

# Evidence of Progress in Making Nursing Practice Visible Using Standardized Nursing Data: a Systematic Review

Tamara G. R. Macieira, BSN, PhD student<sup>1</sup>, Madison B. Smith, BSN, RN, PhD student<sup>1</sup>  
Nicolle Davis, BSN, RN, SCRNP, PhD student<sup>1</sup>, Yingwei Yao, PhD<sup>1</sup>, Diana J. Wilkie, PhD,  
RN, FAAN<sup>1</sup>, Karen Dunn Lopez, PhD, MPH<sup>2</sup>, Gail Keenan, PhD, RN, FAAN<sup>1</sup>  
<sup>1</sup>University of Florida, Florida; <sup>2</sup>University of Illinois at Chicago, Illinois

## Abstract

*Nursing care documentation in electronic health records (EHRs) with standardized nursing terminologies (SNTs) can facilitate nursing's participation in big data science that involves combining and analyzing multiple sources of data. Before merging SNTs data with other sources, it is important to understand how such data are being used and analyzed to support nursing practice. The main purpose of this systematic review was to identify studies using SNTs data, their aims and analytical methods. A two-phase systematic process resulted in inclusion and review of 35 publications. Aims of the studies ranged from describing most popular nursing diagnoses, outcomes, and interventions on a unit to predicting outcomes using multi-site data. Analytical techniques varied as well and included descriptive statistics, correlations, data mining, and predictive modeling. The review underscored the value of developing a deep understanding of the meaning and potential impact of nursing variables before merging with other sources of data.*

## Introduction

The main frontline providers of care are nurses who also represent the largest category of health workers in the hospital setting. Among the 2.8 million registered nurses currently working in the United States (U.S.), 61% work in hospitals<sup>1</sup> whereas 19% of 297,100<sup>2</sup> pharmacists and 41.9% of 854,698 physicians in practice work in hospitals.<sup>3,4</sup>

Nurses are responsible 24 hours each day for continuously identifying care issues, implementing and adjusting care prescribed by themselves and other providers to achieve desired patient outcomes. To date, however, it has been difficult to effectively evaluate the impact of nursing on patient outcomes. The growing use of electronic health records (EHRs) to document care now offers the opportunity to use the data captured in practice for discovering knowledge to transform health care. Thus, the documentation entered by nurses into EHRs, for the first time ever, is a potential source for discovering the impact of nursing care on patient outcomes and using the knowledge to improve care. In this article, we report our systematic review of studies that utilized nursing EHRs data to answer a variety of research questions from describing nursing care for a specific population to predicting patient outcomes. The publications reviewed provide a foundation for identifying future paths of inquiry involving nursing and other data retrievable from EHRs.

The use of standardized nursing terminologies (SNTs) to document nursing care enables the easy retrieval and analysis of nursing data while also representing the nurse's clinical reasoning.<sup>5</sup> The integration of nursing data into large datasets requires the frequent and rapid input of new valid information from EHRs.<sup>6</sup> These can be achieved through the use of controlled vocabularies in EHRs, which helps overcome the major challenges of aggregation, processing and analysis associated with unstructured text data.<sup>7-8</sup> In nursing, SNTs are controlled vocabularies that represent nursing care as nursing diagnoses, interventions and outcomes.<sup>8</sup> The SNT coded data retrieved from EHRs can be analyzed alone or merged with other EHRs data. The use of SNTs to document nursing practice is a big step toward supporting the aggregation of nursing data to large datasets and big data science.

Different sets of SNTs are used to document nursing care. The American Nurses Association (ANA) recognizes and supports the use of certain nursing terminologies to guide and document care if those have clear and unambiguous concepts, are coded with a unique identifier per concept, and if those terminologies were tested for reliability, clinical usefulness and validity.<sup>9</sup> The following nursing terminologies are recognized by ANA: NANDA-International (NANDA-I)<sup>10</sup>; Nursing Interventions Classification (NIC)<sup>11</sup>; Nursing Outcomes Classification (NOC)<sup>12</sup>; International Classification for Nursing Practice (ICNP)<sup>13</sup>; Omaha System<sup>14</sup>; Clinical Care Classification (CCC)<sup>15</sup>; and the Perioperative Nursing Data Set (PNDS).<sup>16</sup> While ICNP, Omaha System, CCC and PNDS sets contain diagnoses, interventions and outcomes terms; NANDA-I (diagnoses), NIC (interventions) and NOC (outcomes) are three separate terminologies. Since NANDA-I, NIC and NOC are very often used together, we will refer to them as a terminology set (NNN).

Systematically reporting and analyzing studies that used SNTs nursing data retrieved from EHRs is important to understand the analytic issues related to the complexity and richness of data generated from the use of these

terminologies. Given the growing emphasis on using existing health care datasets, there is a need for new statistical, computational, and visualization methods to analyze EHRs data, given their complexity and volume.<sup>6</sup> For nursing to join this effort, an important first step is to identify and examine the studies that analyze EHR data coded with SNTs.<sup>17</sup>

To date, there are few reviews that report secondary analysis of SNTs nursing data. One recently published systematic review<sup>18</sup> described study focus, sample characteristics and frequency of publications that studied SNTs. The authors also identified a limited set of studies in which SNTs nursing data were being analyzed.<sup>18</sup> However, substantial evaluation and discussion on the analysis of SNTs nursing data were not performed. Another review<sup>19</sup> described characteristics of nursing research data (ranging from patient demographics, social history, medical history, medications, among others) and evaluated if a specific index metadata system represents sufficiently nursing data. Our systematic review differentiates from the previous investigations as we focused only on studies that analyzed nursing data coded with SNTs retrieved from EHRs. The earlier literature review<sup>19</sup> included all types of unstructured nursing data, and different types of studies, such as controlled clinical trials. The authors also did not restrict their selection criteria for nursing data retrieved from EHRs.<sup>19</sup>

The systematic review presented below was conducted to describe and critically analyze the body of studies in which secondary analyses of data coded with the ANA recognized SNTs was performed. These nursing data were documented during the delivery of nursing care and retrieved from EHRs. We believe that the findings from the present systematic review will emphasize the importance of coded nursing data and will encourage a wider use of SNTs that will allow greater participation of nursing in big data science initiatives.

## **Objective**

The objective of this systematic review was to uncover the state of the science related to the use of standardized nursing data (coded with SNTs) retrieved from EHRs to answer research questions, describe the analytical techniques employed, and outline the lessons learned applicable big data science and nursing.

## **Methods**

### *Search strategy*

A comprehensive literature search was conducted to identify publications in which secondary analysis was performed on data extracted from EHRs and documented using terms from the ANA recognized nursing terminologies (or sets). There were two phases used in this process: 1) study selection and 2) data collection process.

Building from the work of Tastan et al.<sup>18</sup> that reported secondary uses of SNTs documentation in studies up to 2011, the databases PubMed and CINAHL were searched using keywords encompassing all ANA recognized SNTs, along with “nursing” and “electronic health records” from 2010 to 2017. Keywords were defined for each database with the help of a librarian, who was a specialist in Consumer Health, Nursing and Health Education and Behavior. The limiters since 2010, abstract available and published in English were used. Potential publications identified in both databases were downloaded into a reference management program (EndNote X7, Thompson Reuters ISI ResearchSoft), in which duplicates were deleted and abstracts were reviewed. Finally, grey literature search was conducted using Google Scholar to identify possible publications not captured by the traditional search methods, including relevant publications not in PubMed or CINAHL, but in computer science databases like IEEE. Name of authors of publications already reviewed and included in this study’s sample were individually searched in Google Scholar. Publications pertinent to the subject were reviewed and included.

### *Study selection*

The studies identified using the specified search strategy were submitted to abstract review according to the following inclusion criteria: 1) publications that conducted secondary analysis of nursing data retrieved from EHRs, 2) nursing data were coded with one of the ANA recognized nursing terminologies and documented at the point-of-care, as part of the institution’s routine, and 3) publications were published in English. Following PRISMA guidelines<sup>20</sup>, Figure 1 shows the number of publications first identified, retained after selection from abstracts, and reviewed as full text.

The team utilized a two-stage selection process. In the first stage, two doctoral students in nursing informatics (TM, ND) independently conducted the literature search in the databases using the same keywords and found the same number of studies. A guide created by the authors was used to systematically review the abstracts. The reviewers first identified if the article contained at least one of the ANA recognized SNTs. If the abstract described the use of a SNT, the reviewer would continue to verify if the data analyzed were retrieved from EHRs. In the case of an affirmative answer, the reviewers would further analyze the abstract to assess if the SNT coded nursing data were documented as

part of routine clinical practice and not solely for research purposes. Publications selected by one reviewer but not the other were discussed until agreement was reached. In the second stage, TM and MS (doctoral students) conducted a full-text review for further eligibility screening.

#### *Data collection process*

Data from the eligible publications were independently collected by TM and MS. Tables were created to summarize and analyze the content of the articles. The following information was extracted: study foci; sample characteristics (number of health records); variables (dependent, independent, descriptive); and statistical analysis. Data extracted from the selected articles were also reviewed by two faculty members, one specialist in nursing informatics and one statistician. Consensus agreement was reached. The present study aimed to report the diverse statistical analysis methods utilized and studies foci addressed through the use of SNTs coded nursing data. We did not intend to report the results of those publications.

#### *Quality appraisal*

Quality appraisal of publications is common in systematic reviews, however, it is challenging when the publications being reviewed are of the secondary data analysis nature. For the purpose of this review, quality appraisal of articles was evaluated regarding completeness of reporting of information. An adapted version of the STROBE Statement-checklist for observational studies<sup>21</sup> was used, including the following items from the original checklist: study size, clear definition of variables, and description of all statistical methods used. Clear definition of variables was scored as 1 (not defined), 2 (partially defined), and 3 (well defined). These ratings considered if the variables in the publications were described with sufficient clarity to be replicated (collected) in further studies. As the previous stages of the review, authors independently (TM, MS) evaluated the completeness of the items mentioned in rounds until agreement was achieved.

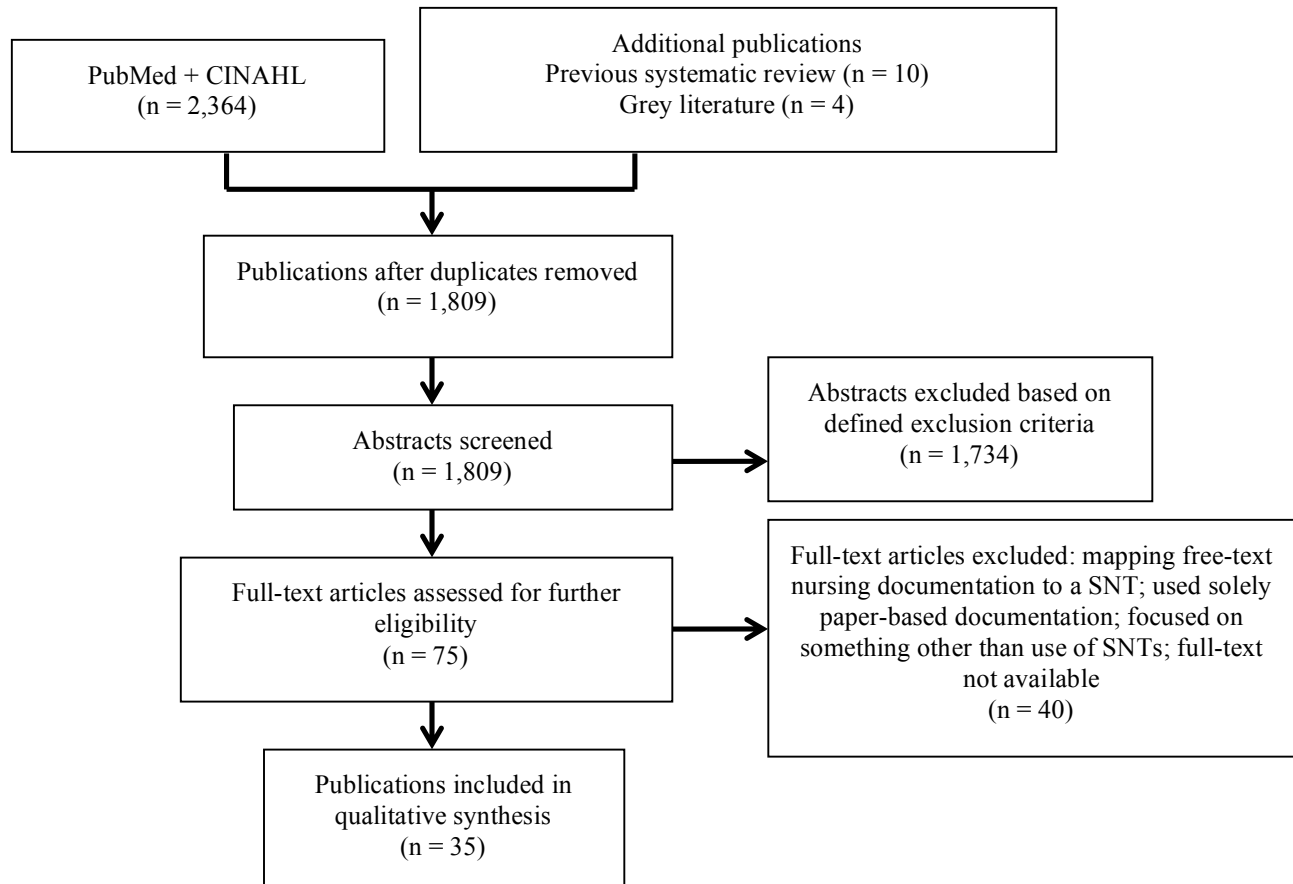
### **Results**

Abstracts of 1,809 publications were examined and 75 publications were retained according to the inclusion criteria. In the data collection process (second stage), of the 75 publications remaining, 40 were excluded due to different reasons as described in Figure 1. A total of 35 publications were included in our systematic review.

The studies included represent considerable diversity in terms of sample sizes, ranging from 29 to 379,601 health records. The largest sample sizes represent more than one unit in a single hospital or multiple hospitals, and more than one primary health care center. In these cases, the same EHR and SNTs were implemented across the institutions. Types of records portrayed in the studies encompass daily entries, or shift entries of nursing information into the EHRs at point-of-care. Episodes of care represent the entire history of a hospitalization, typically consisting of several entries for a unique patient hospitalization. For the purpose of this study, health records were defined as representing unique patients and health records entries were defined as daily/shift documentation of care.

In this review, we found studies for only three of the five terminology sets. There were no studies found for CCC and PNDS. There were 21 NNN articles, 10 Omaha System articles, three articles on ICNP, and one article that evaluated both NANDA-I and ICNP. Although all articles were published in English, we identified studies conducted in countries where English is not an official language. All articles on ICNP were conducted in South Korea, while all articles using the Omaha System were conducted in the US. The studies focusing on NNN were from Spain and U.S.

Descriptive, dependent and independent variables of each study were collected. In studies of descriptive nature, the variables analyzed were the nursing diagnoses, interventions and outcomes relevant to each set of SNTs. For studies in which statistical tests for associations were conducted, the majority of independent variables were the sets of SNTs, along with other factors related to delivery of care and the healthcare environment. Dependent variables in those studies varied significantly, including for example, occurrence of falls<sup>22</sup>, pain<sup>23-24</sup>, gender, age and medical diagnoses<sup>25</sup>, nursing cost<sup>26-27</sup>, among others.



**Figure 1.** PRISMA flow diagram illustrating the selection process of publications.

### *Research focus*

Research foci encompassed different questions that could be answered by SNT coded nursing data. Twelve studies focused on characterization of units and patients regarding the most common nursing diagnoses, nursing interventions and nursing outcomes<sup>28-39</sup> (Table 1). In two studies, authors described potential flaws in the terminologies regarding their content<sup>40-41</sup> (Table 1).

Study of the nursing care through the use of SNTs included differences between standardized and non-standardized plans-of-care<sup>38</sup>; comparison of nursing care provided and standard nursing care from guidelines<sup>32</sup>; and testing different search strategies using nursing interventions to analyze the incidence of a medical complication.<sup>42</sup> The benefits of using SNTs in EHRs to increase availability, validity, and reliability of data for statistical analyses and other research purposes were also described.<sup>34</sup>

SNTs NOC and Omaha System were used across different studies, measuring changes in outcomes from admission to discharge.<sup>43-45</sup> Both terminologies have a grading scale to be used by nurses daily to rate the nursing outcomes for each patient. In the studies<sup>43-45</sup>, authors compared ratings given at admission and discharge to patients as a way of measuring effectiveness of care. Association of those ratings with other variables from EHR was also performed.<sup>43</sup> One study focused on demonstrating differences between two statistical methods in describing changes in ratings for the Omaha System terminology.<sup>46</sup>

**Table 1.** Research foci.

SNT	Research focus
NNN	To describe availability of plans-of-care data <sup>34</sup>
	To describe most prevalent NNN of an SNT-based electronic nursing documentation <sup>37</sup>
	To describe the most frequent NNN documented for hospitalized older patients with a primary discharge diagnosis-related group of pneumonia <sup>30</sup>
	To describe the most frequent NNN documented for hospitalized older patients with a discharge primary diagnosis of heart failure <sup>33</sup>
	To report pain care from admission to discharge or death for end-of-life patients <sup>24</sup>
	To determine statistically significant changes in mean scores of outcomes from admission to discharge for specific medical diagnoses <sup>45</sup>
	To determine changes in outcomes scores from admission to discharge, and describe interventions for pediatric patients with dehydration <sup>44</sup>
	To measure the cost of delivering high surveillance, among hospitalized elders at risk for falling <sup>27</sup>
	To measure association among patient characteristics, unit characteristics, medical, pharmacy, and nursing interventions, and falls for older adults <sup>22</sup>
	To measure the association between high surveillance and low surveillance and the occurrence of failure to rescue <sup>48</sup>
	To establish association between degree of severity of problems and psychiatric diagnosis, number of nursing diagnoses, age, gender <sup>25</sup>
	To determine associations between the variables nursing characteristics and nursing cost per acute care episode <sup>26</sup>
	To identify patient and nurse-related factors associated with nursing outcomes for end-of-life patients with death anxiety <sup>51</sup>
	To report of NOC scores linked with the most frequent nursing interventions, age and length of stay <sup>36</sup>
	To describe the difference between plans-of-care with NNN and plans-of-care without NNN in relation to intermediate health outcomes <sup>38</sup>
	To measure association between the variable nurse continuity and the occurrence of pressure ulcers <sup>56</sup>
OS	Describe data mining techniques to predict whether a patient would meet the expected pain related outcomes <sup>23</sup>
	To develop predictive models that show how patient and nursing variables impact comfortable death outcome in end-of-life patients <sup>52</sup>
	To predict current NOC rating for each nurse shift based on previous shift ratings and factors related to nurse shift <sup>53</sup>
	To create predictive models that determine whether a patient with pain problems will be re-admitted to a hospital <sup>55</sup>
	To describe infomarkers that mark a shift in pain outcomes from standard care to palliative care among patients that died during hospitalization <sup>54</sup>
	To describe client problems, interventions, categories and targets, along with outcomes documented for patients in a low-birth weight program <sup>28</sup>
	To describe all Omaha System data elements across multiple homecare settings using two different EHR vendors <sup>29</sup>
	To describe client risk index, family home visiting interventions between groups of low- and high-risk <sup>31</sup>
	To describe free text entry for the category Signs and Symptoms of the Omaha Classification <sup>40</sup>
	To describe free text entries associated with Omaha System on a software <sup>41</sup>
ICNP	To describe ratings of knowledge, behavior, and status on admission and discharge for Latina adolescent and adult mothers with mental problems <sup>43</sup>
	To demonstrate differences between <i>p</i> -values and Cohen's <i>d</i> in describing clinically meaningful changes in scores for Omaha System problems <sup>46</sup>
	To determine which group of interventions predicts the outcome hospitalization for frail and non-frail elders <sup>47</sup>
	To measure association between patient, support system factors and nursing interventions and improvement in urinary and bowel incontinence <sup>49</sup>
	To determine the association between nurses, Omaha System interventions and client characteristics, and variability in health literacy <sup>50</sup>
N/ICNP	To describe variability in nursing intervention to prevent and treat pressure ulcer patients <sup>32</sup>
	To describe nursing preventive pressure-ulcer interventions and to compare with measures from two published guidelines <sup>35</sup>
	To describe narrative nursing statements of cancer patients treated with cisplatin-based chemotherapy <sup>35</sup>
N/ICNP	To test different computerized search strategies, to analyze the incidence and clinical characteristics of contrast media hypersensitivity <sup>42</sup>
N/ICNP	To describe key nursing diagnoses from the two terminologies for patients with heart failure <sup>39</sup>

Key = NNN (NANDA-I, NOC, NIC); OS (Omaha System); N/ICNP (NANDA-I and ICNP).

NOC and NIC were used to estimate nursing cost.<sup>26-27</sup> Jenkins and Welton<sup>26</sup> used the NOC rating scale as a measure of nursing intensity, and in turn, nursing cost. A different measure of nursing cost was used by Shever et al.<sup>27</sup> In their study, NIC was used to measure the cost of delivering the nursing intervention surveillance. Surveillance was used as a measure of intensity of treatment, based on the average number of times a day that it was delivered (low or high). The number of times the intervention was performed was associated with the total hospital cost for the population under study.<sup>27</sup>

In two studies the researchers examined relationships between nursing interventions and outcomes.<sup>47-48</sup> Other (n=4) studies in the sample focused on establishing associations among a broader number of variables, such as patient characteristics, support system factors; nursing characteristics, nurse continuity, degree of severity of problems; medical diagnoses, pharmacy and nursing interventions, and nursing outcomes.<sup>22,25,50-51</sup> Lodhi et al.<sup>51</sup>, Lodhi et al.<sup>52</sup>, and Lodhi et al.<sup>53</sup> identified patient and nurse-related factors associated with meeting expected nursing outcomes.

Almasalha et al.<sup>23</sup> focused on the use of SNTs to predict whether a patient would meet the expected outcomes by the end of a hospitalization period. Yao et al.<sup>54</sup> identified specific elements of SNT related to changes in outcomes that could be used as an indicator for the adjustment from standard nursing care to palliative nursing care. Lastly, Lodhi et al.<sup>55</sup> explored the creation of different predictive models for re-admission of patients that were diagnosed with pain during their last hospitalization.

**Table 2.** Progression of nursing science and knowledge.

Dataset	Year	Sample*	Statistical analyses
1 university hospital, 2 large community hospitals, and 1 small community hospital	2012 <sup>34</sup>	40,747	Descriptive statistics
	2013 <sup>23</sup>	569	Descriptive statistics, Data mining, Chi-square test, t-test
	2013 <sup>24</sup>	596	Chi-square, Wald tests
	2014 <sup>51</sup>	432	Data mining, Chi-square
	2015 <sup>52</sup>	438	Data mining, Logistic regression, Chi-square
	2015 <sup>53</sup>	160	Data mining, Logistic regression, Pearson's two proportion z-test
	2015 <sup>56</sup>	840	Descriptive statistics, Logistic regression
3 community hospitals	2015 <sup>54</sup>	901	Known-group comparative analysis, ANOVA, Tukey's post-hoc test
	2017 <sup>55</sup>	2,300	Data mining
	2007 <sup>44</sup>	29	Descriptive statistics, t-test
1 academic medical center	2011 <sup>30</sup>	451	Descriptive statistics
	2011 <sup>33</sup>	302	Descriptive statistics
1 homecare and 1 maternal home visiting program	2008 <sup>27</sup>	7,851	Propensity scores, Regression with generalized estimating equations
	2011 <sup>22</sup>	7,851	Generalized estimating equations
1 public health agency	2010 <sup>40</sup>	3,388	Descriptive statistics
	2011 <sup>41</sup>	61,701**	Descriptive statistics
	2007 <sup>28</sup>	75	Descriptive statistics
	2011 <sup>31</sup>	486	Descriptive statistics
15 home health agencies	2013 <sup>46</sup>	1,016	Paired samples t-test, Effect sizes (Cohen's d)
	2015 <sup>50</sup>	141	Logistical mixed-effects model
	2010 <sup>29</sup>	2,900	Descriptive statistics
1 university hospital	2011 <sup>47</sup>	1,750	Data mining, Logistic regression
	2011 <sup>49</sup>	2,072	Descriptive statistics, Logistic regression, Data mining, Chi-square
1 university hospital	2011 <sup>32</sup>	41,891	Descriptive statistics
	2012 <sup>42</sup>	759**	Descriptive statistics
	2012 <sup>35</sup>	Study 1: 427 Study 2: 355	Study 1: Descriptive statistics, ANOVA Study 2: Descriptive statistics

Key = \* health records; \*\* health records entries

#### Statistical analysis

A total of 15 studies used only descriptive statistics to analyze their data. The following statistical analysis methods were present among the remaining studies: propensity scores<sup>26,48</sup>; logistic regression and other regression methods<sup>25-27,47-48,52,55-56</sup>; generalized estimating equations<sup>22,27</sup>; data mining techniques<sup>22,47,51-53,55</sup>; chi-square<sup>23-24,51-53</sup>; t-test<sup>23,39,44-46</sup>;

ANOVA<sup>35,54</sup>; Pearson correlation<sup>25,53</sup>; Tukey's post-hoc test<sup>54</sup>; logistical mixed-effects model<sup>50</sup>; effect sizes (Cohen's d)<sup>46</sup>; general linear mixed methods models<sup>43</sup>; and known-group comparative analysis.<sup>54</sup>

Data from same institutions were used across different publications to answer a diverse number of research questions. Table 2 shows the progression of science and knowledge in nursing across the years. Publications are grouped by dataset (each group of institutions used the same EHR). It can be seen (see Table 2) that earlier publications used mostly descriptive statistics and advanced to generalized estimating equations, logistic regressions, and data mining procedures. Sample sizes varied across publications and more specific and complex statistical analysis were related to smaller parts of the datasets.

#### *Quality of publications included*

Among the 35 publications included in this systematic review, all described their study size, which was represented by number of health records or health records entries, and listed statistical analyses used. Definitions of variables were scored for the 35 publications; 74% of the publications<sup>22-25,27-33,37,39-40,44-48,51-56</sup> scored a 3 (variables well defined), 23% of the publications<sup>26,34-35,38,41,42,49-50</sup> scored a 2, and 3% of the publications<sup>43</sup> scored a 1. Overall, the studies were rated as good for completeness of reporting of information.

#### **Discussion**

The present systematic review provides important information about the types of studies that have been conducted on SNTs coded nursing data retrieved from EHRs. The 35 studies were identified and reviewed through a comprehensive systematic process. The SNTs NANDA-I, NOC, NIC and Omaha System were those implemented most often across different EHRs. Specifically, NNN was represented more in the hospital setting, while Omaha System was present in the primary level of healthcare. Worth noting however is that both of these terminologies can be used interchangeably between hospitals and primary health care centers.

There was great diversity in the questions addressed in each study and the variables analyzed. Characterization of nursing care through the most common nursing diagnoses, interventions and outcomes was the aim of some studies. Others focused on determining significant changes in outcomes from admission to discharge, measure cost of nursing interventions, establishing associations using variables inherent to nursing care and other variables found in EHRs. The creation and use of predictive outcomes models were also studied. This range of study foci show the potential of obtaining and translating new knowledge from nursing data coded with SNTs and the holistic approach of nursing care. It also brings the attention of researchers to the possibility of asking new questions aimed at improving patient outcomes, nursing practice and the healthcare system.

Among the studies included in the sample, most investigators either utilized descriptive analysis (mean, standard deviation, frequency, percentage, etc.) to describe terminology usage in different healthcare settings or relied on classical tests (e.g., ANOVA, Tukey's Post Hoc, Chi-square, t-test, Mann Whitney U test, correlation test) to determine bivariate associations. Although of limited nature, descriptive studies are a first step to understand a dataset and the target population under study. Descriptive statistics are also very efficient and helpful in the identification of predictors with low variability that could be excluded in further analysis.<sup>57</sup> The descriptive studies included in this systematic review provide a foundation for the nursing scientific community to move forward to the use of more sophisticated statistical methods.

The association analyses were those between the adoption of SNTs and patients' outcomes, between nursing diagnosis and episode length, and between nursing interventions and patient outcomes. A number of articles conducted multivariate analysis utilizing regression analysis (e.g., generalized linear mixed regression, generalized estimating equations, logistic regression, etc.) incorporating techniques such as propensity score and clustering to examine the effects of patient, nurse, and care characteristics on patient outcomes. Finally, there were a few studies<sup>51-53,56</sup> using data mining methods to identify hidden patterns among nursing care data or to construct predictive models for patient outcomes. Interesting results were reported in these studies. For example, the pain management, medication, management, and positioning interventions were collectively associated with pain relief among end-of-life patients.<sup>23</sup> Continuity of nursing care was not associated with occurrence of pressure ulcers<sup>56</sup> and predictive modeling on large EHR SNTs data can be used as foundation in nursing decision support.<sup>53</sup>

Our systematic review revealed strengths of using SNTs to code nursing data. Data stored in digital format that are unstructured or not standardized pose difficult challenges to data processing and analysis.<sup>19,57</sup> Unstructured data may obscure linkages among elements, adversely impact the validity of the data collected, and inadvertently fuel misunderstandings and errors. Words that describe care might have different meanings from unit to unit, and between

institutions, due to the fact that standards of care and cultures can vary considerably.<sup>57</sup> The processing of unstructured data is also more time-consuming and labor-intensive when compared to SNT coded nursing data.

Nursing data coded with SNTs facilitates comparisons of patient care within and across institutions.<sup>17</sup> The evidence generated in turn can enhance the decision-making process<sup>17</sup> and demonstrate the impact of nursing care making nursing practice visible.<sup>57</sup> The use of valid standardized nursing terminologies makes extraction of nursing data easier<sup>57</sup>, reduces the cost of reporting quality measures<sup>58</sup>, and produces interoperable data.<sup>18</sup>

There are some limitations of this systematic review. The search focused on two databases, which excluded research studies that could have been published in books or sources not indexed by the selected databases, although some articles from other databases were identified through grey literature search. Also, the performed search strategies and the search terms used to uncover the literature may have excluded publications on the subject that used different search terms or keywords. That the focus of our review was on study methods rather than study results limits the overall conclusions that are informed by this review and point to the need for a review of study results.

A limitation identified among the studies included in this systematic review is the scarce integration and analysis of SNT coded nursing data with other parts of EHRs. In our sample, only four studies included other variables from EHRs in their analysis, such as medical diagnoses, pharmaceutical treatments and pressure ulcer rating scale.<sup>22,25,27,32</sup> To be a part of big data science, nursing data not only need to be coded with SNTs, but also should be combined with other parts of the EHRs in order to enable big data analysis. An example of big data analysis in the healthcare field would be to examine relationships among nursing diagnoses, medical conditions and laboratory results and their impact on patients' outcomes.<sup>10</sup>

A main lesson learned for nursing and big data science is the extraordinary potential of coding nursing care with SNTs and using the data to understand the impact of nursing and continuously expand nursing science.<sup>10</sup> Our systematic review highlights the importance of SNTs to enable faster retrieval and processing of nursing data and the application of big data analysis. Big data science can benefit from the integration of nursing data to other datasets. The care provided by nurses is unique and an important contribution to patient outcomes. Thus, including the nursing component in analyses involving health care outcomes is essential to learning the improvements nurses can make to enhance the overall patient outcomes achieved. In the absence of analyzable nursing care data, the impact of nursing will continue to be elusive. This review provided a lens into the uses and potential benefits of SNTs now and in the future.

## Conclusion

The use of SNTs to code nursing data enables the aggregation of this fundamental part of EHRs to big data datasets. Secondary data analyses of nursing data using methods such as data mining and clustering techniques should be used more often as a way of finding more meaningful results that could change positively the nursing practice and impact directly the care received by patients. This systematic review underscored the value of developing a deep understanding of the meaning and potential impact of nursing variables before merging with other sources of data. Further research is needed to analyze nursing data along with other parts of the EHRs, revealing possible associations among care variables from different health professions.

## References

1. Bureau of Labor Statistics, U.S. Department of Labor. Occupational outlook handbook, 2016-17 edition, Registered Nurses. Washington: United States Department of Labor; 2015. Available from <https://www.bls.gov/ooh/healthcare/registered-nurses.htm#tab-3>
2. Bureau of Labor Statistics, U.S. Department of Labor. Occupational outlook handbook, 2016-17 edition, Pharmacists. Washington: United States Department of Labor; 2015. Available from <https://www.bls.gov/ooh/healthcare/pharmacists.htm#tab-3>
3. U.S. Department of Health and Human Services, Health Resources and Services Administration, National Center for Health Workforce Analysis. The U.S. health workforce chartbook. Rockville, Maryland; 2013. 40 p. Available from <https://bhw.hrsa.gov/sites/default/files/bhw/nchwa/chartbookpart1.pdf>
4. Statista. Statista dossier on physicians in the U.S. Statista; 2016. 91 p. Available from <https://www.statista.com/study/12867/physicians-in-the-us-statista-dossier/>
5. Gallagher-Lepak S. Nursing diagnosis basics. In Herdman TH, Kamitsuru S, editors. NANDA international nursing diagnoses: definitions & classification, 2015–2017. Oxford: Wiley Blackwell; 2014. p. 21-30.
6. Baro E, Degoul S, Beuscart R, Chazard, E. Toward a literature-driven definition of big data in healthcare. *BioMed Res Int.* 2015;639021-9:1-9



7. Van Poucke S, Thomeer M, Hadzic A. 2015, big data in healthcare: for whom the bell tolls? *Crit. Care.* 2015; 19(1):171-2.
8. Bernhart-Just A, Lassen B, Schwendimann R. Representing the nursing process with nursing terminologies in electronic medical record systems: a swiss approach. *Comput Inform Nurs.* 2010; 28(6):345-52.
9. Lundberg C, Warren JJ, Brokel J, et al. Selecting a standardized terminology for the electronic health record that reveals the impact of nursing on patient care. *Online J Nurs Inform.* 2008;12(2):1-20.
10. Herdman TH, Kamitsuru S, editors. *NANDA international nursing diagnoses: definitions & classification, 2015–2017.* Oxford: Wiley Blackwell; 2014.
11. Bulechek G, Butcher H, Dochterman J, Wagner C, editors. *Nursing interventions classification (NIC).* 6th ed. St. Louis: Elsevier; 2013.
12. Moorhead S, Johnson M, Maas M, Swanson E, editors. *Nursing outcomes classification (NOC).* 5th ed. St. Louis: Elsevier; 2013.
13. International Council of Nurses. *International classification for nursing practice (ICNP®).* 2015. Available from <http://www.icn.ch/what-we-do/international-classification-for-nursing-practice-icnpr/>
14. Martin KS. *The Omaha system: a key to practice, documentation, and information management.* 2nd ed. Omaha: Health Connections Press; 2005.
15. Saba V, editor. *Clinical care classification (CCC) system version 2.5.* 2nd ed. New York: Springer; 2012.
16. Beyea S, editor. *Perioperative nursing data set.* 2nd ed. Denver: AORN; 2002.
17. Pruinelli L, Delaney C, Garcia A, Caspers B, Westra B. Nursing management minimum data set: cost-effective tool to demonstrate the value of nurse staffing in the big data science era. *Nurs Econ.* 2016;34(2):66-89.
18. Tastan S, Linch GCF, Keenan GM, et al. Evidence for the existing american nurses association-recognized standardized nursing terminologies: a systematic review. *Int J Nurs Stud.* 2014;51(8):1160-70.
19. Kim H, Jang I, Quach J, Richardson A, Kim J, Choi J. Explorative analyses of nursing research data. *West J Nurs Res.* 2017;39(1):5-19.
20. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med.* 2009;151(4):264-70.
21. von Elm E, Altman DG, Egger M, et al. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Int J Surg.* 2014;12:1495-9.
22. Titler MG, Shever LL, Kanak MF, Picone DM, Qin R. Factors associated with falls during hospitalization in an older population. *Res Theor Nurs Pract.* 2011;25(2):127-52.
23. Almasalha F, Xu D, Keenan G, et al. Data mining nursing care plans of end-of-life patients: a study to improve healthcare decision making. *Int J Nurs Knowl.* 2013;24(1):15-24.
24. Yao Y, Keenan G, Al-Masalha F, et al. Current state of pain care for hospitalized patients at end of life. *Am J Hosp Palliat Care.* 2013;30(2):128-36.
25. Escalada-Hernandez P, Munoz-Hermoso P, Gonzalez-Fraile E, et al. A retrospective study of nursing diagnoses, outcomes, and interventions for patients with mental disorders. *Appl Nurs Res.* 2015;28(2):92-8.
26. Jenkins P, Welton J. Measuring direct nursing cost per patient in the acute care setting. *J Nurs Adm.* 2014;44(5):257-62.
27. Shever LL, Titler MG, Kerr P, Qin R, Kim T, Picone DM. The effect of high nursing surveillance on hospital cost. *J Nurs Scholarsh.* 2008;40(2):161-9.
28. Orlygsdottir B. Use of NIDSEC-Compliant CIS in community-based nursing-directed prenatal care to determine support of nursing minimum data set of objectives. *Comput Inform Nurs.* 2007;25(5):283-93.
29. Westra BL, Oancea C, Savik K, Marek KD. The feasibility of integrating the Omaha system data across home care agencies and vendors. *Comput Inform Nurs.* 2010;28(3):162-71.
30. Head BJ, Scherb CA, Reed D, et al. Nursing diagnoses, interventions, and patient outcomes for hospitalized older adults with pneumonia. *Res Gerontol Nurs.* 2011;4(2):95-105.
31. Monsen KA, Radosevich DM, Kerr MJ, Fulkerson JA. Public health nurses tailor interventions for families at risk. *Public Health Nurs.* 2011;28(2):119-28.
32. Park H, Cho I, Chung E. Exploring use of a clinical data repository containing international classification for nursing practice-based nursing practice data. *Comput Inform Nurs.* 2011;29(7):419-26.
33. Scherb CA, Head BJ, Maas ML, et al. Most frequent nursing diagnosis, nursing interventions and nursing sensitive patient outcomes of hospitalized older adults with heart failure: part 1. *Int J Nurs Terminol Classif.* 2011;22(1):13-22.
34. Keenan G, Yakel E, Yao Y, et al. Maintaining a consistent big picture: meaningful use of a web-based POC EHR system. *Int J Nurs Knowl.* 2012;23(3):119-33.

35. Park H, Cho I, Ahn H. Use of narrative nursing records for nursing research. *Proceedings of the NI 2012 Congress*; 2012 June 23-27; Montreal, Canada. 316-21.
36. Tseng H, Moorhead S. The use of standardized terminology to represent nursing knowledge: nursing interventions relevant to safety for patients with cancer. *Stud Health Technol Inform*. 2014;201:298-303.
37. Park H, Lee E. Incorporating standardized nursing languages into an electronic nursing documentation system in Korea: a pilot study. *Int J Nurs Knowl*. 2015;26(1):35-42.
38. Rivas PFJ, Martín-Iglesias S, del Cerro JLP, Arenas CM, López MG, Lagos MB. Effectiveness of nursing process use in primary care. *Int J Nurs Knowl*. 2016;27(1):43-8.
39. Rabelo-Silva ER, Cavalcanti ACD, Caldas MCRG, et al. Advanced nursing process quality: comparing the international classification for nursing practice (ICNP) with the NANDA-international (NANDA-I) and nursing interventions classification (NIC). *J Clin Nurs*. 2016;26:379-87.
40. Melton GB, Westra BL, Raman N, et al. Informing standard development and understanding user needs with Omaha system signs and symptoms text entries in community-based care settings. *AMIA 2010 Annual Symposium Proceedings*. 2010;512-6.
41. Farri O, Monsen K, Westra B, Melton G. Analysis of free text with Omaha system targets in community-based care to inform practice and terminology development. *Appl Clin Inform*. 2011;2(3):304-16.
42. Kim Y, Kim K, Kim H, et al. Surveillance of contrast-media-induced hypersensitivity reactions using signals from an electronic medical recording system. *Ann Allergy Asthma Immunol*. 2012;108(3):167-71.
43. Garcia C, McNaughton D, Radosevich DM, Brandt J, Monsen K. Family home visiting outcomes for latina mothers with and without mental health problems. *Public Health Nurs*. 2013;30(5):429-38.
44. Scherb CA, Stevens MS, Busman C. Outcomes related to dehydration in the pediatric population. *J Pediatr Nurs*. 2007;22(5):376- 82.
45. Scherb CA. Outcome research: making a difference in practice. *Outcome Manag*. 2002;6(1):22-6.
46. Johnson K, McMorris B, Raynor L, Monsen K. What big size you have! using effect sizes to determine the impact of public health nursing interventions. *Appl Clin Inform*. 2013;4(3):434-44.
47. Monsen KA, Westra BL, Oancea SC, Yu F, Kerr MJ. Linking home care interventions and hospitalization outcomes for frail and non-frail elderly patients. *Res Nurs Health*. 2011;34(2):160-68.
48. Shever LL. The impact of nursing surveillance on failure to rescue. *Res Theory Nurs Pract*. 2011;25(2):107-26.
49. Westra BL, Savik K, Oancea C, Choromanski L, Holmes JH, Bliss D. Predicting improvement in urinary and bowel incontinence for home health patients using electronic health record data. *J Wound Ostomy Continence Nurs*. 2011;38(1):77-87.
50. Monsen KA, Chatterjee SB, Timm JE, Poulsen JK, McNaughton DB. Factors explaining variability in health literacy outcomes of public health nursing clients. *Public Health Nurs*, 2015;32(2):94-100.
51. Lodhi MK, Cheema U, Stifter J, et al. Hospitalized end of life patients' death anxiety captured from a structured electronic record: difference by patient and nurse characteristics. *Res Gerontol Nurs*. 2014;7(5):224-34.
52. Lodhi MK, Ansari R, Yao Y, Keenan GM, Wilkie DJ, Khokhar AA. Predictive modeling for comfortable death outcome using electronic health records. *Proceedings of the 2015 IEEE International Conference on Big Data*; 2015 Oct 29-Nov 01; Santa Clara, USA; 2015. p. 409-15.
53. Lodhi MK, Stifter J, Yao Y, et al. Predictive modeling for end-of-life pain outcome using electronic health records. In Perner P, editor. *Advances in data mining: applications and theoretical aspects. ICDM 2015: Proceedings of Lecture Notes in Computer Science*; 2015 July 11-24; Hamburg, Germany. Springer; 2015. p. 56-68.
54. Yao Y, Stifter J, Ezenwa MO, et al. Informarkers for transition to goals consistent with palliative care in dying patients. *Palliat Support Care*. 2015;13(5):1-8.
55. Lodhi MK, Ansari R, Yao Y, Keenan GM, Wilkie D, Khokhar AA. Predicting hospital re-admissions from nursing care data of hospitalized patients. *Proceedings of the 2016 Industrial Conference on Data Mining*; 2017 Dec 13-15; New York, USA. 2017. p. 1-15.
56. Stifter J, Yao Y, Lodhi MK, et al. Nurse continuity and hospital-acquired pressure ulcers: a comparative analysis using an electronic health record "big data" set. *Nurs Res*. 2015;64(5):361-71.
57. Khokhar A, Lodhi MK, Yao Y, Ansari R, Keenan G, Wilkie DJ. Framework for mining and analysis of standardized nursing care plan data. *West J Nurs Res*. 2017;39(1):20-41.
58. Westra BL, Latimer GE, Matney SA, et al. A national action plan for sharable and comparable nursing data to support practice and translational research for transforming health care. *J Am Med Inform Assoc*. 2015;22(3):600-7.