bouzillé G, Morival C, Westerlynck R, et al. An automated detection system of drug-drug interactions from electronic patient records using big data analytics. *Studies in Health Technology & Informatics*. 2019;264: 45–49.

120. Shi C, Dumville JC, Cullum N. Evaluating the development and validation of empirically-derived prognostic models for pressure ulcer risk assessment: a systematic review. *International Journal of Nursing Studies*. 2019;89: 88–103.

121. Li HL, Lin SW, Hwang YT. Using nursing information and data mining to explore the factors that predict pressure injuries for patients at the end of life. *Computers, Informatics, Nursing: CIN*. 2019;37(3): 133–141.

122. Park SK, Park HA, Hwang H. Development and comparison of predictive models for pressure injuries in surgical patients: a retrospective case-control study. *Journal of Wound, Ostomy, and Continence Nursing: Official Publication of the Wound, Ostomy and Continence Nurses Society/ WOCN*. 2019;46(4): 291–297.

123. Flett HM, Delparte JJ, Scovil CY, Higgins J, Laramee MT, Burns AS. Determining pressure injury risk on admission to inpatient spinal cord injury rehabilitation: a comparison of the FIM, spinal cord injury pressure ulcer scale, and Braden scale. *Archives of Physical Medicine and Rehabilitation*. 2019;100(10): 1881–1887.

124. Crane N, Pool N, Chang I, Rogan S, Stocker C, Raman S. A dedicated paediatric logistic organ dysfunction score—adjusted pressure injury risk assessment scale is required for tertiary paediatric ICUs. *Cardiology in the Young*. 2019;29(3): 455–456.

125. Zhang N, Yu X, Shi K, Shang F, Hong L, Yu J. A retrospective analysis of recurrent pressure ulcer in a burn center in Northeast China. *Journal of Tissue Viability*. 2019;28(4): 231–236.

126. Duvall J, Karg P, Brienza D, Pearlman J. Detection and classification methodology for movements in the bed that supports continuous pressure injury risk assessment and repositioning compliance. *Journal of Tissue Viability*. 2019;28(1): 7–13.

127. Ohura N, Mitsuno R, Sakisaka M, et al. Convolutional neural networks for wound detection: the role of artificial intelligence in wound care. *J Wound Care*. 2019;28(Suppl 10): S13–S24.

128. Ehsani H, Mohler MJ, Golden T, Toosizadeh N. Upper-extremity function prospectively predicts adverse discharge and all-cause COPD readmissions: a pilot study. *International Journal of Chronic Obstructive Pulmonary Disease*. 2019;14: 39–49.

129. Katlic MR, Coleman J, Khan K, Wozniak SE, Abraham JH. Sinai abbreviated geriatric evaluation: development and validation of a practical test. *Annals of Surgery*. 2019;269(1): 177–183.

130. Morgan DJ, Bame B, Zimand P, et al. Assessment of machine learning vs standard prediction rules for predicting hospital readmissions. *JAMA Network Open*. 2019;2(3): e190348.

131. Shahrokni A, Tin A, Alexander K, et al. Development and evaluation of a new frailty index for older surgical patients with cancer. *JAMA Network Open*. 2019;2(5): e193545.

132. Sokoreli I, Cleland JG, Pauws SC, et al. Added value of frailty and social support in predicting risk of 30-day unplanned re-admission or death for patients with heart failure: an analysis from OPERA-HF. *International Journal of Cardiology*. 2019;278: 167–172.

133. Wang N, Gallagher R, Sze D, Hales S, Tofler G. Predictors of frequent readmissions in patients with heart failure. *Heart, Lung & Circulation*. 2019;28(2): 277–283.

134. Nijhawan AE, Metsch LR, Zhang S, et al. Clinical and sociobehavioral prediction model of 30-day hospital readmissions among people with HIV and substance use disorder: beyond electronic health record data. *Journal of Acquired Immune Deficiency Syndromes*. 2019;80(3): 330–341.

135. Mahajan SM, Ghani R. Combining structured and unstructured data for predicting risk of readmission for heart failure patients. *Studies in Health Technology and Informatics*. 2019;264: 238–242.

136. Brittan MS, Campagna EJ, Keller D, Kempe A. How measurement variability affects reporting of a single readmission metric. *Journal for Healthcare Quality: Promoting Excellence in Healthcare*. 2019;41(3): 160–164.
137. Chandra A, Rahman PA, Sneve A, et al. Risk of 30-day hospital readmission among patients discharged to skilled nursing facilities: development and validation of a risk-prediction model. *Journal of the American Medical Directors Association*. 2019;20(4): 444–450.e2.
138. Deschepper M, Eeckloo K, Vogelaers D, Waegeman W. A hospital wide predictive model for unplanned readmission using hierarchical ICD data. *Computer Methods and Programs in Biomedicine*. 2019;173: 177–183.
139. Durojaiye OC, Kritsotakis EI, Johnston P, Kenny T, Ntziora F, Cartwright K. Developing a risk prediction model for 30-day unplanned hospitalization in patients receiving outpatient parenteral antimicrobial therapy. *Clinical Microbiology and Infection*. 2019;25(7): 905.e1–905.e7.
140. Franckowiak TM, Raub JN, Yost R. Derivation and validation of a hospital all-cause 30-day readmission index. *American Journal of Health-System Pharmacy: AJHP: Official Journal of the American Society of Health-System Pharmacists*. 2019;76(7): 436–443.
141. Kabue S, Greene J, Kipnis P, et al. The impact of pharmacy-specific predictors on the performance of 30-day readmission risk prediction models. *Medical Care*. 2019;57(4): 295–299.
142. Khera S, Kolte D, Deo S, et al. Derivation and external validation of a simple risk tool to predict 30-day hospital readmissions after transcatheter aortic valve replacement. *EuroIntervention*. 2019;15(2): 155–163.
143. Kim LD, Pfoh ER, Hu B, et al. Derivation and validation of a model to predict 30-day readmission in surgical patients discharged to skilled nursing facility. *Journal of the American Medical Directors Association*. 2019;20(9): 1086–1090.e2.
144. Kwon JY, Karim ME, Topaz M, Currie LM. Nurses "seeing forest for the trees" in the age of machine learning: using nursing knowledge to improve relevance and performance. *Computers, Informatics, Nursing: CIN*. 2019; 37(4): 203–212.
145. Mahajan SM, Ghani R. Using ensemble machine learning methods for predicting risk of readmission for heart failure. *Studies in Health Technology and Informatics*. 2019;264: 243–247.
146. Mahajan SM, Mahajan A, Burman P, Heidenreich P. Can we do more with less while building predictive models? A study in parsimony of risk models for predicting heart failure readmissions. *CIN: Computers, Informatics, Nursing*. 2019;37(6): 306–314.
147. Pauly V, Mendizabal H, Gentile S, Auquier P, Boyer L. Predictive risk score for unplanned 30-day rehospitalizations in the French universal health care system based on a medico-administrative database. *PLoS One*. 2019; 14(3): e0210714.
148. Cediel G, Sandoval Y, Sexter A, et al. Risk estimation in type 2 myocardial infarction and myocardial injury: the TARRACO risk score. *The American Journal of Medicine*. 2019;132(2): 217–226.
149. Chen S, Kong N, Sun X, Meng H, Li M. Claims data-driven modeling of hospital time-to-readmission risk with latent heterogeneity. *Health Care Management Science*. 2019;22(1): 156–179.
150. Baig M, Hua N, Zhang E, et al. Predicting patients at risk of 30-day unplanned hospital readmission. *Studies in Health Technology and Informatics*. 2019;266: 20–24.
151. Nakamura MM, Toomey SL, Zaslavsky AM, et al. Potential impact of initial clinical data on adjustment of pediatric readmission rates. *Academic Pediatrics*. 2019;19(5): 589–598.
152. Zhang S, Tan S, Jiang Y, et al. Sarcopenia as a predictor of poor surgical and oncologic outcomes after abdominal surgery for digestive tract cancer: a prospective cohort study. *Clinical Nutrition*. 2019;38(6): 2881–2888.
153. Holzgrefe RE, Wilson JM, Staley CA, Anderson TL, Wagner ER, Gottschalk MB. Modified frailty index is an effective risk-stratification tool for patients undergoing total shoulder arthroplasty. *Journal of Shoulder and Elbow Surgery*. 2019;28(7): 1232–1240.
154. McAlister F, van Walraven C. External validation of the hospital frailty risk score and comparison with the hospital-patient one-year mortality risk score to predict outcomes in elderly hospitalised patients: a retrospective cohort study. *BMJ Quality & Safety*. 2019;28(4): 284–288.
155. McConachie SM, Raub JN, Trupianio D, Yost R. Development of an iterative validation process for a 30-day hospital readmission prediction index. *American Journal of Health-System Pharmacy: AJHP: Official Journal of the American Society of Health-System Pharmacists*. 2019;76(7): 444–452.
156. Robinson R, Bhattarai M, Hudali T. Vital sign abnormalities on discharge do not predict 30-day readmission. *Clinical Medicine & Research*. 2019; 17(3–4): 63–71.
157. Sieck C, Adams W, Burkhart L. Validation of the BOOST risk stratification tool as a predictor of unplanned 30-day readmission in elderly patients. *Quality Management in Health Care*. 2019;28(2): 96–102.
158. Traven SA, Reeves RA, Sekar MG, Slone HS, Walton ZJ. New 5-factor modified frailty index predicts morbidity and mortality in primary hip and knee arthroplasty. *The Journal of Arthroplasty*. 2019;34(1): 140–144.
159. Goltz DE, Ryan SP, Hopkins TJ, et al. A novel risk calculator predicts 90-day readmission following total joint arthroplasty. *The Journal of Bone and Joint Surgery American Volume*. 2019;101(6): 547–556.
160. Woldu SL, Sanli O, Clinton TN, Lotan Y. Validating the predictors of outcomes after radical cystectomy for bladder cancer. *Cancer*. 2019; 125(2): 223–231.
161. Hatcher VH, Galet C, Lilienthal M, Skeete DA, Romanowski KS. Association of clinical frailty scores with hospital readmission for falls after index admission for trauma-related injury. *JAMA Network Open*. 2019;2(10): e1912409.
162. Nadkarni D, Minocha A, Harpaldas H, et al. Predicting resource-dependent maternal health outcomes at a referral hospital in Zanzibar using patient trajectories and mathematical modeling. *PLoS One*. 2019;14(3): e0212753.
163. Menger V, Spruit M, van Est R, Nap E, Scheepers F. Machine learning approach to inpatient violence risk assessment using routinely collected clinical notes in electronic health records. *JAMA Network Open*. 2019;2(7): e196709.