NEAT: Nurse Effort Assessment Tool—Human Factors Considerations in Designing for Appropriate Staffing

Proceedings of the Human Factors and Ergonomics Society Annual Meeting
2024, Vol. 68(1) 1680–1684

DOI: 10.1177/10711813241261347
journals.sagepub.com/home/pro
Sage

Scott Good[1], Michael W. Boyce[2,3], Leigh V. Evans[2,3], and Mark Sevilla[2,3]

Abstract

This research has developed an algorithm to measure behavioral health acuity that relies predominantly on patient-level information from electronic health records (EHRs) but also has nurse-input-driven parameters contributing to the tool, known as the Nurse Effort Assessment Tool (NEAT). The overall objective is to create the NEAT to address the increasing acuity of behavioral patients by moving away from subjective patient assessment to using an automated algorithm that draws predominantly from the rich data in EHRs. Data from over 9,000 patient records provided statistically significant results that the subjective rating both correlates and serves as significant predictors for the NEAT score. Future research looks to expand the algorithm and validate it across multiple clinical sites.

Keywords

inpatient psychiatric care, patient acuity, human factors, nurse staffing, behavioral health

Introduction

The quantity and acuity of U.S. residents experiencing mental health challenges has steadily increased over the past 20 years—accelerating since the COVID-19 pandemic (Goodwin et al., 2022). To exacerbate the issue, inpatient care facilities face dramatic staffing challenges, as many nurses have retired or resigned. A recent study (American Nurses Foundation, 2022) shows that insufficient staffing is the most common reason nurses leave the profession; nine of ten nurses indicate that the hospital where they work is understaffed. In addition, hospital expense growth—of which, labor costs are the biggest contributor—has increased at a rate more than twice the average reimbursement rate (American Hospital Association, 2023). Burgeoning mental health needs combined with critical capacity issues and health system cost structures drive the need to more efficiently deploy direct care staff in the mental health inpatient setting.

In response to the increasing acuity seen in psychiatric facilities, the Harbor Performance Initiative (HPI) started a systematic process of developing an algorithm to measure behavioral health acuity that relies predominantly on patient-level information from EHRs. This paper reports on the experience of one HPI member site in collecting data to produce the NEAT measures of nurse staffing effort.

This research focuses on developing patient handoff metrics in inpatient behavioral health settings. Patient handoff is: "the process of sharing pertinent health information during transitions in patient care" (Keebler et al., 2016). Appropriate levels of nurse-to-patient ratios avoid nursing staff burnout, fatigue, cognitive stress, and psychological stress (American Psychiatric Nurse Association [APNA], 2012; Harper & McCully, 2007). In hospital environments, providing the highest quality care for patients is essential. Knowing that there are tools, processes, and procedures to help facilitate and inform healthcare personnel can assist healthcare workers. The need for appropriate assessment of behavioral health patients continues to grow with the complexity and comorbidities of patients. Mental illness is "a mental state which is characterized by alternation in thinking, mood, or behavior (or some combination thereof) and associated with significant distress and impaired functioning for an extended period" (Canadian Federation of Mental Health Nurses

[1]Crescendo Consulting Group, Portland, ME, USA
[2]Yale School of Medicine, New Haven, CT, USA
[3]Yale New Haven Health, New Haven, CT, USA

Corresponding Author:
Michael W. Boyce, Department of Emergency Medicine, Yale School of Medicine, 20 York Street, New Haven, CT 06510, USA.
Email: michaelwboyce@gmail.com

[CFMHN], 2014). Human Factors Engineering (HFE) is positioned to support appropriate staffing measures, environmental design, and tailored intervention approaches to maximize the safety of patients and staff. Barriers to an appropriate handoff process include clinician fatigue, poor teamwork, time constraints, and lack of knowledge of effective processes (Cannon, 2021; Keebler et al., 2016). Currently, staffing loads for nursing are predominantly based on ratios of patients per staff member with nurse assessments of patient care needs, being the most commonly researched approach. However information about the patient's acuity level can be generated from multiple sources, and more granular ratios can be developed with more precise and complex data.

Related Work

Algorithmic Approaches

Researchers have recently developed algorithmic approaches to estimate patient needs and nurse staffing. A Discrete Event Simulation (DES) model adjusted two proven drivers of workload: (a) nurse workload capacity, and (b) patient acuity based on historical data. This simulation modeled 35 different conditions, each comprising a combination of the different acuity levels and nurse-patient ratios compared to subjective nursing assessment models (Qureshi et al., 2020). Another algorithmic approach utilized EHR abstraction to develop individual patient scores, a product of nine patient need score components (e.g., medications, lines, drains, airways, risks, wounds, orders), to automate the creation of individual patient workload score (Womack et al., 2021). This early-stage algorithm received criticism for arbitrary cut-off points, nurses' concerns over the number of outliers in pediatric and mental health units, and its failure to calculate the number of positions needed.

Existing Measures

Acute Hospital. Four major approaches to determining staffing of nurses in acute hospitals are: (a) professional adjustment, (b) volume-based methods such as patient ratios, (c) patient prototype classification, and (d) time-task approaches (Griffiths et al, 2020). In the U.S., four popular measures of nurse staffing for intensive care units (ICUs) are: (a) the Therapeutic Intervention Scoring System (TISS), (b) the Nursing Activities Score (NAS), (c) the Subjective Workload Assessment Technique (SWAT), and (d) the National Aeronautics and Space Administration-Task Load Index (NASA-TLX). The TISS quantifies type and number of intensive care treatments for each patient to determine the needed number of nursing staff (Moreno & Morais, 1997). The NAS determines nurse workload by measuring the time to address patient needs through seven major categories and 23 activities (Miranda et al., 2003). The SWAT, which employs a human factors approach, measures three dimensions of nurse workload: cognitive, time, and psychological stress burden (Reid & Nygren, 1988). The NASA-TLX also employs a human factors approach and measures six dimensions of a nurse's workload perception: cognitive demand, physical demand, time pressure, performance, effort, and frustration (Lebet et al., 2021; Racy et al., 2021). A common weakness of this field is the lack of research determining which approach is more accurate or appropriate in a given circumstance.

Mental Health Inpatient Facilities. Current nursing mental health staffing ratios rely primarily on surveys of nurses about their education and skill set to determine capacity. A systematic review of 20 studies on ratios of mental health nurses to inpatient psychiatric patients determined that most studies followed the same methodology: estimating a nurse-to-patient ratio by surveying nurses' skill mix and education (Moyo, 2020). Three studies found significant associations between a more highly educated nursing workforce and a lower mortality risk. Patient characteristics in models of nursing-staff needs include primary psychiatric diagnosis, previous psychiatric hospitalization, psychiatric subdiagnoses, procedures, and number of physical subdiagnoses as measured by the Elixhauser Comorbidity Measures (ECM) (American Psychiatric Nurse Association [APNA], 2012; Park et al., 2020).

Measuring Patient Acuity

Accurately measuring patient acuity to determine required nursing effort is essential for estimating optimal nurse staffing needs, but approaches to determining patient acuity vary. A common approach is to rely on patient classification systems based on a nurse assessment (De Groot, 1989; Harper & McCully, 2007). Other approaches rely on the complex and detailed data collected in electronic patient records (Juvé-Udina et al., 2020; Needleman et al., 2011).

Variables of interest. Common measures of adverse outcomes, often referred to as nursing-sensitive indicators, include mortality, nosocomial infections, hospital-acquired pneumonia, unplanned extubation, respiratory failure, cardiac arrest, and failure to rescue (Kane et al., 2007; Needleman et al., 2011; Oner et al., 2021). Patient status variables include previous patient care, medications, activities of daily living, wounds, lines, drains, and airways (Qureshi et al., 2020; Womack et al., 2021). Factors that capture the physical characteristics and operating procedures of a nursing unit include inpatient unit floor plan, task location, walking distance, operating logic, and facility characteristics (Park et al., 2020; Qureshi et al., 2020).

Research Gaps and Open Questions

Concerns have been raised about the potential bias and inaccuracy due to the reliance on subjective ratings of nurse capacity (Lebet et al., 2021). In staffing ratios for inpatient

psychiatric patients, the American Psychiatric Nurses Association issued a 2023 position statement noting the scarcity of well-supported nursing-to-patient ratios and a strong understanding of their relation to outcomes developed and evaluated with data (Johnson et al., 2023).

The behavioral health assessment literature lacks empirically developed algorithmic solutions. More specifically, algorithmic solutions that have been verified to meet the needs of practitioners by practitioners and validated through psychometric analyses that individual variables measure the latent constructs, which can lead to improved understanding. An acceptable middle ground can be achieved by combining a data-driven approach and nurse stakeholder input.

This collaboration brought together a group of researchers seeking to draw a connection between data collected and the level of care a patient would need. Further, it would have the capability to provide quantitative justification related to the expense of resources thereby advocating for appropriately adjusted reimbursement rates. The goal is to develop and iteratively prove that our underlying algorithm is usable and extensible, continuing to adapt to meet the needs of patients. The Nurse Effort Assessment Tool (NEAT) is being deployed across multiple hospital systems to validate its effectiveness.

Our two primary research questions were:

1. How do the subjective scores provided by the nurses associate with the overall NEAT Score?
2. In terms of predictive variance, how much do the individual nurse scores account for?

These led to the following two hypotheses:

H_1: The nurse assessment scores will be positively correlated to the NEAT Score, so as nurse effort increases, so does the overall NEAT Score.
H_2: The nurse assessment scores will represent a significant portion of the variance explained by the overall NEAT score.

Methods

Participants

Participants comprised 9,417 de-identified archived patient records captured between November 2023 and January 2024. Records were collected as an initial pilot test of our algorithm through one participating healthcare facility in the southeastern United States. Records contained the aggregate assessment score and subjective inputs from the nursing staff.

Materials

Patient records were outputted into a Microsoft Excel workbook (Version 2312) and then extracted into an IBM SPSS Statistics data file (Version 29.0.2).

Algorithm Variables

The algorithm includes a proprietary set of 12 measures extracted seamlessly from the EHRs, such as fall risk, medication, and recent admissions, and two measures from the nurses delivering the care. The subjective input from the nurses is because those two values are the nurses' perceived state of the patient: one of which is the current state, and the other is what they believe to be the future state. Subjective data is what one gathers from the patient that you cannot use your five senses to measure (Adam et al., 2023).

The two nursing measures are:

1. Current State: refers to the nurse effort required for current patient care on a 1 to 4 scale: 1 refers to minimal effort, 2 refers to average care needs, 3 refers to above-average needs, and 4 refers to almost constant levels of effort.
2. Future State: refers to the nurse effort required for future patient care on the same 1 to 4 scale mentioned above.

Development and Verification

To understand the relationship between the subjective variables from the nurses and how they played into the overall aggregate score, correlation analyses were used first, followed by regression analysis.

Correlation Assumptions and Analysis Strategy. Bivariate correlation analysis assessed the correlation between the nurse-reported inputs and the overall NEAT score. It is acknowledged that scales such as the 4-point scale the nurses were using could be viewed as ordinal, and thereby violate the assumption of continuous variables within correlation. However, the reason we pursued this analysis technique is based on the way that end users suggested these categories. Since patients can be seen as moving within the stages (average to above average and vice versa) from a care perspective, this comparison would still be beneficial and provide insight for our regression analysis. The correlation also violated the normality assumption according to Shapiro-Wilk ($p \leq .05$); however, the test was run because correlations are known to be somewhat robust to deviations from normality. The other three assumptions were met (pairs, linearity, and no significant outliers). Two correlations were run: NEAT Score versus Current State, and NEAT Score versus Future State.

Regression Assumptions and Analysis Strategy. Linear Regression was chosen as the analysis method to assess the overall nurse score. It is acknowledged (as above) that the nurse scale violates the assumption of continuous variables. This method was done this way because it is an exploratory analysis that we extend to expand both the nursing score to a scalar format (e.g., moving it from a 1 to 100 scale) and because

follow-on analysis is intended to extract each of the 14 variables and look at their overall weighting score. Our data had 20 outliers (out of 9,417 records). Still, we decided to maintain the outliers as we believe they accurately represent patient data and don't represent significant changes to the overall model. We acknowledge that multicollinearity exists due to the correlation between the two variables. Still, in terms of our research question, we are looking at the variance accounted for by both variables (including overlap) and not looking at individual contributions. The assumptions of linearity, homoscedasticity, and independence were met.

Results

In the first analysis phase, we examined the correlation between the subjective measurements and the overall NEAT score. The Pearson correlations indicate a high and significant relationship of 67% between the NEAT Score and the subjective current measurement of patient need (0.671, $p < .001$) and the subjective future measurement of patient need (0.672, $p < .001$). Next, the regression model analysis, which removed duplicates, determined that current and future intervention scores predicted 46% of the NEAT score variance ($p < .001$). In the regression analysis limited to 142 instances where scores shifted for the current and future intervention, these changes predicted 29% of the NEAT score variance ($p < .001$). This finding on the change intervention scores verified the robustness and predictability of the NEAT score.

Discussion

Few research studies have examined approaches to optimize the use of extensive data in EHRs to determine nursing staff levels in inpatient mental health facilities. A strength of NEAT is the inclusion of 12 measures derived based on EHR data and two measures based on nurse assessments of patient level of care. The correlation analysis indicated a high degree of relationship (67.1%; 67.2%) between the EHR measure of patient needs with the nurse assessment of current and future care. The regression analysis, which removed duplicate cases, indicated that current and future nurse scores predicted 46% of the NEAT variance. Therefore, NEAT is capturing nearly half of the patient needs that nurse assessments do.

Implications and Applied Use Cases

This study is the first step in incorporating the rich objective measures documented in the EHR with nurse assessments to determine the level of nursing care patients require. Studies on acute hospital care present an initial list of nurse-sensitive conditions that must be modified to mental health conditions (Kane et al., 2007; Needleman et al., 2011; Oner et al., 2021).

Future Work

Future research could examine how well NEAT nursing levels prevent adverse patient outcomes in mortality, morbidity, longer patient stays, and readmissions. To continue its work on inpatient behavioral health staffing, HPI will implement NEAT across participating inpatient facilities to benchmark the labor resources required to provide quality mental health care. The information gained will provide inpatient mental health facilities with evidence-based staffing requirements during payment rate negotiations.

Conclusion

NEAT advances the field of nursing care by developing an algorithmic approach to determine nurse-to-patient staff ratios (Griffiths et al., 2020). By relying on patient-level data in EHRs, NEAT provides a less time-intensive approach to identifying nurse staffing requirements. Measures of the high correlation between the 12 EHR-derived and two nurse-created measures indicate the potential for future versions of NEAT to improve the power of EHR-derived measures to more closely match the nursing measures.

Acknowledgments

We thank the Harbor Performance Initiative (HPI) which launched a 2017 workgroup to study patient acuity measurement in inpatient behavioral health units. This work led to the development of the NEAT

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Michael W. Boyce https://orcid.org/0000-0002-9869-3989

References

Adam, S., Gold, E., & Burstow, B. (2023). From subjective opinion to medical fact: A critical discourse analysis of mental health nursing education. *Issues in Mental Health Nursing*, *44*(1), 55–63. https://doi.org/10.1080/01612840.2022.2113940

American Hospital Association. (2023). *The financial stability of America's hospitals and health systems is at risk as the costs of caring continue to rise*. https://www.aha.org/costsofcaring

American Nurses Foundation. (2022). *Pulse on the nation's nurses survey series: COVID-19 two-year impact assessment survey*. https://www.nursingworld.org/~4a2260/contentassets/872ebb13c63f44f6b11a1bd0c74907c9/covid-19-two-year-impact-assessment-written-report-final.pdf

American Psychiatric Nurse Association [APNA]. (2012). APNA position statement: Staffing inpatient psychiatric units. *Journal of the American Psychiatric Nurses Association*, *18*(1), 16–22. 10.1177/1078390311432133

Canadian Federation of Mental Health Nurses [CFMHN]. (2014). *Canadian standards for psychiatric-mental health nursing: Standards of practice* (4th ed.). https://live-cfmhn.pantheonsite.io/wp-content/uploads/2019/05/2014-Standards-of-Practice-Final-1.pdf

Cannon, K. K. (2021). *Utilization of a standardized handoff tool (PSYCH) to reduce errors of omission*. [Doctoral dissertation, University of Maryland Baltimore]. UMB Digital Archive. https://archive.hshsl.umaryland.edu/handle/10713/15692

De Groot, H. A. (1989). Patient classification system evaluation: Part 2, system selection and implementation. *The Journal of Nursing Administration*, *19*(7), 24–30.

Goodwin, R. D., Dierker, L. C., Wu, M., Galea, S., Hoven, C. W., & Weinberger, A. H. (2022). Trends in U.S. depression prevalence from 2015 to 2020: The widening treatment gap. *American Journal of Preventive Medicine*, *63*(5),726–733. https://doi.org/10.1016/j.amepre.2022.05.014

Griffiths, P., Saville, C., Ball, J., Jones, J., Pattison, N., & Monks, T., & SAFER NURSING CARE STUDY GROUP. (2020). Nursing workload, nurse staffing methodologies and tools: A systematic scoping review and discussion. *International Journal of Nursing Studies*, *103*, 103487. https://doi.org/10.1016/j.ijnurstu.2019.103487

Harper, K., & McCully, C. (2007). Acuity systems dialogue and patient classification system essentials. *Nursing Administration Quarterly*, *31*(4), 284–299. https://doi.org/10.1097/01.NAQ.0000290426.41690.cb

Johnson, C., Delaney, K. R., Cirpili, A., Marriott, S., & O'Connor, J. (2023). American psychiatric nurses association position: Staffing inpatient psychiatric units. *Journal of the American Psychiatric Nurses Association*. https://doi.org/10.1177/10783903231198247

Juvé-Udina, M. E., González-Samartino, M., López-Jiménez, M. M., Planas-Canals, M., Rodríguez-Fernández, H., Batuecas Duelt, I. J., Tapia-Pérez, M., Pons Prats, M., Jiménez-Martínez, E., Barberà Llorca, M. À., Asensio-Flores, S., Berbis-Morelló, C., Zuriguel-Pérez, E., Delgado-Hito, P., Rey Luque, Ó., Zabalegui, A., Fabrellas, N., & Adamuz, J. (2020). Acuity, nurse staffing and workforce, missed care and patient outcomes: A cluster-unit-level descriptive comparison. *Journal of Nursing Management*, *28*(8), 2216–2229. https://doi.org/10.1111/jonm.13040

Kane, R. L., Shamliyan, T. A., Mueller, C., Duval, S., & Wilt, T. J. (2007). The association of registered nurse staffing levels and patient outcomes: Systematic review and meta-analysis. *Medical Care*, *45*(12), 1195–1204. https://doi.org/10.1097/MLR.0b013e3181468ca3

Keebler, J. R., Lazzara, E. H., Patzer, B. S., Palmer, E. M., Plummer, J. P., Smith, D. C., Lew, V., Fouquet, S., Chan, Y. R., & Riss, R. (2016). Meta-analyses of the effects of standardized handoff protocols on patient, provider, and organizational outcomes. *Human Factors*, *58*(8), 1187–1205. https://doi.org/10.1177/0018720816672309

Lebet, R. M., Hasbani, N. R., Sisko, M. T., Agus, M. S., Nadkarni, V. M., Wypij, D., & Curley, M. A. (2021). Nurses' perceptions of workload burden in pediatric critical care. *American Journal of Critical Care*, *30*(1), 27–35.

Miranda, D. R., Nap, R., de Rijk, A., Schaufeli, W., & Iapichino, G., & MEMBERS of the TISS WORKING GROUP. (2003). Nursing activities score. *Critical Care Medicine*, *31*(2), 374–382. https://doi.org/10.1097/01.CCM.0000045567.78801.CC

Moreno, R., & Morais, P. (1997). Validation of the simplified therapeutic intervention scoring system on an independent database. *Intensive Care Medicine*, *23*(6), 640–644. https://doi.org/10.1007/s001340050387

Moyo, N., Jones, M., Kushemererwa, D., Pantha, S., Gilbert, S., Romero, L., & Gray, R. (2020). The association between the mental health nurse-to-registered nurse ratio and patient outcomes in psychiatric inpatient wards: A systematic review. *International Journal of Environmental Research and Public Health*, *17*(18), 6890. https://doi.org/10.3390/ijerph17186890

Needleman, J., Buerhaus, P., Pankratz, V. S., Leibson, C. L., Stevens, S. R., & Harris, M. (2011). Nurse staffing and inpatient hospital mortality. *The New England Journal of Medicine*, *364*(11), 1037–1045. https://doi.org/10.1056/NEJMsa1001025.

Oner, B., Zengul, F. D., Oner, N., Ivankova, N. V., Karadag, A., & Patrician, P. A. (2021). Nursing-sensitive indicators for nursing care: A systematic review (1997–2017). *Nursing Open*, *8*(3), 1005–1022. https://doi.org/10.1002/nop2.654

Park, S., Park, S., Lee, Y. J., Park, C. S., Jung, Y. C., & Kim, S. (2020). Nurse staffing and health outcomes of psychiatric inpatients: A secondary analysis of national health insurance claims data. *Journal of Korean Academy of Nursing*, *50*(3), 333–348. https://doi.org/10.4040/jkan.19203

Qureshi, S. M., Purdy, N., & Neumann, W. P. (2020). Development of a methodology for healthcare system simulations to quantify nurse workload and quality of care. *IISE Transactions on Occupational Ergonomics and Human Factors*, *8*(1), 27–41.

Racy, S., Davidson, P. M., Peeler, A., Hager, D. N., Street, L., & Koirala, B. (2021). A review of inpatient nursing workload measures. *Journal of Clinical Nursing*, *30*(13–14), 1799–1809. https://doi.org/10.1111/jocn.15676

Reid, G. B., & Nygren, T. E. (1988). The subjective workload assessment technique: A scaling procedure for measuring mental workload. *Advances in Psychology*, *52*, 185–218.

Womack, D., Warren, C., Hayes, M., Stoyles, S., & Eldredge, D. (2021). Evaluation of electronic health record-generated work intensity scores and nurse perceptions of workload appropriateness. *Computers, Informatics, Nursing*, *39*(6), 306–311. https://doi.org/10.1097/CIN.0000000000000687