Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Available online at www.sciencedirect.com

ScienceDirect

Nurs Outlook 69 (2021) 257–264

Nursing Outlook

www.nursingoutlook.org

The nurse COVID and historical epidemics literature repository: Development, description, and summary

Figaro L LorestoJr., PhD, RN[a,b*], *Lisa Nunez, BS*[c], *Lindsey Tarasenko, PhD, RN*[a,b], *Marie St. Pierre, MLIS, AHIP*[a], *Kenneth Oja, PhD, RN*[d], *Mallory Mueller, MSN, RNC-NIC*[a], *Bailey Switzer, BSN, RN, CNOR*[a], *Katherine Marroquin, BS*[a], *Catherine Kleiner, PhD, RN*[a,b]

[a]*Division of Patient Care Services, Children's Hospital Colorado, Aurora, CO*

[b]*University of Colorado College of Nursing, CU Anschutz, Aurora, CO*

[c]*Preventive Medicine and Community Health, University of Texas Medical Branch, Galveston, TX*

[d]*Nursing Education and Research, Denver Health, Denver, CO*

ARTICLE INFO

Article history:
Received 24 August 2020
Received in revised form
10 December 2020
Accepted 29 December 2020
Available online January 30, 2021.

Keywords:
Text mining
Nursing
COVID-19
Resource

ABSTRACT

Background: During COVID-19, a Kaggle challenge was issued to data scientists to leverage text mining to provide high-level summaries of full-text articles in the COVID-19 Open Research Dataset (CORD-19) data set, a data set containing articles around COVID-19 and other epidemics. A question was asked: "What if nursing had something similar?"

Purpose: Describe the development and function of the Nursing COVID and Historical Epidemic Literature and describe high-level summaries of abstracts within the repository.

Method: Nurse-specific literature was abstracted from two data sets: CORD-19 and LitCOVID. LitCOVID is a data set containing the most up-to-date literature around COVID-19. Multiple text mining algorithms were utilized to provide summaries of the articles.

Discussion: As of July 2020, the repository contains 760 articles. Summaries indicate the importance of psychological support for nurses and of high-impact rapid education.

Conclusion: To our knowledge, this repository is the only repository specific for nursing that utilizes text mining to provide summaries.

Cite this article: Loresto, F.L., Nunez, L., Tarasenko, L., Pierre, M.S., Oja, K., Mueller, M., Switzer, B., Marroquin, K., & Kleiner, C. (2021, May/June). The nurse COVID and historical epidemics literature repository: Development, description, and summary. Nurs Outlook, 69(3), 257–264. https://doi.org/10.1016/j.outlook.2020.12.017.

Background

The COVID-19 pandemic has necessitated increased scholarly work around the virus, from the virus's characteristics to the treatment and management of the virus. Published work around the virus is increasing at a rapid pace. The scientific community, in response to the pandemic, has provided open access to this published literature. Two literature repositories, the COVID-19 Open Research Dataset Challenge (CORD-19)

*Corresponding author: Figaro L Loresto Jr., Children's Hospital Colorado, 13123 East 16th Avenue, Box 725, Aurora, CO 80045.
E-mail address: floresto84@gmail.com (F.L. Loresto).

0029-6554/$ -see front matter © 2021 Elsevier Inc. All rights reserved.

https://doi.org/10.1016/j.outlook.2020.12.017

and LitCovid, contain over 230,000 published manuscripts related to COVID-19 and historical epidemics (Center for Security and Emerging Technology, 2020; Chen et al., 2020). LitCovid was developed to house current published literature around COVID-19 and is updated daily to assist researchers and health leaders keep current with the latest research (Chen et al., 2020). The CORD-19 data set contains over 199,000 published literature with over 88,000 full-text articles on both COVID-19 and other historical epidemics (Center for Security and Emerging Technology, 2020).

With the advent of the pandemic, a Kaggle challenge, a data science and machine learning competition, was issued to data scientists to use text mining and natural language processing to data-mine through the CORD-19 data set. The purpose of this challenge was to provide researchers and health leaders with insight into issues and topics regarding the current management of COVID-19. This challenge was a collaborative challenge among the Allen Institute, the White House, the National Institute of Health, and other institutions (Center for Security and Emerging Technology, 2020). Text mining uses natural language processing (NLP) for unstructured data, such as the text of abstracts or full-text papers, to summarize information using various methods (Zengul et al., 2020; Zhao et al., 2018). This body of methodologies provides multiple ways to analyze unstructured data including visualization through word clouds and knowledge maps, and ranking top ckeywords, phrases, and sentences by importance (Abulaish, Parwez, & Jahiruddin, 2019; Garg & Kumar, 2018; Yang et al., 2018; Zhao et al., 2018). NLP is a type of text mining that scans through large amounts of unstructured data to extract meaning and information from the data. NLP methods can include text classification, clustering, and sentiment analysis. Text mining and NLP transforms and characterizes text by using statistical algorithms to provide quality information from unstructured text data (Dreisbach et al., 2019; Zaremba et al., 2009).

Within health care, there have been some demonstrations of the use of these text mining techniques. Using text mining and NLP, NimbleMiner mined through thousands of patient and nurse documentation to identify alcohol and substance abuse clinical notes, enabling users to search through and find alcohol and substance abuse-specific documentation (Topaz et al., 2019). Text mining has also been used to parse and annotate unstructured notes in electronic health records to discover risk factors for patient deterioration (Korach et al., 2020). Further, text mining and NLP techniques have been utilized on literature databases, such as Web of Science, or for specific topics, such as cardiovascular disease, to parse out high-level information from abstracts of the databases or topics (Gal et al., 2019; K. Li et al., 2018).

For nursing, evidence-based practice (EBP) is the standard in developing practice guidelines and is of particular importance within the presence of a pandemic (Melnyk et al., 2009). However, a current challenge for the discipline is that nurses' primary responsibility is often at the bedside. Potential time constraints and competing priorities are barriers to the nurses' ability to do robust literature gathering and appraisals of evidence to implement EBP to address issues presented with the COVID-19 pandemic. Text mining and NLP techniques could help remove this barrier by providing high-level summaries of published literature about nursing topics around the historic epidemics and the COVID-19 pandemic. These summaries could assist in readily evaluating the applicability and meaningfulness of literature to a nurses' clinical question potentially improving access to literature reviews. The question was asked: "What if there was a tool that provided nurses access to high-level information on a particular topic within published literature applicable to nursing around historical epidemics and the COVID-19 pandemic?" In response to this question, a nurse scientist at a children's hospital in the western United States created this resource that housed nursing-specific literature around pandemics and epidemics. The intent of this resource was to provide a literature repository for nurses that houses abstracts with links to literature and provides preliminary assistance in accessing and visualizing the information through text mining and NLP. The repository can be accessed freely at this link: childrenscolorado.shinyapps.io/RN_COVID_Lit/. This repository has been shared with the University of Colorado, College of Nursing, Case Western Reserve University College of Nursing, the American Nurses Association and others.

The aims of this paper are as follows: (a) describe the development, maintenance, and function of the Nursing COVID and Historical Epidemic Literature Repository (NCHELR) and (b) provide a case that demonstrates the use of the repository. For the use case, an algorithm called TextRank was utilized to provide high-level summaries for the abstracts housed in the repository.

Methods

Data Sources

The repository utilizes the CORD-19 and the LitCovid literature data sets. The CORD-19 contains literature around historical epidemics such as Severe Acute Respiratory Syndrome, Middle East Respiratory Syndrome, and Ebola, providing a breadth of information useful for nursing that crosses several countries and continents (Center for Security and Emerging Technology, 2020). LitCovid contains current COVID-19 literature providing the latest published literature around the pandemic (Chen et al., 2020). As there are over 230,000 published papers between these two repositories, it was essential to restrict the abstracts and corresponding links to the full-text papers contained within the NCHELR to relevant nursing topics. Keywords "nursing" or "nurse" were employed on both data sets as a broad approach in selecting topics relevant to

nursing. CORD-19 was primarily used for historical epidemic literature, while LitCovid was utilized for the latest COVID-19 literature.

NCHELR Development

Initial development of the NCHELR began in March 2020. For development, a team was formed to mine through the CORD-19 and LitCovid databases. Using "nurse" and "nursing" as keywords, the abstracts of published literature and corresponding links to the full-text articles were collected to be placed in the repository. A primary goal of the NCHELR was to provide a database of nursing literature abstracts; thus, all published articles regarding nurses or nursing care and regardless of type (i.e., quantitative, qualitative) were included. An experienced medical librarian and two nurse scientists collected the literature from these databases. They reviewed the abstracts to ensure that the literature was related to nursing and that links to full-text articles were present. They characterized the literature by population (adult vs. children) and by COVID or historical epidemic. Then, a research nurse and a research assistant ensured that the following characteristics were collected: title, author, year, complete abstract, and doi or link to the article. Lastly, the NCHELR developer, whom is a nurse scientist, examined the dataset to remove duplicates and to ensure completeness of the dataset. Based on completeness of the data or applicability to nursing, some literature was not included in the repository (i.e., some articles included were animal studies).

Once the developer completed the data preparation for each new round of articles, the data fed into the repository application. Multiple text mining and NLP algorithms were implemented on the abstracts to summarize the unstructured data into high-level information. The repository foundation is coded within the R language and utilizes multiple R packages within ShinyApps as the web-based application for user access (R Core Team, 2017). Figure 1 summarized this process.

The first version of the repository was deployed at the end of March 2020. This version contained only 49 abstracts of published articles with corresponding links to the full-text articles, and a few text mining and NLP algorithms. A beta test was conducted using nurse scientist colleagues employed at various organizations, nurse leaders, a research assistant, and a biostatistician to assess usefulness (how useful is the repository?), meaningfulness (how meaningful is this repository?), and communicability (how well does the repository communicate the information?). Overall, feedback from this round of beta test was generally positive on all three questions. One leader and a clinical education specialist commented that the current repository was accessible for researchers but not necessarily for nurse leaders. She indicated that it would be useful to have a summary of the information and a table that contained the links to the full article. These requested changes were incorporated into the next version of NCHELR. Another beta test was conducted with the same group using the same questions; feedback was positive for this round in which testers found the changes valuable.

NCHELR Overview

The NCHELR is divided into seven different pages that incorporate different algorithms to summarize the information. Supplementary material A provides screenshots of each page. Table 1 provides a brief description of the algorithm and an example of using that algorithm within published literature across different fields. Further, the NCHELR is searchable through sub-selection of the population or epi- and/or pandemic. There is also a Boolean search term option to further sub-select the literature by a specific word. The search term option mines through the titles and the abstracts for that specific word. Within the repository, the papers containing that search term in the title or abstract and/or subselection of the categories are collected in a dataset. The algorithms then utilize the abstracts of these papers to provide a high-level summary.

The second page of the repository application contains a table of the abstracts as defined by search parameters. The table lists the subselection of papers with their authors, title, year, summary sentence, and doi or link to the full-text articles. Further, an overall summary consisting of the five most important sentences from the selection of papers is displayed, providing a summary of the selection of papers. The TextRank algorithm is utilized to provide this summary (Korach et al., 2020). This high-level summary provides the end-user a snapshot of the top entries and a short summary of each abstract, with corresponding links to the full-text articles, allowing end-users the ability to select literature they would like to further investigate and appraise. The third page of the repository application groups "similar" papers into clusters. A clustering algorithm involving "Euclidean" distance between words across the literature is implemented across the selection of abstracts and group papers into three clusters (Nakagawa et al., 2012). This technique allows the user to further sub-select the papers into similar clusters, potentially enabling more intentional digestion of information. Further, a table, similar to what is found on the landing page of the repository application, is displayed to deliver information on the select papers for each cluster. The utility of this function is that the user has the information and links for papers that are similar to each other potentially reducing time in browsing papers.

Word frequencies and word clouds summarize the selection of papers on the fourth page of the repository application. Word frequencies are simple counts of words arranged by most common, while word clouds are visual representations of the most common words as determined by word frequencies (Berlanga et al., 2012; Cidell, 2010). The word frequencies bar plot is

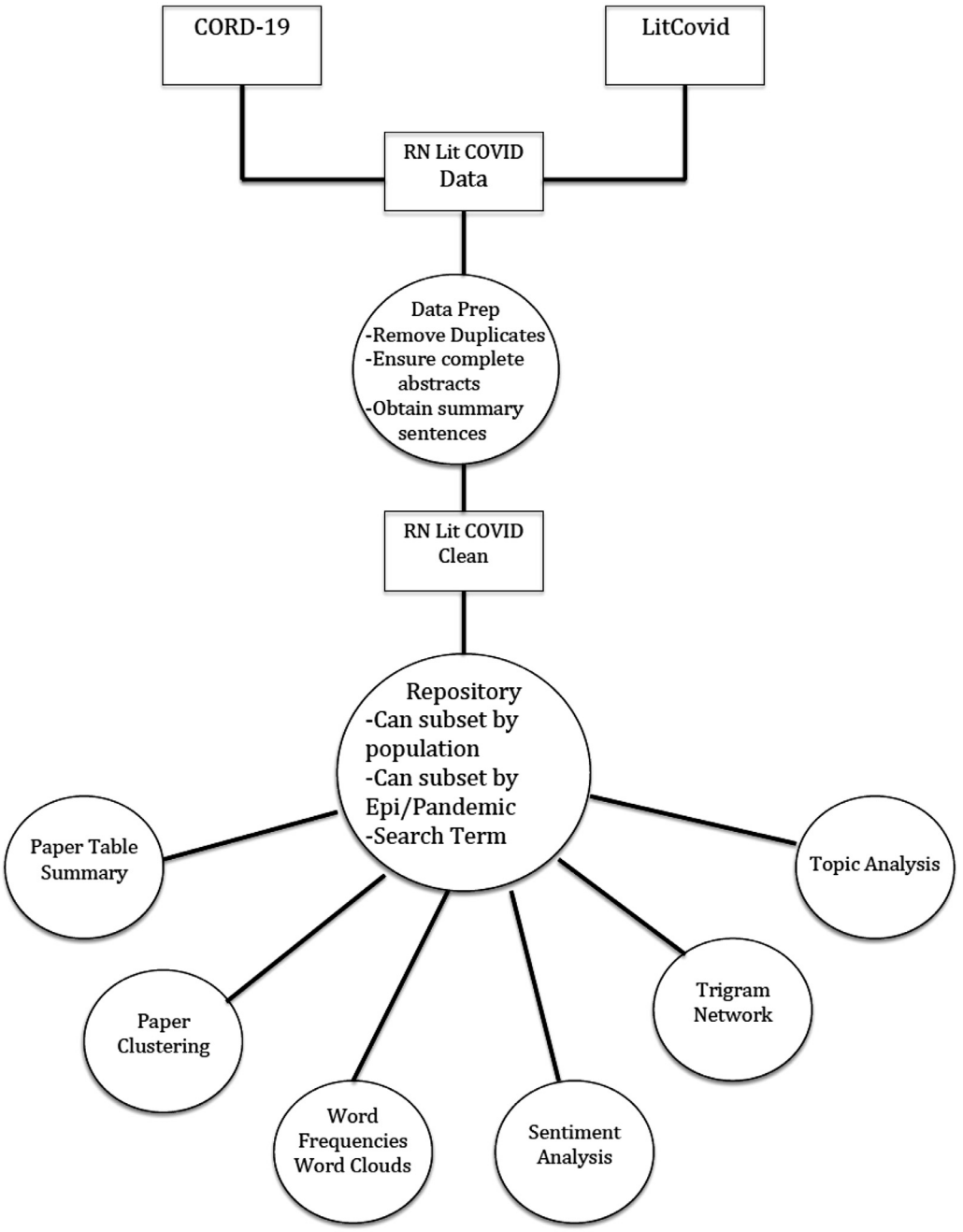

Figure 1 – NCHELR Process Map-Describes the processes for extracting the literature from the CORD-19 and LitCovid data sets, for cleaning the data, and for how the algorithms are utilized.

restricted to the top 10 most common words, while the word cloud is restricted to the top 100 words. Word clouds provide the end-user words that are emphasized across the select abstracts and help to visualize key words that can be used for additional searches and can provide a global sense of the text across abstracts. The fifth page of the repository application utilizes sentiment analysis (an algorithmic methodology that assigns words as positive or negative) to display the top 10 most common positive words and negative words (Dreisbach et al., 2019). This page provides end-users a sense of positive or negative topics within the selection of papers.

The co-occurrence network of the most common trigrams (groups of three words) from selected papers is displayed on the sixth page of the repository application (Chen & Luo, 2019). This page provides the end-user a sense of how the words are related through the co-occurrence network, adding some summative information to the selected papers. The last page, page seven, clusters words from the selected abstracts into three topics. Latent Dirichlet Allocation is an algorithm that utilizes the probability to associate words into the three topics (Li et al., 2019). These topics are the themes for the selected papers providing the end-user a sense of the themes being discussed within the

Table 1 – Text Mining Algorithms Utilized in the Nurse COVID and Historic Epidemic Literature Repository

Algorithm	Brief Description	Examples of Use
TextRank	-Network based algorithm that identifies important sentences and keywords	-Determine factors from EHR notes associated with patient deterioration (Korach et al., 2020).
Cluster analysis – Euclidean distance	-Algorithm that cluster papers based on Euclidean distance of the words contained within a selection of text	-Assessing the impact of an off-campus program for physical therapy students (Nakagawa et al., 2012)
Word frequencies/word clouds	-Counts the most frequent words used in a selection of text -Visualizes those counts with more emphatic words having higher counts	-Exploratory qualitative data analysis for text (Cidell, 2010) -Formative feedback for textual assignments for learners (Berlanga et al., 2012)
Sentiment analysis	-Categorizes and assigns positive or negative sentiments to words	-Describing patient symptoms from electronic patient-authored text data (Dreisbach, 2019)
Trigrams co-occurrence networks	-Relating words based on co-occurrences between a selection of text	-Knowledge map for engineering literature (Chen & Luo, 2019)
Latent Direchlet allocation – Topic analysis	-Utilizes probability to associate words to a predefined number of topics	-Summarizing personal goals among patients undergoing bladder cancer (Li et al., 2019)

selected papers. All of these features provide nurse end-users fast guidance on which full-text articles to further explore based on their topic of interest.

NCHELR Example

As a use case of the NCHELR, we provide an (a) overall summary of all the abstracts of literature housed within the repository; (b) by COVID-19 papers; and (c) by historical epidemics. We currently use a TextRank, an extractive and text summarization technique, to generate this summary. TextRank is a graph-based ranking algorithm that can be applied to a variety of natural language processing applications. Derived from Google PageRank and other graph-based algorithms, it provides a mechanism to rank sentences by importance (Korach et al., 2020; Mihalcea & Tarau, 2004). Graph-based ranking algorithms decide the importance of a vertex within a graph based on global information recursively drawn from the entire graph. For example, a vertex that links to another will cast a vote giving that vertex a score. This score is subsequently ranked according to importance. Text is used to calculate and link related sentences and then rank them based on importance (Mihalcea & Tarau, 2004). Thus, the recommendations made by TextRank provide a sensible summary of a selection of abstracts. We looked at the top five most important sentences for each group to assess an overall summary of the topics discussed.

NCHELR Maintenance

COVID-19 literature is constantly produced and published; therefore, a maintenance strategy was implemented to keep the NCHELR updated with the growing literature. The maintenance strategy included a team of the medical librarian and the nurse scientist developer. The structure developed to create the repository is also utilized for maintenance (see Figure 1). The medical librarian utilizes the LitCOVID database on a biweekly basis to collect recently published abstracts and links to the full-text papers. The nurse scientist developer ensures the completeness of the papers and updates the repository. The NCHELR is updated monthly.

Findings

A total of 760 published papers related to nursing are housed in the repository as of July 2020. COVID-19 published papers numbered approximately 511, while published paper classified as historical numbered 302. There were 56 papers that data extractors considered as applicable to COVID-19 and historical categories.

Table 2 summarizes the results of the TextRank analysis. When examining all 760 papers within the repository, important sentences emphasized topics germane to nursing, such as the psychological state of nurses, the necessity of rapid continuing education, and health care delivery changes (Badahdah et al., 2020; Dickerson & Graebe, 2020; Pahuja & Wojcikewych, 2020; Rabb, 2020; Zhang et al., 2020). For the COVID-19 papers, the emphasis was on the changes in health care delivery, specifically for vulnerable populations and utilization of telehealth (Pahuja & Wojcikewych, 2020; Rabb, 2020; Vendekerckhove et al., 2020). Further, nurses' psychological concerns were the first and fifth most important sentences in this selection of papers (Que et al., 2020; Zhang et al., 2020). For the historical papers, topics highlighted mostly discussed training and adherence to respiratory guidelines, specifically relating to the H1N1 outbreak (Choi & Kim, 2018; Lam & Hung, 2013; Maroldi et al., 2017; Martel et al., 2013; Turnberg et al., 2008). By having the abstracts of these articles in one location, nurse leaders and clinicians can further select articles and seek full text articles to inform practice.

Table 2 – Five Most Important Sentences from the Nurse COVID Literature Repository by Category

All Papers	COVID Papers	Historic Papers
"Understanding nurses' psychological change process during the care for patients with COVID-19 is imperative for healthcare leaders" (Zhang, 2020)	"Understanding nurses' psychological change process during the care for patients with COVID-19 is imperative for healthcare leaders" (Zhang, 2020)	"Our objective was to determine the compliance with respiratory hygiene of triage nurses at 2 university hospital centers and to identify factors influencing compliance to the respiratory hygiene principles of emergency health care workers" (Martel et al., 2013)
"The current COVID-19 pandemic has affected every one, but presents profound consequences for patients with kidney disease, health care providers, and biomedical researchers (Rab, 2020)	"The current COVID-19 pandemic has affected every one, but presents profound consequences for patients with kidney disease, health care providers, and biomedical researchers (Rab, 2020)	"Conclusion: The study points out the need to provide in-service training for professionals on the transmission of microorganism in primary health care to ensure adequate level of risk perception and knowledge" (Maroldi et al., 2017)
"The COVID-19 pandemic has created the need for rapid development and implementation of nursing continuing professional development (NCPD) to scale up nurses and other health care providers to meet a surge in critically ill patients" (Dickerson & Graebe, 2020)	"The novel coronavirus SARS-COV-2 (COVID-19) pandemic is changing how we deliver expert palliative care" (Pahuja & Wojcikewych, 2020)	"Introduction: The primary aim of this study was to explore the perception of Hong Kong emergency nurses regarding their work during the human swine influenza pandemic outbreak" (Lam & Hung, 2013)
"The novel coronavirus SARS-COV-2 (COVID-19) pandemic is changing how we deliver expert palliative care" (Pahuja & Wojcikewych, 2020)	"During the COVID-19 pandemic, cardiologists try to minimize the risk for their patients by using telehealth to provide continuing care" (Vandekerckhove, 2020)	"Methods: The study examined health care worker adherence to CDC recommended respiratory infection control practices in primary care clinics and emergency departments of 5 medical centers in King County, Washington, using a self-administered questionnaire" (Turnberg, 2008)
"Results: The study revealed a high prevalence of stress, anxiety, and poor psychological well-being, especially among females, young health care workers, and those who interacted with known or suspected COVID-19 patients" (Badahdah et al., 2020)	"Conclusions: Psychological problems are pervasive among healthcare workers during the COVID-19 pandemic" (Que et al., 2020)	Therefore, this study was conducted to identify nursing students' knowledge, attitudes, practices, and risk perceptions of infection prevention related to occupational exposure to Zika virus infection, and to identify correlations among the related variables" (Choi & Kim, 2018)

Discussion and Recommendations

This paper describes the development and function of the Nurse COVID and Historical Epidemic Literature Repository. The repository is a tool to link frontline nurses, nurse researchers, and nurse leaders to evidence and information applicable to nursing during their initial stages of investigation in a less time intense way. The NCHELR, with its use of text mining algorithms and NLP, can provide nurses access to published papers to help make immediate evidenced-based decisions in the field. Text mining and NLP have been known to provide valuable knowledge and information useful for an organization (Chen & Luo, 2019).

There are limitations to text mining and NLP methods. As numerous algorithms could be used, the algorithms' efficacy can be called into question (Dreisbach et al., 2019). Further, this repository should not be used as a replacement for a proper appraisal of the published papers as outlined by the steps of EBP (Fineout-Overhold et al., 2010). It should provide some guidance on what is discussed within the selection of papers and lower the barriers of literature reviews during the pandemic.

Another limitation worth noting is that abstracts are utilized for the algorithms. Abstracts widely vary in quality, formats, and communicating methods and findings embedded within the full text article. Therefore, poor abstracts will result in poor results of the algorithms; however, abstracts were chosen for the following reasons. First, abstracts of full text articles are freely available without having to pay a fee. The development of this application was done in a hospital-setting and not an academic institution; thus, the in-house library has limited access to full text papers. Second, the computational resource was more manageable for abstracts compared to full text articles. There are several ways to mitigate this computational resource, such as implementing parallel computing

within the code; however, the authors felt that this repository needed to be developed quickly to respond to the need for access to COVID-19 literature, and implementing parallel computing would have delayed the development. Though this limitation may be a disadvantage to the validity of the results of the algorithms, the intent of the repository was first and foremost an initial tool for end-users to use the results of the algorithms to narrow down COVID-19 literature specific to their clinical question.

As an application example of the repository, the TextRank algorithm was utilized on all the literature housed within the repository to provide a high-level summary. It was found that frontline nurses' psychological status is of utmost relevance (Badahdah et al., 2020; Que et al., 2020; Zhang et al., 2020). Leaders should be aware of assessing their staff's resiliency and the psychological impact of a pandemic to provide support resources for their staff. Researchers could further study the impact of COVID-19 on the nursing workforce, particularly their long-term psychological status. Further, high-impact rapid education is necessary for nurses' safety for both COVID-19 and historical epidemics (Dickerson & Graebe, 2020; Maroldi et al., 2017). Investing in these educational resources could potentially provide better management of the pandemic. Anecdotally, these conclusions are sensible and are readily noted. However, the algorithms do dictate that these are a few of the most important topics for nursing.

To our knowledge, this is the only literature repository that utilizes text mining and NLP algorithms on published literature related to nursing and specifically for frontline nurses, nurse researchers, and nurse leaders. This repository intends to be a tool to be used by nurses to gain knowledge specific to COVID-19 and historical epidemics in an easily accessible way. The advantages of this repository compared to other databases are as follows: (a) this a repository-specific to nursing literature; and (b) it utilizes text mining and NLP to provide high-level summaries. This link, childrenscolorado.shinyapps.io/RN_COVID_Lit/, provides access to the repository.

Authors' Contributions

Figaro Loresto: Conceptualization, Methodology, Software, Formal Analysis, Visualization, Writing-Original Draft, Writing-Review & Editing; **Lisa Nunez**: Writing-Original Draft, Writing-Review & Editing; **Lindsey Tarasenko**: Conceptualization, Investigation, Data Curation, Writing-Review & Editing; **Marie St. Pierre:** Investigation, Data Curation, Writing-Review & Editing; **Kenneth Oja:** Investigation, Data Curation, Writing-Review & Editing; **Mallory Mueller:** Conceptualization, Writing-Review & Editing; **Bailey Switzer:** Investigation, Data Curation,; **Katherine Marroquin:** Investigation, Data Curation; **Catherine Kleiner:** Conceptualization, Resources, Supervision, Writing-Review & Editing.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.outlook.2020.12.017.

REFERENCES

Abulaish, M., Parwez, M. A., & Jahiruddin (2019). DiseasE: A biomedical text analytics system for disease symptom extraction and characterization. *Journal of Biomedical Informatics, 100*, 1–15.

Badahdah, A., Khamis, F., Mahyijari, N. A., Balushi, M. A., Hatmi, H. A., Salmi, I. A., . . ., Noomani, J. A. (2020). The mental health of health care workers in Oman during the COVID-19 pandemic. *International Journal of Social Psychiatry*, doi:10.1177/0020764020939596.

Berlanga, A. J., van Rosmalen, P., Boshuizen, H. P. A., & Sleop, P. B (2012). Exploring formative feedback on textual assignments with help of automatically created visual representations. *Journal of Computer Assisted Learning, 28*, 146–160.

Center for Security and Emerging Technology. (2020). *COVID-19 open research dataset (CORD-19)*. *Kaggle*. Retrieved from cset.georgetown.edu/research/covid-19-open-research-dataset-cord-19/. (Accessed December 2, 2020).

Chen, H., & Luo, X. (2019). An automatic literature knowledge graph and reasoning network modeling framework based on ontology and natural language processing. *Advanced Engineering Informatics, 42*, 1–17.

Chen, Q., Allot, A., & Lu, Z. (2020). Keep up with the latest coronavirus research. *Nature, 579*(7798), 193, doi:10.1038/d41586-020-00694-1.

Choi, J. S., & Kim, K. K. (2018). Infection-control knowledge, attitude, practice, and risk perception of occupational exposure to Zika virus among nursing students in Korea: A cross-sectional survey. *Journal of Infection and Public Health, 11*(6), 840–844, doi:10.1016/j.jiph.2018.07.002.

Cidell, J. (2010). Content clouds as exploratory qualitative data analysis. *Area, 42*(4), 514–523.

Dickerson, P., & Graebe, J. (2020). Nursing continuing professional development: A paradigm shift. *Journal of Continuing Education in Nursing, 51*(7), 297–299, doi:10.3928/00220124-20200611-02.

Dreisbach, C., Koleck, T. A., Bourne, P. E., & Bakken, S. (2019). A systematic review of natural language processing and text mining of symptoms from electronic patient-authored text data. *International Journal of Medical Informatics, 125*, 37–46.

Fineout-Overhold, E., Melnyk, B. M., Stillwell, S. B., & Williamson, K. M. (2010). Evidence-based practice stepy by step: Critical appraisal for the evidence: part I. *American Journal of Nursing, 110*(7), 47–52, doi:10.1097/01.NAJ.0000383935.22721.9c.

Gal, D., Thijs, B., Glanzel, W., & Sipido, K. R. (2019). Hot topics and trends in cardiovascular research. *European Heart Journal, 40*, 2363–2374.

Garg, M., & Kumar, M. (2018). The structure of word co-occurrence network for microblogs. *Physica A, 512*, 698–720.

Korach, Z. T., Yang, J., Rossetti, S. C., Cato, K. D., Kang, M. J., Knaplund, C., . . ., Zhou, L. (2020). Mining

clinical phrases from nursing notes to discover risk factors of patient deterioration. *International Journal of Medical Informatics, 135*(104053), 1–8.

Lam, K. K., & Hung, S. Y. M. (2013). Perceptions of emergency nurses during the human swine influenza outbreak: A qualitative study. *International Emergency Nursing, 21*(4), 240–246, doi:10.1016/j.ienj.2012.08.008.

Li, K., Rollins, J., & Yan, E. (2018). Web of Science use in published research and review papers 1997-2017: A selective, dynamic, cross-domain, content-based analysis. *Scientometrics, 115*, 1–20.

Li, Y., Rapkin, B., Atkinson, T. M., Schofield, E., & Bochner, B. H. (2019). Leveraging Latent Dirichlet Allocation in processing free-text personal goals among patients undergoing bladder cancer surgery. *Qual Life Res, 28*(6), 1441–1455.

Maroldi, M. A. C., da Silva Felix, A. M., Lima Dias, A. A., Kawagoe, J. Y., Padoveze, M. C., Ferreira, S. A., . . ., Figueiredo, R. M. (2017). Adherence to precautions for preventing the transmission of microorganisms in primary health care: A qualitative study. *BMC nursing, 16*(49), doi:10.1186/s12912-017-0245-z.

Martel, J., Bui-Xuan, E. F., Carreau, A. M., Carrier, J. D., Larkin, E., Vlachos-Mayer, H., & Dumas, M. E. (2013). Respiratory hygiene in emergency departments: Compliance, beliefs, and percpetions. *American Journal of Infection Control, 41*(1), 14–18, doi:10.1016/j.ajic.2011.12.019.

Melnyk, B. M., Fineout-Overhold, E., Stillwell, S. B., & Williamson, K. M. (2009). Evidenced-based practice: step by step: igniting a spirit of inquiry: An essential foundation for evidence-based practice. *American Journal of Nursing, 109*(11), 49–52, doi:10.1097/01.NAJ.0000363354.53883.58.

Mihalcea, R., & Tarau, P. (2004). Textrank: Bringing order into text. *Proceedings of the Conference on Empirical Methods in Natural Langauge Processing, 2004*, 404–414.

Nakagawa, K., Yamada, K., Asakawa, Y., & Yamaguchi, H. (2012). What kinds of impressions did physical therapy students receive through participation in off-campus classes? An analysis using text-mining. *Journal of Physical Therapy Science, 24*, 1063–1068.

Pahuja, M., & Wojcikewych, D. (2020). Systems barrier to assessment and treatment of COVID-19 positive patients at the end of life. *Journal of Palliative Medicine*, doi:10.1089/jpm.2020.0190.

Que, J., Shi, L., Deng, J., Liu, J., Zhang, L., Wu, S., . . ., Lu, L. (2020). Psychological impact of the COVID-19 pandemic on healthcare workers: A cross-sectional study in China. *General Psychiatry, 33* e100259, doi:10.1136/gpsych-2020-100259 *2020*.

R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Core Team. Retrieved from https://www.R-project.org. (Accessed August 3, 2020).

Rabb, H. (2020). Kidney diseases in the time of COVID-19: Major challenges to patient care. *Journal of Clinical Investigation, 130*(6), 2749–2751, doi:10.1172/JCI138871.

Topaz, M., Murga, L., Bar-Bachar, O., Cato, K., & Collins, S. (2019). Extracting alcohol and substance abuse status from clinical notes: The added value of nursing data. *MEDINFO*, doi:10.3233/SHT1190386 *2019*.

Turnberg, W., Daniell, W., Seixas, N., Simpson, T., Van Buren, J., Lipkin, E., & Duchin, J. (2008). Appraisal of recommended respiratory infection control practices in primary care and emergency department settings. *American Journal of Infection Control, 36*(4), 268–275, doi:10.1016/j.ajic.2007.08.004.

Vendekerckhove, P., Vandekerckhove, Y., Tavernier, R., Jaegher, K. D., & de Mul, M. (2020). Leveraging user experience to improve video consultations in a cardiology practice duirng the covid-19 pandemic: Initial insights. *Journal of Medical Internet Research, 22*(6), e19771, doi:10.2196/19771.

Yang, L., Li, K., & Huang, H. (2018). A new network model for extracting text keywords. *Scientometrics, 116*, 339–361, doi:10.1007/s11192-018-2743-5.

Zaremba, S., Ramos-Santacruz, M., Hampton, T., Shetty, P., Fedorko, J., Whitmore, J., . . ., Pot, D. (2009). Text-mining of PubMed abstracts by natural language processing to create a public knowledge base on molecular mechanisms of bacterial enteropathogens. *BMC Bioinformatics, 10*(177), 1–11.

Zengul, F. D., Lee, T., Delen, D., Almehmi, A., Ivankova, N. V., Mehta, T., & Topuz, K. (2020). Research themes and trends in ten top-ranked nephrology journals: A text mining analysis. *American Journal of Nephrology, 51*, 147–159.

Zhang, Y., Wei, L., Li, H., Pan, Y., Wang, J., Li, Q., . . ., Wei, H. (2020). The psychological change process of frontline nurses caring for patients with COVID-19 during its outbreak. *Issues in Mental Health Nursing, 41*(6), 525–530, doi:10.1080/01612840.2020.1752865.

Zhao, Y., Fesharaki, N. J., Liu, H., & Luo, J. (2018). Using data-driven sublanguage pattern mining to induce knowledge models: Application in medical image reports knowledge representation. *BMC Medical Informatics and Decision Making, 18*(61), 1–15.