Available online at www.sciencedirect.com

ScienceDirect

journal homepage: http://www.elsevier.com/locate/kontakt

ELSEVIER

Original research article

Relationships among interventions and health literacy outcomes for sub-populations: A data-driven approach

*Martin Michalowski [a, *], Robin R. Austin [a], Michelle A. Mathiason [a], Sasank Maganti [b], Erica Schorr [a], Karen A. Monsen [a]*

[a] *University of Minnesota School of Nursing, Minneapolis, MN, USA*
[b] *University of Minnesota Computer Science and Engineering, Minneapolis, MN, USA*

ARTICLE INFO

Received: 2018-07-31
Accepted: 2018-11-11
Published online: 2018-11-30

Keywords:
Statistical learning
Knowledge
Health literacy
Nursing interventions
Omaha System

ABSTRACT

Objectives: The goals of this study were to examine relationships among health literacy and outcomes for sub-populations identified within a large, multi-dimensional Omaha System dataset. Specific aims were to extract sub-populations from the data using Latent Class Analysis (LCA); and quantify the change in knowledge score from pre- to post-intervention for common sub-populations.
Design: Data-driven retrospective study using statistical modeling methods.
Sample: A set of admission and discharge cases, captured in the Omaha System, representing 65,468 cases from various health care providers.
Measures: Demographic information and the Omaha System terms including problems, signs/symptoms, and interventions were used as the features describing cases used for this study. Development of a mapping of demographics across health care systems enabled the integration of data from these different systems.
Results: Knowledge scores increased for all five sub-populations identified by latent class analysis. Effect sizes of interventions related to health literacy outcomes varied from low to high, with the greatest effect size in populations of young at-risk adults. The most significant knowledge gains were seen for problems including Pregnancy, Postpartum, Family planning, Mental health, and Substance use.
Conclusions: This is the first study to demonstrate positive relationships between interventions and health literacy outcomes for a very large sample. A deeper analysis of the results, focusing on specific problems and relevant interventions and their impact on health literacy is required to guide resource allocation in community-based care. As such, future work will focus on determining correlations between interventions for specific problems and knowledge change post-intervention.

* **Author for correspondence:** Martin Michalowski, PhD, University of Minnesota School of Nursing, 5-140 Weaver-Densford Hall, 308 Harvard Street SE, Minneapolis, Minnesota 55455, USA; e-mail: martinm@umn.edu
http://dx.doi.org/10.1016/j.kontakt.2018.10.009

KONTAKT XX/4: 345–352 • ISSN 1212-4117 (Print) • ISSN 1804-7122 (Online)

Please cite the article: Michalowski M, Austin RR, Mathiason MA, Maganti S, Schorr E, Monsen KA. Relationships among interventions and health literacy outcomes for sub-populations: A data-driven approach. Kontakt 2018; 20(4): e319–e325; http://dx.doi.org/10.1016/j.kontakt.2018.10.009

Introduction

The association between health literacy and health disparities is well documented in the literature [1–3]. Today's complex health environment often requires individuals and families to play a more active role in their health. Unfortunately, limited health literacy can be a barrier to achieving optimal health, particularly for disadvantaged populations [3]. Several national organizations and initiatives have asserted the importance of health literacy from the individual level to a broader population health perspective [4–7].

Nurses are well positioned to optimize health literacy and are known to influence population health [2, 8, 9]. In this work, we explore the relationships between nursing interventions and health literacy outcomes for various cases collected during routine documentation in community care settings (e.g. public health nursing and home care physical therapy, or hospice care). Using a data-driven approach we analyzed case data for 65,468 cases from various health care providers.

Background

Approximately 80 million U.S. adults (35% of the population) are thought to have limited health literacy and an increased risk for poor health outcomes [4, 10]. Health literacy is the degree to which individuals have the capacity to obtain, process, and understand the basic health information and services needed to make appropriate health decisions [4]. Limited health literacy rates are higher among older adults, minority, economically challenged, and individuals with less than a high school education [2, 11]. Numerous studies and reports have identified health literacy as a major health barrier and a challenge to achieving optimal health outcomes [11]. Despite increased awareness and numerous national initiatives, more than one third of adults in the United States still demonstrate limited health literacy; for example, those with limited health literacy are less likely to receive preventative care services, and over 70 million people have difficulty following directions on a medication label [10–12]. The estimated annual cost of limited health literacy to the U.S. economy is between $106 billion and $238 billion [13].

Health literacy has been identified as a necessary element to increase health equity and reduce health disparities [2, 7, 9]. The nursing profession is well positioned to help alleviate health inequity by engaging in patient-centered health literacy practices to minimize negative outcomes associated with limited health literacy and improve the skill and abilities needed in complex health systems [7, 9, 14–16]. Public health nursing data sources, using the Omaha System, have been used to highlight nursing interventions provided during home care visits [17–19]. A recent study revealed that home visits by public health nurses can improve health literacy [19].

Omaha System

The Omaha System is the only multi-disciplinary clinical terminology that taxonomically classifies the whole of health concepts and interventions, and also serves to measure health Knowledge, Behavior, and Status [17]. It is one of only 12 standardized terminologies recognized by the American Nurses Association (ANA) [20]. The Omaha System Data Collaborative within the University of Minnesota Center for Nursing Informatics consists of de-identified, precisely coded Omaha System data for a wide range of cases. Chart 1 provides an overview of the Omaha System.

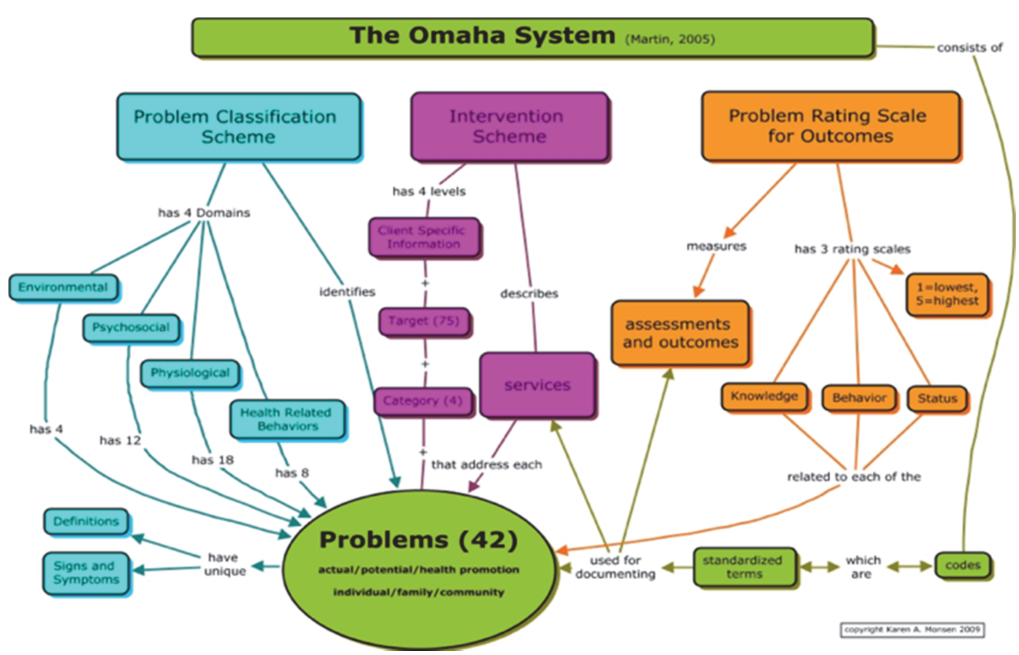

Chart 1 – Overview of the Omaha System, used with permission

Patients may have one or more admissions that consist of one or more visits. Each patient admission is considered a *case* and has demographics, problems, signs/symptoms, interventions, and problem-specific baseline and final rating scores. Each case has one or more interventions. The database structure (Chart 2) uses an entity-relationship (ER) model enabling precise linkages between the variables for optimal sample selection.

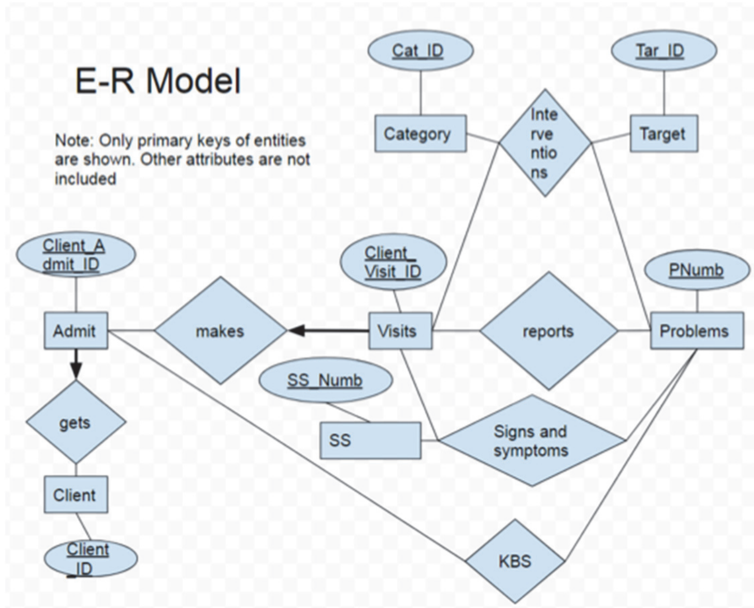

Chart 2 – Omaha System entity – relationship model (© Monsen, 2017)

In the Omaha System Data Collaborative, there are 111,570 patient admissions. Each patient admission has, on average, 19 signs/symptoms. There are a total of 3,268,089 interventions with an average of 29 interventions per patient admission. Of the total sample, 16.1% are aged 65 years and older ($N = 17{,}908$). The most frequent problems for adults aged 65 and older are Health care supervision, Neuro-musculo-skeletal function, Skin, Circulation, Pain, Personal care, Respiration, Nutrition, Urinary function, and Medication regimen.

Materials and methods

The aim of our study was to determine the relationships among interventions and the health literacy of identified sub-populations. Towards this end, our analysis was divided into two main components. The first was a data-driven clustering of cases in our dataset to create sub-populations. After creating these sub-populations we studied the knowledge score changes from pre- to post-intervention for the cases in each sub-population. Both of these steps are described in detail below.

Study design and sample

Omaha System data from a de-identified database for admissions between 2012 and 2017 from multiple sources was evaluated for patterns of health literacy. Cases with health literacy evaluated by Knowledge at admission and discharge were included in analysis. In total, there were 65,468 cases that met these inclusion criteria. Case demographics were evaluated at admission to characterize each sub-population. Knowledge, Behavior and Status (KBS) were summarized across problems at admission and discharge.

Analysis methods

Latent Class Analysis (LCA) was used to determine the best division of the entire sample into sub-populations [21, 22]. The number of classes was determined based on minimizing the Bayesian Information Criterion (BIC) and Akaike's Information Criterion (AIC). From 3 to 10 sub-populations were evaluated, with 5 sub-populations producing low BIC and AIC with high entropy. Each sub-population is a group of cases; and all five sub-populations were tested for differences in demographics, admission and discharge KBS, and number of interventions.

Using the Centers for Disease Control and Prevention's (CDC's) definition for health literacy [24], we qualified those cases with a knowledge score of 1 (no knowledge) and 2 (minimal knowledge) as having *limited* health literacy, and those with a knowledge score of 3 (basic knowledge), 4 (adequate knowledge), and 5 (superior knowledge) as having positive health literacy. Our analytical goal was then two-fold: (1) determine the identified sub-populations' base health literacy level (limited or positive), and (2) determine if health literacy improved post-intervention.

Data analysis

KBS change scores were calculated by taking the difference between the score at initial admission and the last reported score for a given case (note: a case can have multiple admission and discharge events). Average change scores are reported as the total change divided by the number of cases within the sub-population. The change in knowledge for problems that were documented fewer than 100 times within a sub-population was not employed in the analysis. This filtering approach eliminated those problems that were outliers for each discovered sub-population.

Results

Applying LCA to the available data resulted in 5 sub-populations. The number of cases within the sub-populations from 1 to 5 were: 1,041 (sub-population 1), 7,719 (sub-population 2), 16,778 (sub-population 3), 35,497 (sub-population 4), and 4,433 (sub-population 5). The dot plot age distribution of each sub-population is shown in Chart 3, where the *x*-axis indicates age and the *y*-axis is the sub-population number. Circles in each row represent the proportional distribution of the age of cases within each sub-population. As Chart 3 shows, sub-populations 1 and 2 are mainly comprised of cases aged 55 years and older, which is typical of elder home care. The age of sub-populations 3, 4, and 5 skews towards the lower range typical of newborns and young parents.

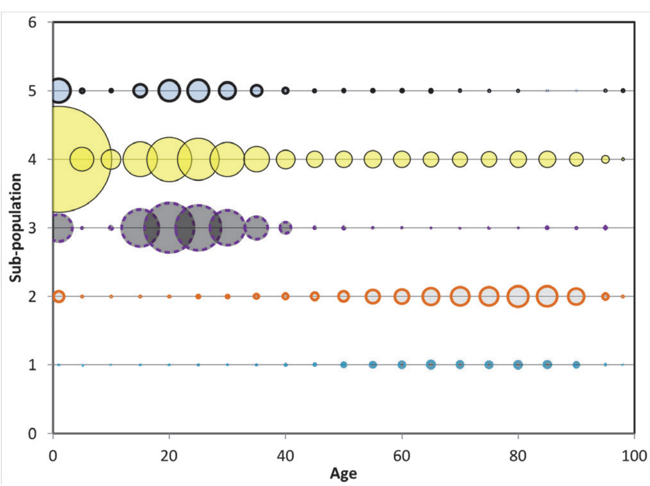

Chart 3 – Age distribution per sub-population

Identified gender was reported as 58.8% and 57.5% female in sub-populations 1 and 2 respectively, while sub-populations 3, 4, and 5 were 94.7%, 61.0%, and 84.5% respectively. The gender distribution correlated with our characterization of sub-populations 1 and 2 as elder care cases. It also allowed us to further break down sub-populations 3, 4, and 5.

Chart 4 presents overall and per sub-population mean knowledge scores at initial admission (Admit) and after discharge (Discharge). Overall health literacy increased, with most of the knowledge gain accounted for in sub-populations 2, 3, and 4. Sub-populations 1 and 5 exhibit a ceiling effect with respect to knowledge gain, as their baseline knowledge scores were already relatively high.

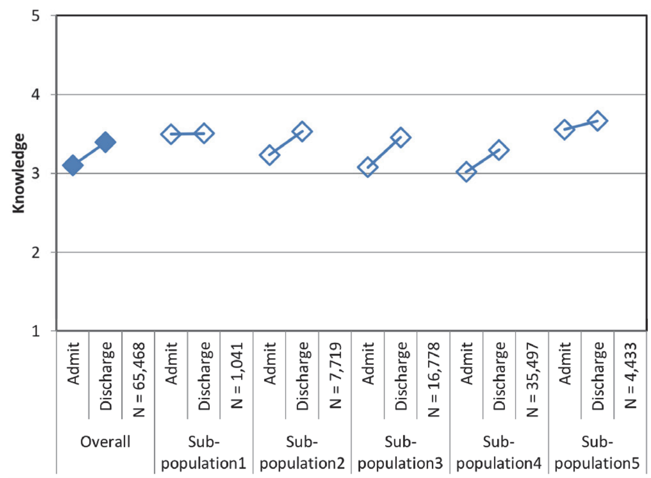

Chart 4 – Mean knowledge scores at admission and last discharge

Given the similarities of sub-populations 1 and 2 (elder care) and sub-populations 3, 4, and 5 (young adults), we considered these two meta-groupings when comparing sub-populations. To delve deeper into the knowledge gain for cases within each sub-population, we plotted the weighted knowledge gain by age as shown in Chart 5 (elder care) and 6 (young adults). The weights were applied as the proportion of cases for a given age within the sub-population with equal weight per sub-population.

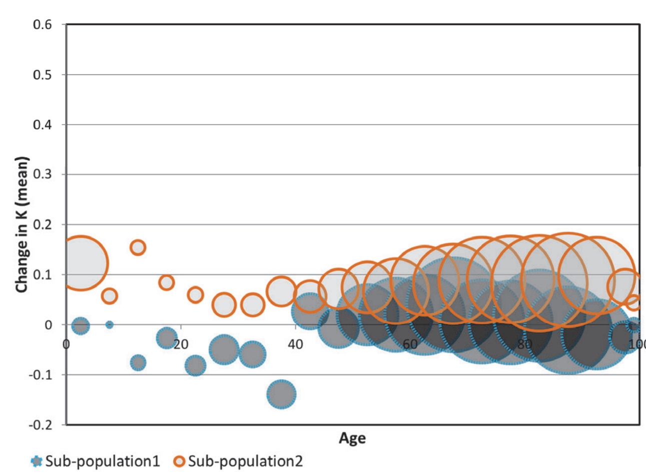

Chart 5 – Weighted mean knowledge gain for elder care sub-populations

For sub-populations 1 and 2, most of the knowledge gain came from the eldest members (Chart 5). For sub-populations 3, 4, and 5, the majority of knowledge gain comes from cases in their 20s for sub-populations 3 and 4, and for newborns in sub-population 5 (Chart 6). The latter is an artifact of the way in which data is collected, where knowledge is scored for parents of newborns but attributed to the newborn case and thus the younger age.

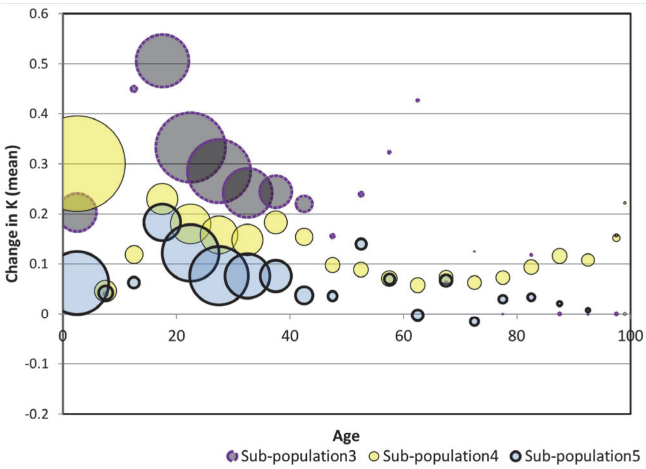

Chart 6 – Weighted mean knowledge gain for young adult sub-populations

To complete the picture of knowledge gain within sub-populations, we studied the knowledge gain for the nine most common problems in the data set. These problems were Income, Mental health, Caretaking/parenting, Growth and development, Pregnancy, Postpartum, Substance use, Family planning, and Health care supervision. The results are reported in Table 1, where green indicates low gain, blue indicates medium gain, and black indicates high gain. While problems associated with a case don't correspond to a diagnosis, they do indicate the case's suspected problems and act as a proxy for the sub-populations.

Results in Table 1 confirm the insignificant to low knowledge gain for the elder care sub-populations (1 and 2). Knowledge gain was predominantly shown for problems related to home care (Caretaking/parenting, Growth and development, and Health care supervision). Sub-populations 3 and 5 represent at-risk young adults, with significant gains seen for problems related to childbirth and family development. Sub-population 4 is primarily made up of newborns, and we see the medium to high knowledge gain for problems related to pregnancy, postpartum, and family development (planning, parenting, and growth and development). As we noted previously, knowledge scores for each case involving a newborn are collected from the mother of the child.

Our final analysis studied the median number of interventions applied to each sub-population. Chart 7 shows that on average, a median number between 100 and 200 interventions were applied to the elder care sub-populations. Sub-population 1, which exhibits an insignificant gain in knowledge has the higher median number of interventions applied to its cases. Sub-population 2, which started at a lower health literary score and showed higher gains, has a lower median number of interventions applied and they are more focused in the 60–80 years old range.

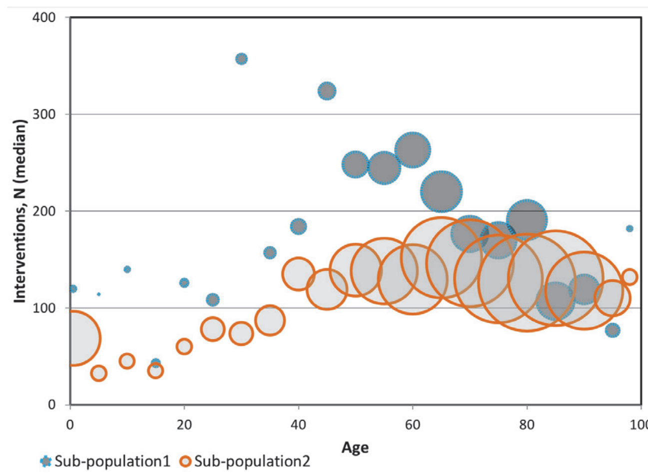

Chart 7 – Weighted median interventions applied within elder care sub-population

Chart 8 demonstrates the median number of interventions applied in cases for Sub-populations 3, 4, and 5. On average, 50 to 120 interventions are applied per case, with the majority of these applied in the 0–35 years old range. Sub-populations 3 and 5 (young at-risk parents) exhibit higher median numbers of interventions when compared to sub-population 4 (newborns). This trend of

Table 1 – Knowledge gain per sub-population for nine of the most frequent problems

Problem	Label	Sub-population 1	Sub-population 2	Sub-population 3	Sub-population 4	Sub-population 5
PB0001	Income	0.03	0.06	0.37[a]	0.28[a]	0.12
PB0012	Mental health	0.05	0.09	0.34[a]	0.32[a]	0.12
PB0014	Caretaking/parenting	0.02	0.25[a]	0.44[b]	0.37[a]	0.20[a]
PB0017	Growth and development	–	0.22[a]	0.53[b]	0.39[a]	0.18
PB0048	Pregnancy	–	–	0.58[b]	0.38[a]	0.44[b]
PB0049	Postpartum	–	–	0.51[b]	0.41[b]	0.24[a]
PB0039	Substance use	0.01	0.03	0.29[a]	0.20[a]	0.09
PB0040	Family planning	–	–	0.64[c]	0.50[b]	0.21[a]
PB0041	Health care supervision	0.01	0.32[a]	0.20[a]	0.25[a]	0.10

Note: [a] indicates low gain; [b] indicates medium gain; [c] indicates high gain.

a lower median number of interventions between sub-populations 3 and 5 and sub-population 4 is consistent across ages, suggesting that the cases in sub-populations 3 and 5 are likely more complex and necessitate more nursing interventions.

When comparing the results presented in Charts 7 and 8 and cross-referencing them with the knowledge gains reported for the top 9 problems in Table 1, we see that the number of interventions does not correlate to knowledge gain. Age and the problem being addressed are better indicators of knowledge gain according to our findings. We discuss this relationship below.

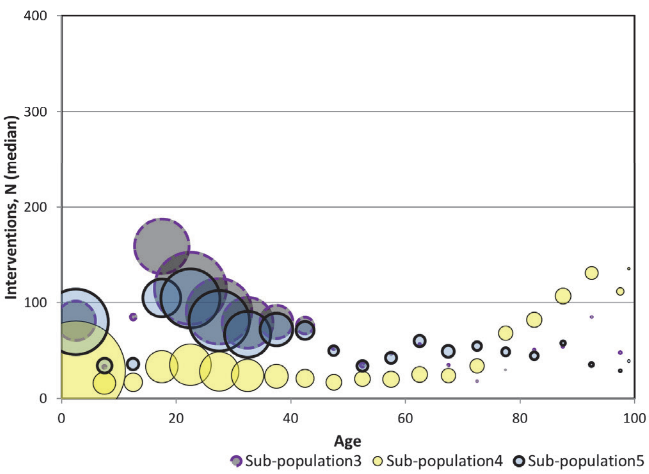

Chart 8 – Weighted median interventions applied within young adult sub-populations

Discussion

Overall and by sub-population, cases served by multi-disciplinary clinicians in community settings exhibited an increase in knowledge with interventions consistent with a low/moderate/high effect [25]. Knowledge followed a similar pattern of change to Behavior and Status, as seen in Chart 9 where the *x*-axis represents the sub-population and the *y*-axis the mean change. Interventions with the most positive relationships between interventions and outcomes are for those applied to the populations with the lowest baseline literacy and with younger cases.

An analysis of relationships among interventions and health literacy outcomes across sub-populations suggests that the most effective interventions are those applied to young at-risk adult sub-populations. This aligns with previous research, using the Omaha System and similar age ranges, which found that public health nursing interventions explained variability in health literacy outcomes [18, 19]. Nurses and the multi-disciplinary health care team play a critical role in promoting health literacy and facilitating patient communication, and understanding health literacy should to be an integral component of every patient's plan of care [9, 15, 18]. Specifically, we found that the high-risk group received or required more interventions to increase

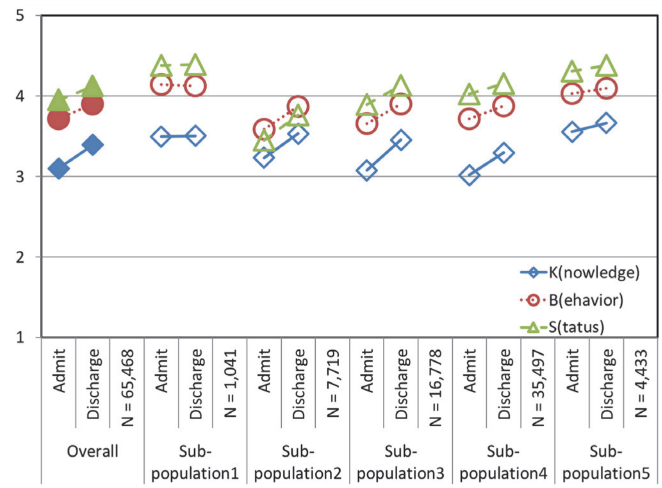

Chart 9 – Change for Knowledge, Behavior, and Status across all sub-populations

knowledge scores [18]. This result is similar to our findings, where sub-population 3 (young at risk adults) had increased weighted mean knowledge scores when compared to the other younger sub-populations (4 and 5).

On average, most cases in this study began with a "basic knowledge" score (3 out of 5) consistent with the definition of health literacy as "the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions" [4]. Current literature suggests a higher level of health literacy knowledge is needed to navigate the health systems, influence patient engagement, and facilitate motivation over one's health [7, 9]. This indicates a knowledge level of 4 (adequate) or 5 (superior) would be ideal as a definition of benchmark attainment in health literacy, vs. the CDC benchmark of 3 – or basic knowledge of adopted a priori for this study [24]. Prospective studies examining health literacy benchmarks should be conducted to validate optimal benchmarks for future research in diverse sub-populations using existing Omaha System Knowledge scores.

Our big data-driven approach to health literacy highlights knowledge gained per problem to improve health literacy in specific groups within the population. Specific problems such as Family planning, Mental health, and Substance use had a positive association with increased knowledge scores. This suggests that individuals with serious healthcare concerns may benefit the most from interventions, and underscores the critical importance of providing sufficient interventions for such individuals [18, 19, 23]. Furthermore, there is potential for nurses and other clinicians to use such data in practice, policy, education, and research.

Limitations of the study were the usual challenges of research using data points in the sample set that were collected from different electronic records and combined together as a single set. During this merging process, some mappings of terms and concepts needed to be done manually, potentially introducing errors. Further limiting the study was that the reported median number of interventions per case was a total of all interventions, independent of the problem addressed. Examining interventions by problem type or domain would facilitate further exploration of specific intervention-outcome relationships for health literacy. Finally, additional demographic information such as marital status and race were not used when extracting classes using LCA. Further research is needed to understand sub-population membership using additional variables in the LCA models.

This is the first study to attempt total-population analysis of health literacy using the Omaha System. An interesting finding was that the five latent classes could be identified by their mean ages and types of problems (1 and 2 vs. 3, 4, and 5). In addition, all LCA-derived sub-populations encompassed cases across the lifespan. Therefore our findings must be considered preliminary, taking into account other factors that may influence results. Future research will be directed towards examining the relationship between overall change in knowledge and change in behavior and status pre- and post-intervention, as well as problem-specific changes. Based on these findings, there is potential to provide valuable insights into the creation and implementation of problem-specific interventions.

Conclusions

This exploratory analysis showed that interventions were positively associated with knowledge uptake, and subsequently improved behavior and status. Our future work will focus on determining correlations between interventions for specific problems and knowledge change post-intervention. Preliminary analysis shows that the highest effect of interventions on improving health literacy occurred for problems such as Pregnancy, Postpartum, Family planning, Mental health, and Substance use. Further analysis will enable a more nuanced look at how nursing interventions affect populations and ultimately lead to improved patient outcomes.

Conflict of interests

The authors declare no conflict of interests regarding this article.

Acknowledgements

The authors would like to acknowledge the Omaha System Partnership and Data Collaborative within the University of Minnesota Center for Nursing Informatics.

REFERENCES

[1] Hudson S, Rikard RV. The Case for Health Literacy Moving from Equality to Liberation; 2018.
[2] Logan RA, Wong WF, Villaire M, Daus G, Parnell TA, Willis E, et al. Health Literacy: A Necessary Element for Achieving Health Equity. Discuss Pap Inst Med 2015: 1–9.
[3] Rudd RE. Improving Americans' Health Literacy. N Engl J Med 2010;363:2283–5. Doi: 10.1056/NEJMp1008755.
[4] Institute of Medicine. Health literacy: A prescription to end confusion. Washington, DC; 2004.
[5] The National Academies of Sciences Engineering and Medicine. Health Literacy: Past, present, and future – Workshop summary; 2015. Doi: 10.17226/21714.
[6] Office of Disease Prevention and Health. National action plan to improve health literacy. Washington, DC Dep Heal Hum Serv 2010;1–73.
[7] Barton AJ, Allen PE, Boyle DK, Loan LA, Stichler JF, Parnell TA. Health Literacy: Essential for a Culture of Health. J Contin Educ Nurs 2018;49:73–8. Doi: 10.3928/00220124-20180116-06.
[8] Baur C. Calling the nation to act: Implementing the national action plan to improve health literacy. Nurs Outlook 2011;59:63–9. Doi: 10.1016/j.outlook.2010.12.003.
[9] Loan LA, Parnell TA, Stichler JF, Boyle DK, Allen P, VanFosson CA, et al. Call for action: Nurses must play a critical role to enhance health literacy. Nurs Outlook 2018;66:97–100. Doi: 10.1016/j.outlook.2017.11.003.
[10] Kutner M, Greenberg E, Jin Y, Paulsen C. Health literacy: A review. Vol. 22; 2006. Doi: 10.1592/phco.22.5.282.33191.
[11] Berkman N, Sheridan SL, Donahue KE, Halpern DJ. Annals of Internal Medicine Review Low Health Literacy and Health Outcomes: An Updated 2014;155.
[12] Bennett I, Chen J. The contribution of health literacy to disparities in self-rated health status and preventive health behaviors in older adults. Ann Fam Med 2009;7:204–11. Doi: 10.1370/afm.940.Department.
[13] Vernon JA, Trujillo A, Rosenbaum SJ, DeBuono B. Low health literacy: implications for national health policy. Hsrc Himmelfarb Gwu Edu 2007:1–18.
[14] Baur C. Calling the nation to act: Implementing the national action plan to improve health literacy. Nurs Outlook 2011;59:63–9. Doi: 10.1016/j.outlook.2010.12.003.
[15] Speros CI. Promoting Health Literacy: A Nursing Imperative. Nurs Clin North Am 2011;46:321–33. Doi: 10.1016/j.cnur.2011.05.007.
[16] Barrett SE, Puryear JS, Westpheling K. Health Literacy Practices in Primary Care Settings : Examples From the Field 2008;36.
[17] Martin K. The Omaha System: A Key to Practice. Documentation and Information Management. Reprinted 2nd Ed. Omaha NE: Health Connections Press; 2005.
[18] Monsen KA, Peterson JA, Mathiason MA, Kim E, Lee S, Chi C-L, Pieczkiewicz DS. Data Visualization Techniques to Showcase Nursing Care Quality. Comput Inform Nurs 2015;33(10):417–26. Doi: 10.1097/CIN.0000000000000190.
[19] Monsen KA, Peterson JJ, Mathiason MA, Kim E, Votava B, Pieczkiewicz DS. Discovering Public Health Nurse-Specific Family Home Visiting Intervention Patterns Using Visualization Techniques. WJNR 2017;39(1):127–46. Doi: 10.1177/0193945916679663.
[20] American Nurses Association. ANA Recognized Terminologies and Data Element Sets. [online] [cit.

2018-09-26]. Available from: http://ana.nursingworld.org/npii/terminologies.html
[21] [Software] PL & PL (Version 1.3.2). No Title 2015.
[22] Lanza ST, Dziak JJ, Huang L, Wagner A, Collins LM. PROC LCA & PROC LTA users' guide (Version 1.3.2); 2015.
[23] Monsen KA, Chatterjee SB, Timm JE, Kay Poulsen J, McNaughton DB. Factors explaining variability in health literacy outcomes of public health nursing clients. Public Health Nurs 2015;32(2):94–100.
[24] Centers for Disease Control and Prevention. "What is Health Literacy?" [online] [cit. 2018-09-26]. Available from: https://www.cdc.gov/healthliteracy/learn/index.html
[25] Cohen, J. Statistical Power Analysis for the Behavioral Sciences. Routledge; 1988.