The quantification of nursing workload in Intensive Care Units

Margadant, C.C.

Publication date
2020

Document Version
Final published version

License
Other

THE QUANTIFICATION OF NURSING WORKLOAD IN INTENSIVE CARE UNITS

Charlotte Margadant

The quantification of nursing workload in Intensive Care Units

Charlotte Margadant

Financial support by the National Intensive Care Evaluation (NICE), Itémedical and Chipsoft for the publication of this thesis is gratefully acknowledged.

Cover design and lay-out: Judith van der Werf
Printing: Drukkerij Haveka

The quantification of nursing workload in Intensive Care Units

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van Rector Magnificus
prof. dr. ir. K.I.J. Maex
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op woensdag 30 september 2020, te 13.00 uur

door
Charlotte Claire Margadant
geboren te Apeldoorn

Promotiecommissie

Promotor:	
Prof. dr. N.F. de Keizer	AMC-UvA
Copromotores:	
Dr. S. Brinkman	AMC-UvA
Dr. J.J. Spijkstra	VU Medisch Centrum
Overige leden:	
Prof. dr. A. Abu-Hanna	AMC-UvA
Prof. dr. M.B. Vroom	AMC-UvA
Dr. F. Paulus	AMC-UvA
Prof. dr. B.M. Buurman-Van Es	AMC-UvA
Prof. dr. A.R.J. Girbes	Vrije Universiteit Amsterdam

Faculteit der Geneeskunde

Table of contents

CHAPTER 1

GENERAL INTRODUCTION

1.1 Introduction

A common description of Intensive care is the 'continuous monitoring and treatment of seriously ill patients using special medical equipment and services'. This requires special services. To provide these patients with optimal care, a close cooperation between the nurse and the physician is considered to be of great importance [1]. Intensive nurse-physician collaborations (e.g. sharing responsibilities and making decisions together) are significantly associated with health outcomes in critical care settings, such as a decline in infections and mortality rates [2,3]. Intensive Care Unit (ICU) nurses do not only have to work closely with physicians, but the fact that critically ill patients require continuous attention and on occasion rapid interventions makes that they have to work more autonomous compared to nurses on the general wards [4,5].

In Europe, the costs of ICUs represent approximately 20% of total hospital costs [6]. The demanding level of care and monitoring are the main reasons why ICU care is expensive. ICU needs high staffing levels, the use of expensive medical devices and equipment, complex procedures, and extensive IT infrastructure. According to Bittner and colleagues and Miranda and colleagues [7,8], an average of 50% of an ICU budget is spend on nursing staff. From an economical point of view, it is therefore important to establish how many nurses are required to perform good quality care. Excessive budgetary cuts targeted to reduce the number of nursing staff are likely to increase the nursing workload, which may have a negative impact on patient safety and survival chances [2,9]. Moreover, reducing the number of nurses may also negatively impact the nurses well-being [10]. Clear insight into the differences in workload over different shifts and for different type of patients is absent. It is no surprise therefore, that the quantification of an optimal Patient to Nurse Ratio (PNR), as how the workload is often expressed, has been a topic of debate for years [9,11,12].

To determine an optimal PNR one really needs an instrument to measure the needed care objectively. Without such an accurate method to measure needed nursing time it is difficult to assess the optimal level of care and makes ICUs at risk for under- or overstaffing of nurses. Additionally, the lack of a validated measurement method hampers the development of national norms for adequate staffing [13]. In order to take the capacity planning of nurses beyond rough estimates and assumptions, an accurate method for measuring nursing workload is essential.

In this thesis, we examined which nursing workload models are currently used in ICUs and how valid they are. We described variability in needed nursing time over different patient- and contextual groups. We investigated whether workload is associated with in-hospital mortality, and we developed a novel instrument to quantify the nursing workload at ICUs in the Netherlands. In doing so, we intended to improve the quality of healthcare for both patients and nurses.

1.2 Quality of care

In line with the definition of Lohr and colleagues (1990) [14], we define quality of care as 'the degree to which health services for individuals and populations increase the likelihood of desired outcomes and are consistent with current professional knowledge.' The Institute of Medicine in the United States of America suggests that quality of care encompasses various aspects, and that improving the quality of care hinges on: safety (to avoid possible harm resulting from received care), effectiveness (to provide care based on scientific evidence), patient centeredness (to provide care which suits the needs, preferences, and values of the patient), timeliness (to reduce wait- and delay times for both caregivers and care receivers), efficiency (to keep the costs of care in balance with the required care), and equity (to provide the same care to all patients, independent of personal characteristics such as gender, ethnicity or geological origin) [15-17]. In this thesis we primarily focused on the efficiency aspect of care.

In efficiency of care, costs are kept in balance with the required care [17,18]. To achieve high efficiency within a hospital setting, good management is needed, which should lead to lower costs with equal quality of care or a higher quality of care delivered with equal costs [19]. Hospital management often wants to minimize the number of nurses while keeping good outcome for patients and a healthy workload for nurses. For this reason, insight in the optimal workload per nurse is of importance. Currently, there are no clear guidelines for the number of patients a nurse can take care off. Furthermore, it is unknown whether this patients per nurse ratio is a reasonable measure for quality of care. Current guidelines in the United States over the PNR have yielded disappointing results [20]. In our studies, we focus on the improvement of nursing capacity planning in ICUs, and therefore target the efficiency aspect of quality of care.

1.3 National Intensive Care Evaluation

In 1996 a group of intensivists established the National Intensive Care Evaluation (NICE). The registry serves to allow different ICUs to benchmark their data against each other, with the goal to improve the overall effectiveness of Dutch ICU care [21]. Since its inception, the NICE registry is constantly working towards improving the quality of data provided by the ICUs, by providing site-visits to check the data quality, by making quality reports and data validation reports, by answering questions regarding the data of ICUs, etc. At the start NICE collected a limited set of data, called the minimal dataset, which included data on the severity of illness of patients and their outcomes in terms of length of stay and mortality. Over the years the data has been extended by incorporating more information on quality indicators (e.g. the number of intensivists and ICU nurses for each ICU, the duration of mechanical ventilation and glucose regulation), complications, daily organ failure, nursing workload, and quality indicators designed to make rapid corrections possible (information on the management of pain, blood transfusion, respiration, and antibiotics). At present, all Dutch ICUs participate in the NICE registry and collect the minimal dataset, the other modules are voluntarily and are collected by Dutch ICUs according to their usefulness, which is different for each individual participant. As part of the monitoring process, the registry provides semi-annual feedback reports and data quality reports, and offers its participants the possibility to keep track on their ICU data online on a continuous basis between the semi-annual reports. Additionally, the registry is used to perform scientific research with the collected data. Both the reports and the online monitoring tools are meant to make detection of points for improvement in ICUs possible, which can then be addressed by initiating quality improvement activities. The effect of these activities can be monitored by using the registry which closes Deming's Plan Do Check Act cycle [22].

In recent years, there has been an explicit need to optimize the efficiency at Dutch ICUs due to budget constraints and a shortage of ICU nurses. In this light, the NICE registry has developed the nursing capacity module, in cooperation with the Dutch society of ICU nurses, the V&VN IC [23], which is meant to offer insights into the nursing workload at an ICU. Such knowledge could be used to deploy nurses in the most efficient way. The nursing capacity module consists of a combination of the already existing Nursing Activities Score (NAS) and the Therapeutic Intervention Scoring System (TISS) [24,25], supplemented by five additional variables, which we thought to be lacking in two existing instruments, to get a more accurate picture of work done by nurses at Dutch ICUs. To get a complete picture and to gain a better understanding of the average workload per nurse, the number of qualified nurses (ICU, cardiac care, medium care, and recovery), and the number of non-qualified (student) nurses, are also registered in the module.

1.4 The importance of quantifying nursing workload

Importance for ICU managers

ICU managers are expected to manage the ICU as efficiently as possible. They have to provide enough personnel and enough means for optimal quality of patient care and at the same time they are also responsible for the welfare of their employees. To be able to do so ICU managers need valid instruments to measure nursing workload. These instruments should be able to give information on capacity planning to accomplish the best outcomes. ICU nurses may experience a high workload due to stressful situations and insufficient time for their patients [26]. The systematic review of Chuang et al.(2016) [27] gives a prevalence of 6 to 47% of ICU nurses and physicians who suffer from burnout. One should also realize that the work on the ICU can create a lot of emotional stress, although this subject, important as it may be, was not part of the thesis. A work pressure which is too high may result in an increase of nurses being absent for longer periods of time [28]. It is clear that in a period of rising nursing costs, nurse shortages, and a competing job market, it is important to deploy nurses in the most optimal way and to avoid loss of staff because of illness. Quantifying nursing workload enables ICU managers to efficiently plan the needed nursing capacity.

Importance for ICU nurses

In recent years, the workload of ICU nurses has risen due to increased rotation of patients, inadequate staffing of nurses due to shortages, and increased demand of nursing care due to complexity of patients [29]. Currently, workload planning of nurses in hospitals is mainly based on financial arguments, process optimization (instead of planning on nursing workload), and historical data [30]. Since these measures fall short of accurately portraying the workload of nurses, Twigg and colleagues (2009) argue that nurses are likely to experience a high workload, especially when patients need complex and extensive amounts of care [30]. A high workload has negative consequences for ICU nurses; it can increase the chance of burnout and greater job dissatisfaction [31]. Additionally, nurses with a higher workload are more likely to experience emotional exhaustion and somatization [32]. When nursing workload is quantified well, nurses could be deployed as efficiently as possible without over demanding, which will decrease their risk of physical and mental exhaustion.

Importance for patients

Patient care profits from an adequate staffing of nurses. Many studies have shown that the quality of care increases with adequate numbers of nurses. Lower numbers of nurse staffing per patient are associated with higher mortality rates and longer hospital admissions for patients, both inside and outside ICU [2,9,10,33]. Patients at the ICU which are under the care of nurses with a high workload are at higher risk of decubitus, pneumonia, and sepsis [34]. Also their morbidity and mortality rates rise significantly with higher nursing workload [30,35,36]. The ICU- and hospital length of stay of ICU patients reduces with high-intensity ICU physician staffing [37]. Therefore, capacity planning is of great importance to ensure that patients are not unnecessarily at risk for negative health outcomes. However, there is no clear guideline for how many nurses are enough to take care of the ICU patients.

Until now, only the PNR is used for this purpose, but this measure does not show consequent results [9,11,12,20]. As every patient generates its own workload and this workload may be quite different from patient to patient, one can imagine that using a fixed number of patients per nurse can give very different results regarding the work that has to be done. This raises the question whether the PNR is a good measure for the quantification of nursing workload. Many studies have been performed on the optimal number of ICU patients per nurse, since it has been a topic of debate for years [9,11,33]. However, these studies give mixed results [34].

1.5 Current workload models

It is of great importance to quantify nursing workload in ICUs to limit unnecessary negative patient outcomes [2,9,10]. This quantification can be done in several ways, for instance by scoring the amount and type of nursing actions per shift or in 24 hours, adjust points to every specific nursing actions, and determine how many points a nurse should be able to do. Instead of points an alternative can be used to adjust a specific amount of nursing time for each specific action. Over the last five decades, many nursing workload models have been developed for ICUs. However, only two of these models are widely used. These models are the Therapeutic Intervention Scoring System (TISS) and the Nursing Activities Score (NAS).

Therapeutic Intervention Scoring System

In 1974, Cullen and colleagues [24] developed the TISS originally in order to classify severity of illness of ICU patients. Later, the TISS was used more and more to classify nursing workload of ICU patients. The original TISS exists of 76 therapeutic interventions that receive one to four points based on the severity of illness of the patient. However, nursing workload appeared to be only partly related to severity of illness, since less severely ill patients could also generate high nursing workload, which makes the TISS less adequate to assess nursing workload. The TISS only maps medical proceedings which are only a small part of all the care that has to be done in the ICU. An example of a patient which does not score high in severity of illness, but takes very intensive nursing care, is a patient recovering from a serious illness with an agitated delirium, necessitating constant and continuous bedside care throughout the day.

Nursing Activities Score

The NAS was developed in 2003 by Miranda et al. [25] to correct the disadvantages of the TISS and to focus more on the caring process. The NAS describes activities that largely represent the work actually done by ICU nurses at the bedside in caring for the patients. It measures the nursing workload needed for each individual patient. The points assigned to the nursing activities provide a measure of time consumption in caring for the patients instead of representing the severity of illness.

The TISS has been developed more than 40 years ago and the NAS some 15 years ago. Since ICU healthcare is rapidly evolving, some of the nursing activities from the current daily ICU practice are absent in the TISS and NAS. These activities might be needed to make a valid tool to measure nursing workload today.

1.6 Objectives

To support ICU nurses and managers with capacity planning of ICU nurses, we aim to quantify nursing workload more precisely to deploy nurses in the most efficient way. An accurate model is needed to measure the amount of nursing workload and therewith needed nursing time per patient. The conducted studies in this thesis are undertaken to study the performance of current nursing workload models in Dutch ICUs, and to develop a new improved workload model in order to organize ICU care more efficiently, accounting for both the quality of care for ICU patients and the well-being for ICU nurses.

The objective of this thesis is to develop and evaluate models to quantify nursing workload in order to improve the quality and efficiency of care in ICUs. To realize this objective, four research questions are addressed in this thesis:

1. Which nursing workload models are used in ICUs and how valid are they to estimate nursing workload?

2. Is nursing workload associated with in-hospital mortality in the Dutch ICU population?

3. How could nursing workload be quantified in the most accurate way?

4. How important are patient characteristics and contextual factors for the amount of nursing workload generated?

1.7 Outline of this thesis

Chapter 2 of this thesis describes a systematic literature review on nursing workload models used in ICU care and their ability to quantify nursing time. To check the accuracy of the models in quantifying nursing time, we payed special attention to their content validity, reliability, and validity. In Chapter 3 we performed a validation study on the currently most used nursing workload model, the NAS, to find out whether this model is in need of revisions after more than 15 years since its development. In Chapter 4 we determined the association between nursing workload and in-hospital mortality and tested the effect of case-mix on this association. In Chapter 5 we developed a nursing workload model based on time measurements of the nursing activities in the NAS, TISS, and five additional activities. In Chapter 6 we looked at frequencies and differences in nursing workload over several patient characteristics and contextual factors, to determine whether these are relevant in the process of nursing capacity planning. Finally, an overall discussion of our findings along with their strengths, limitations and implications is given in Chapter 7.

References

1. Manojlovich, M., & DeCicco, B. (2007). Healthy work environments, nurse-physician communication, and patients' outcomes. American Journal of Critical Care, 16(6), 536-543.
2. Boev, C., & Xia, Y. (2015). Nurse-physician collaboration and hospital-acquired infections in critical care. Critical Care Nurse, 35(2), 66-72.
3. Wheelan, S. A., Burchill, C. N., &Tilin, F. (2003). The link between teamwork and patients' outcomes in intensive care units. American Journal of Critical Care, 12(6), 527-534.
4. Varjus, S. L., Suominen, T., &Leino-Kilpi, H. (2003). Autonomy among intensive care nurses in Finland. Intensive and Critical Care Nursing, 19(1), 31-40.
5. Castellan, C., Sluga, S., Spina, E., et al (2016). Nursing diagnoses, outcomes and interventions as measures of patient complexity and nursing care requirement in Intensive Care Unit. Journal of advanced nursing, 72(6), 1273-1286.
6. Tan, S. S., Bakker, J., Hoogendoorn, M. E., Kapila, et al (2012). Direct cost analysis of intensive care unit stay in four European countries: applying a standardized costing methodology. Value in Health, 15(1), 81-86.
7. Bittner, M. I., Donnelly, M., Van Zanten, A. R., Andersen, J. S., Guidet, B., Cabello, J. J. T., ... & Schmutz, A. (2013). How is intensive care reimbursed? A review of eight European countries. Annals of intensive care, 3(1), 37.
8. Reis Miranda, D., & Jegers, M. (2012). Monitoring costs in the ICU: a search for a pertinent methodology. Acta Anaesthesiologica Scandinavica, 56(9), 1104-1113.
9. Penoyer, D. A. (2010). Nurse staffing and patient outcomes in critical care: a concise review. Critical care medicine, 38(7), 1521-1528.
10. Carayon, P., Alvarado, C. et al. Workload and Patient Safety Among Critical Care Nurses. Crit Care Nurs Clin North Am. 2007;19(2):121–9.
11. Costa, D.K.. Organizing critical care for the 21st century. JAMA: Journal of the American Medical Association. 2016;315(8):751-752.
12. Debergh, D. P., Myny, D., Van Herzeele, I., Van Maele, G., Miranda, D. R., & Colardyn, F. (2012). Measuring the nursing workload per shift in the ICU. Intensive care medicine, 38(9), 1438-1444.
13. Quality standard Dutch ICUs [Internet]. Available at: https://www.zorginzicht.nl/.
14. Lohr, KN. Medicare: A Strategy for Quality Assurance. Washington, DC: National Academy Press; 1990.
15. Wolfe, A. (2001). Institute of Medicine Report: crossing the quality chasm: a new health care system for the 21st century. Policy, Politics, &Nursing Practice, 2(3), 233-235.
16. Campbell, S. M., Roland, M. O., & Buetow, S. A. (2000). Defining quality of care. Social science & medicine, 51(11), 1611-1625.
17. American Quality Alliance (AQA). AQA principles of "efficiency" measures. June 2009. Accessed January 1st, 2020.

18. Jha, A. K., Orav, E. J., Dobson, A., Book, R. A., & Epstein, A. M. (2009). Measuring efficiency: the association of hospital costs and quality of care. Health Affairs, 28(3), 897-906.
19. Leatherman, S., Berwick, D., Iles, D., Lewin, L. S., Davidoff, F., Nolan, T., & Bisognano, M. (2003). The business case for quality: case studies and an analysis. Health affairs, 22(2), 17-30.
20. Scruth, E. (2018). Nurse-patient staffing relationship in the ICU: More than just numbers. Critical care medicine, 46(10), 1681-1682.
21. Dutch National Intensive Care Evaluation (NICE) foundation (2020). URL https://www.stichting-nice.nl/dd/#561
22. Moen, R., & Norman, C. (2006). Evolution of the PDCA cycle.
23. V&VN Intensive Care (V&VN IC) (2020). URL https://www.venvn.nl/afdelingen/intensive-care/
24. Cullen, D. J., Civetta, J. M., Briggs, B. A., & Ferrara, L. C. (1974). Therapeutic intervention scoring system: a method for quantitative comparison of patient care. Critical care medicine 1974; 2(2), 57-60.
25. Miranda, D. R., Nap, R., de Rijk, A., et al.: Nursing activities score. Critical care medicine 2003; 31(2), 374-382.
26. Moss, M., Good, V. S., Gozal, D., Kleinpell, R., & Sessler, C. N. (2016). A Critical Care Societies collaborative statement: burnout syndrome in critical care health-care professionals. A call for action. American journal of respiratory and critical care medicine, 194(1), 106-113.
27. Chuang, C. H., Tseng, P. C., Lin, C. Y., Lin, K. H., & Chen, Y. Y. (2016). Burnout in the intensive care unit professionals: a systematic review. Medicine, 95(50).
28. Costa, D. K., & Moss, M. (2018). The cost of caring: emotion, burnout, and psychological distress in critical care clinicians. Annals of the American Thoracic Society, 15(7), 787-790.
29. Palese, A., Comisso, I., Burra, M., DiTaranto, P. P., Peressoni, L., Mattiussi, E., &Lucchini, A. (2016). Nursing Activity Score for estimating nursing care need in intensive care units: findings from a face and content validity study. Journal of nursing management, 24(4), 549-559.
30. Twigg D, Duffield C. A review of workload measures: a context for a new staffing methodology in Western Australia. Int J Nurs Stud. 2009;46:131-9.
31. Aiken, L. H., Clarke, S. P., Sloane, D. M., Sochalski, J., & Silber, J. H. (2002). Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction. Jama, 288(16), 1987-1993.
32. Greenglass, E. R., Burke, R. J., & Fiksenbaum, L. (2001). Workload and burnout in nurses. Journal of community & applied social psychology, 11(3), 211-215.
33. Lang, T. A., Hodge, M., Olson, V., Romano, P. S., & Kravitz, R. L. (2004). Nurse–patient ratios: a systematic review on the effects of nurse staffing on patient, nurse employee, and hospital outcomes. JONA: The Journal of Nursing Administration, 34(7), 326-337.
34. Kleinpell, R., Ward, N. S., Kelso, L. A., Mollenkopf Jr, F. P., & Houghton, D. (2015). Provider to patient ratios for nurse practitioners and physician assistants in critical care units. American Journal of Critical Care, 24(3), e16-e21.

35. Numata, Y., Schulzer, M., Van Der Wal, R., et al. Nurse staffing levels and hospital mortality in critical care settings: Literature review and meta-analysis. J Adv Nurs. 2006;55(4):435–48.
36. Duffield, C., Diers, D., O'Brien-Pallas, L., Aisbett, C., Roche, M., King, M., &Aisbett, K. (2011). Nursing staffing, nursing workload, the work environment and patient outcomes. Applied nursing research, 24(4), 244-255.
37. Pronovost PJ, Angus DC, Dorman T, Robinson KA, Dremsizov TT, et al. Physician staffing patterns and clinical outcomes in critically ill patients: a systematic review. J Am MedAssoc. 2002;288:2151-62.
38. Wilcox ME, Chong CA, Niven DJ, Ruben¬feld GD, Rowan KM, et al. Do intensivist staffing patterns influence hospital mortal¬ity following ICU admission? A systematic review and meta-analyses. Crit Care Med. 2013;41:2253-74.

CHAPTER 2

WORKLOAD SCORING SYSTEMS IN THE INTENSIVE CARE AND THEIR ABILITY TO QUANTIFY THE NEED FOR NURSING TIME: A SYSTEMATIC LITERATURE REVIEW

Marga E. Hoogendoorn, Charlotte C. Margadant, Sylvia Brinkman, Jasper J, Haringman, Jan Jaap Spijkstra, Nicolette F. de Keizer

International Journal of Nursing Studies 2019; 103408

Abstract

Introduction

The Intensive Care Unit is a resource intense service with a high nursing workload per patient resulting in a low ratio of patients per nurse. This review aims to identify existing scoring systems for measuring nursing workload on the Intensive Care and assess their validity and reliability to quantify the needed nursing time.

Methods

We conducted a systematic review of the literature indexed before 01/Mar/2018 in the bibliographic databases MEDLINE, Embase, and Cinahl. Full-text articles were selected and data on systems measuring nursing workload on the Intensive Care and translation of this workload into the amount of nursing time needed was extracted.

Results

We included 71 articles identifying 34 different scoring systems of which 27 were included for further analysis as these described a translation of workload into nursing time needed. Almost all systems were developed with nurses. The validity of most scoring systems was evaluated by comparing them with another system (59%) or by using time measurements (26%). The most common way to translate workload-scores into nursing time needed was by categorizing the Nurse:Patient-ratios. Validation of the Nurse:Patient-ratios was mostly evaluated by comparing the results with other systems or with the actual planning and not with objective time measurements.

Conclusion

Despite the large attention given to nursing workload systems for Intensive Care, only a few systems objectively evaluated the validity and reliability of measuring nursing workload with moderate results. The Nursing Activity Score system performed best. Poor methodology for the translation of workload scores into Nurse:Patient-ratio weakens the value of nursing workload scoring systems in daily Intensive Care practice.

Introduction

The Intensive Care Unit in a hospital is a labor-intense service due to its highly complex patients and the consequently high amount of care they require. Therefore, nurses can care for only a limited number of patients. A high workload and a low Nurse:Patient-ratio have shown to be associated with an increased risk for nosocomial infections in Intensive Carepatients, unplanned extubations, and an increased risk of mortality (Neuraz et al., 2015; Aragon Penoyer, 2010; Lee et al., 2017; West et al., 2014; Driscoll et al., 2018). Although there is evidently a high need for nursing capacity, there are also constraints on the healthcare budget and the availability of educated Intensive Carenurses. It is clear that resources should be used as efficiently as possible, which means avoiding understaffing as well as overstaffing. Therefore, for managerial as well as financial reasons, quantification of the Nurse:Patient-ratio is an important issue as the costs for nursing staff comprise about 40% of total Intensive Care costs (Costa, 2016; Tan et al., 2012). Application of scoring systems to measure the amount of nursing time needed per patient, mostly translated into a Nurse:Patient-ratio, could provide insight into the required nursing capacity. This is increasingly important for Intensive Care management who has to focus on both quality and cost, including the implementation of guidelines on Nurse:Patient-ratios (Adviescommissie Kwaliteit van het Zorginstituut, 2016; Australian College of Critical Care Nurses Ltd (ACCCN), 2002). The application of a reliable nursing workload classification system might optimize both Intensive Care and hospital costs, availability of Intensive Care beds and improve patient outcome. Due to this importance, many systems have been developed for this purpose over the years. However, the validation and application of those systems in daily practice varies strongly. The objective of this study is to systematically review the literature to identify existing scoring systems used to measure the amount of nursing care needed for Intensive Care patients, evaluate the validity and reliability and evaluate which system is most useful in daily practice in terms of quantification of the required nursing capacity.

Methods

Search strategy

We searched the databases MEDLINE, Cinahl, and Embase for original studies with the primary aim to develop or validate a scoring system to quantify the nursing time needed for Intensive Care patients. We checked the references of the included publications for relevant publications. We searched all literature up till 01/Mar/2018. As the earliest publications on workload scoring systems, of which some are still in use, date from the early seventies we did not restrict the commencing date. We also checked the Cochrane-database for published reviews on this topic. The search strategy included MeSH terms and keywords for 'nursing', 'Intensive Care', 'scoring system', 'classification' and 'workload'. The exact search queries are presented in Appendix A. The results were first independently assessed by two reviewers (MH and CM) based on title and abstract. If there was no abstract available, but the title indicated potential relevance, the article was selected for full-text reading. The full text of the selected articles was independently judged by the same two reviewers on the inclusion criteria. Differences in selection of articles were solved by discussion and in case of disagreement resolved by a third reviewer (NdK). We used a PRISMA flow chart for reporting the results of the search (www.equator-network.org/ reporting-guidelines/prisma/) (Liberati et al., 2009).

Selection criteria

Papers were selected when they adhered to all of the following inclusion criteria:

- It concerned an original study on either:
- The development of a new scoring system to measure nursing workload or;
- The update of an existing scoring system to measure nursing workload or;
- The validation or evaluation of an existing scoring system to measure nursing workload.
- The scoring system quantified the workload into the needed amount of nursing time based on points, classes, levels of care or absolute amount of time.
- The setting was an adult Intensive Care Unit.
- The language of the articles was English, German or Dutch.

We excluded articles about scoring systems without a quantification of the nursing time needed. We also excluded articles with Intensive Care Units situated in a burn center because the nursing care in a burn center is not comparable with other Intensive Care Units. References from reviews and included articles were checked on relevance and included if they met the inclusion criteria.

Data extraction from selected articles

The two reviewers (MH and CM) extracted all relevant information from the selected articles by filling in a data extraction form. This form contained the following items: name of the scoring system, study aim, country, setting and number of participants (Intensive Care Units, patients and nurses), kind of nursing interventions or activities measured by the scoring system, methods used to select nursing interventions or activities in case of development of a scoring system, methods used to measure reliability and validity of the system, results regarding reliability and validity and methods used to translate the workload measurement in needed nursing time.

Assessment of validity and reliability of scoring systems

For all included full papers the validity and reliability of the scoring systems were assessed using the following criteria. Content validity: we considered a scoring system content-valid when nursing professionals participated in the selection of interventions and activities included in the scoring system, and when expert-consensus in focus groups or Delphi rounds were used or when a Content Validity Index for the overall system was at least 0.9 (Polit and Beck, 2012, 2016).

Reliability: we assessed data on inter-rater reliability (level of agreement between the scores of different nurses scoring the nursing interventions of the same patient) and intra-rater reliability (level of agreement between assessment and reassessment of the nursing intervention scores of a patient by the same nurse). The following statistical tests and cut-off values were used for the assessment of the reliability: Cohen's Kappa and the Intra-Class Coefficient (ICC). For the Kappa we used the ranges of kappa according Landis and Koch meaning a value of 0.41–0.60 as moderate; 0.61–0.80 as substantial, and 0.81–1 as almost perfect agreement (Landis and Koch, 1977). For evaluation of the ICC we used the Cronbach's alpha with a cut-off point of 0.70 for an acceptable reliability (DeVon et al., 2007).

Validity: we defined the validity as to which extent interventions or activities of a scoring system actually measured the true outcome i.e. needed nursing time. We distinguished two methods to assess the validity:

1. By comparing the results of a scoring system with the 'gold standard' observed time-measurements.

2. By comparing a newly developed scoring system with an already existing system.

We considered method 2 a weaker method for validation. The following statistical methods were used for the assessment of the validity: linear regression equation (r2) and the Correlation Coefficient (Pearson's r or Spearman's rs). For interpretation of the results we did categorize the results as a weak (r/r2 < 0.25), moderate (0.25 ≤ r/r2 < 0.75) or strong (r/r2 ≥ 0.75) correlation (Knapp, 2000). We used the same methods to assess the validity of the translation of the measurement of nursing time into the need for nursing staff, often translated into a N:P-ratio.

M.E. Hoogendoorn, C.C. Margadant and S. Brinkman et al./International Journal of Nursing Studies 101 (2020) 103408

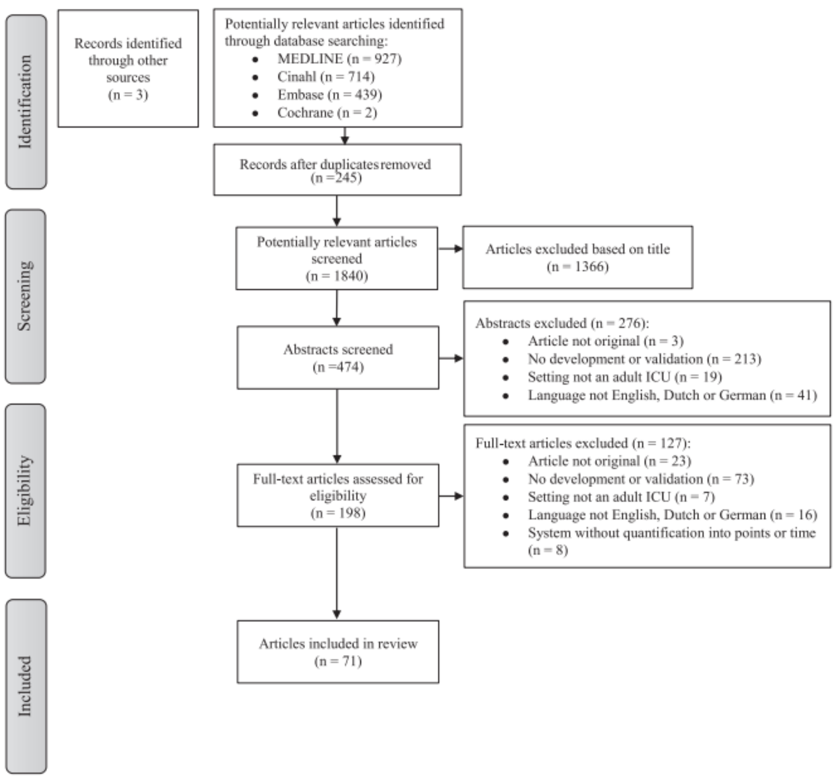


Figure 1. Prisma flow diagram

Table 1. Scoring systems.

Scoring systems (year 1st publication)	**Score per shift or per 24h**	**Main content**	**Scoring method**
Systems measuring nursing and/or therapeutic interventions			
1. TISS-76 (1974) (Cullen et al., 1974) Therapeutic Intervention Scoring System	Per 24h	Classification of medical interventions	- 76 medical interventions - 1–4 points per intervention - Total score categorized into four levels of care with expected Nurse:Patient-ratio
2. PRN-system (1978) (Chagnon et al., 1978a) Project Research in Nursing	Per 24h	Classification of nursing interventions	- 8 categories of nursing interventions - 35 tasks per category - 1–20 points per nursing task - 1 point is 5 min
3. NISS (1978) (Greenburg et al., 1978) Nursing Intervention Scoring System	Per shift	Classification of nursing interventions	- 15 categories of nursing interventions - 1–4-points scale of nursing care (preventive till compensatory nursing) - 1 point is 6.75 min
4. Classification system of Jackson Memorial Medical Centre (1979) (Hudson et al., 1979)	Per shift	Classification of nursing interventions	- 44 nursing interventions - 1–6 points score per intervention - 1 point is 4 min
5. NDS (1983) (Garfield et al., 2000) Nurse Dependency Scoring System	Per shift	Classification of nurse dependency	- 6 categories of nurse-dependency - Score of 0–4 per category - Total score categorized to different Nurse:Patient-ratios
6. Computerized Acuity System (1986) (Adams and Johnson, 1986)	Per shift	Classification of nursing activities	- Software program calculating direct and indirect nursing activities from a care plan - Calculation of points representing the nursing time per activity - 1 point is 1 min
7. PINI (1988) (Prescott and Philips, 1988) Patient Intensity for Nursing Index	Per shift	Classification of nursing intensity	- 12 dimensions of nursing care (i.e. complexity tasks, complications, mobility) - Level I–V per dimension, representing the complexity of nursing care (basic till intense/high or complex)

8. TOSS (1991) (Iapichino, 1991) Time Oriented Score System	Per 24h	Classification of nursing activities	- List of nursing 14 activities - Categories of estimated time per activity
9. NCR-11 (1992) (Hjortso et al., 1992) Nursing Care Recording System	Per shift	Classification of nursing activities	- Description of nursing contribution to 11 categories of nursing and medical procedures - 1–3 points per category, points representing intensity - Total points categorized to estimated nursing time
10. WICSS (1993) (van Aken and de Vries, 1993)	Per shift	Classification of nursing activities	- 111 nursing activities in direct and indirect nursing patient care - 1–20 points per activity - 1 point is 6 min
11. Acuity tool (1995) (Marsee et al., 1995)	Per 24h	Classification of nursing intensity	- Five categories of nursing care intensity; minimal till life-support - Estimated hours of nursing time per category
12. CCPD (1996) (Donoghue et al., 2001) Critical Care Dependency System	Per shift	Classification of nursing intensity	- 7 nursing activity related groups - 1–4 points of nursing intensity per activity, 1 is low and 4 is high intensity - Total points categorized to estimated nursing time
13. TISS-28 (1996) (Miranda et al., 1996) Therapeutic Intervention Scoring System	Per 24h	Classification of medical interventions	- Simplified version of the TISS-76 with 28 therapeutic medical interventions - 1–4 points per intervention. - Total score categorized into four levels of care with expected Nurse:Patient-ratio
14. CritScore (1996) (Scribante et al., 1996)	Per 24h	Classification of medical interventions	- 70 therapeutic medical interventions - 1–4 points per intervention - Total score categorized into four levels of care with expected Nurse:Patient-ratio
15. NEMS (1997) (Reis Miranda et al., 1997)	Per shift or per 24h	Classification of medical interventions	- Simplified version of the TISS-28 with 9 activities performed on an IC - 1–8 points per activity - Total score categorized into four levels of care with expected Nurse:Patient-ratio
16. Acuity System (1999) (Lacovara, 1999)	Per 24h	Classification of estimated nursing time	- Assignment of patient to level I–V description of the expected nursing time for level V

17. ICNSS (2001) (Pyykko et al., 2000) Intensive Care Nursing Scoring System	Per shift	Classification of nursing intensity	- 16 different health problems of patients - 1 point (preventive) – 4 points (compensatory) - Points representing intensity of nursing care - Total score categorized to different Nurse:Patient-ratios
18. Perroca's instrument (2002) (Perroca and Gaidzinski, 2002)	Per 24 h	Classification of nursing intensity	- Nine areas of the care process - complexity of care per area graded from 1to 4. - Total score categorized to levels of care with a description of expected nursing time for an ICU patient.
19. NAS (2003) (Miranda et al., 2003) Nursing Activity Score	Per shift or per 24 h	Classification of nursing activities	- 23 nursing activities - A score 1.2–32 points per nursing activity - Points representing the required nursing time per activity.
20. Nurse Workload (NWL)-Patient Category Scoring System (2003) (Adomat and Hicks, 2003)	Per 24 h	Classification of nursing and medical interventions	- Score based on the TISS with additional scores for therapeutic and nursing interventions. - Total score categorized to different N:P-ratios.
21. CNIS (2003) (Yamase, 2003) Comprehensive Nursing Intervention Score	Per 24 h	Classification of nursing interventions	- List of 73 nursing interventions - 4-grade workload score in 6 aspects per intervention: nursing time needed, number of nurses, muscular extension, mental stress, skill, job intensity
22. Workload indicator for Nursing (WiN)-score (2009) (Sermeus et al., 2009)	Per 24 h	Classification of nursing interventions	- A list of nursing interventions based on the Nursing Interventions Classification - Points representing estimated time per intervention - 1 point is 1 min
Scoring systems based on an expected N:P-ratio			
23. [No name] (1980) (Evans et al., 1980)	Per shift and per 24 h	Classification of Nurse:Patient-ratio	- Categorization of patients according the expected N:P-ratio based on nursing time.
24. SGI-Grading system of the Swiss Society of Intensive Care (1997) (Jakob and Rothen, 1997)	Per shift	Classification of Nurse:Patient-ratio	- Categorization of patients according to the estimated number of patients per nurse

25. American Association of Critical Care Nurses Synergy Model (1998) (Curley, 1998)	Per 24 h	Classification of Nurse:Patient-ratio	- Description of indicators for nursing time divided under 8 dimensions of patient care. - Categorization according Nurse:Patient-ratio - Description of patient care for 1:1-ratio
26. Time weighted nursing demand (2000) (Volpatti et al., 2000)	Per shift	Classification of Nurse:Patient-ratio	- Description of needed time for patient categories based on a N:P-ratio
27. Association of UK University Hospitals (AUKUH) Acuity Tool (2008) (Hurst et al., 2008)	Per 24 h	Classification of Nurse:Patient-ratio	- Description of patient criteria, based on patient dependency and nursing activities - Classification in four levels of care - Expected nursing time and Nurse-Patient ratio per level of care

Results search strategy

Fig.1 shows a PRISMA flowchart with the results of the literature search strategy. Starting with 1840 unique articles we finally we included 71 articles for analysis. Amongst these 71 included articles, 30 articles reported on the development of a scoring system to measure the amount of nursing care. Nineteen articles (also) reported on reliability of an existing scoring system. Forty four articles reported (also) on validation of a scoring system. In total 17 articles reported (also) on the validation of the translation of the measured nursing time into a Nurse:Patient-ratio for calculation of the need for nursing staff.

Methods used for classification of needed nursing time

From the 71 included articles, we identified 27 different scoring systems with a translation of workload into nursing time needed. Table 1 provides an overview of these 27 systems with the name and the abbreviation as used in daily practice and in this article, a description of their main content and the year of development. This table also shows that the way in which the needed nursing capacity is classified varies largely. There are differences in both content (nursing or medical interventions) and way of categorizing the care (points, time or Nurse:Patient-ratio). Twelve systems (44%) were based on a list of nursing interventions or a combination of nursing and medical interventions with either a description of minutes per intervention (n = 3), or points per intervention (n = 9). Those points were translated into minutes per point (n = 4) or translated into (a categorization of) expected nursing time per shift or an expected Nurse:Patient-ratio (n = 5). Nine systems (33%) were based on the level of dependency of the patient or a category of nursing care (i.e. preventive or minimal care to compensatory or intensive care), with a description of time in minutes or hours per category (n = 2), or points per category (n = 7). The points were translated either into minutes per point (n = 1) or translated into (a categorization of) expected nursing time per shift or an expected Nurse:Patient-ratio (n = 6). One system (4%) was based on a computerized calculation of activities from a care-plan with points per activity and a translation from points to minutes. In five systems

(19%) the classification was based on the expected Nurse:Patient-ratio with a description of the patient category per Nurse:Patient-ratio.

Validity and reliability of the scoring systems

Content validity

Information on the content validity was reported for 20 out of 27 (74%) scoring systems. A summary of the results of the content validity of the scoring systems is presented in Table 2.

In 17 of these 20 systems (85%), nurses participated in the selection and weighing of the interventions. For the TISS-76, the interventions were selected by physicians, but the actual weighing of the points was done by a team of physicians and nurses. The interventions included in the PINI were based on nineteen other scoring systems. The interventions included in the SGI-Grading System were based on a retrospective dataset without involvement of nurses. The Content Validity Index was only described for the Acuity-tool with a value of 0.85, which was lower than the considered threshold index of 0.9.

Reliability

Information about the inter- or intra-rater reliability was reported for 12 out of 27 (44%) scoring systems. A summary of the results of the reliability of the scoring systems is presented in Table 2. For 10 systems (37%) the inter-rater reliability and the intra-rater reliability were considered substantial to almost perfect (Cronbach's alpha 0.71–1.00, Kappa > 0.65). The results of the remaining 2 systems (PINI and NAS) showed varying results from slight to substantial agreement (Prescott and Philips, 1988; Prescott et al., 1991, 1989; Philips et al., 1992; Stuedahl et al., 2015). The interventions which include categories of a subjective estimation of time by the nurse (e.g. the hygienic procedures took more than 2 h per shift in NAS) showed lower reliability (Kappa of 0.02–0.12) (Stuedahl et al., 2015).

Table 2. Results on the validity and reliability of the scoring systems.

Scoring systems	Nr. Of articles	Content validity		Reliability		Validity		Results
		Score	**Explanation**	**Score**	**Results**	**Score**	**Compared with:**	
TISS-76	5	+/−	Selection interventions without nurses, weighting interventions with nurses (Cullen et al., 1974; Keene and Cullen, 1983)	+	-	NS	Time observations (Dick et al., 1992) TISS (Keene and Cullen, 1983) Non-ICU system: Oulu patient classification system (Lundgren-Laine and Suominen, 2006)	r = 0.89 r = 0.57
PRN-system	3	+	Selection interventions with nurses (Chagnon et al., 1978a)	NS	-	+/−	Actual staff ICU (Chagnon et al., 1978a) Non-ICU system (O'Brien-Pallas et al., 1989) - GRASP - Medicus	r2 = 0.77–0.90 r2 = 0.80–0.84
NISS	2	+	Selection interventions with nurses (Greenburg et al., 1978)	NS	-	+/−	Non-ICU system (Alexander et al., 1984)	r2 = 0.61
Classification system of Jackson Memorial Medical Centre	2	+	Selection interventions with nurses (Hudson et al., 1979)	NS	-	+/−	Time observations (Sheppard and Garland, 1983)	r not stated
NDS	1	NS	-	NS	-	+/−	TISS (Garfield et al., 2000)	r = 0.33
Computerized Acuity System	1	+	Selection interventions with nurses (Adams and Johnson, 1986)	NS	-	-	Non-ICU system for nursing diagnosis (Adams and Johnson, 1986)	No correlation, r not stated
PINI	4	-	Selection interventions based on 19 other systems, without nurses (Prescott and Philips, 1988)	+/−	Kappa 0.29–0.86 (Prescott and Philips, 1988; Prescott et al., 1991;	+/−	Time measurements (Prescott et al., 1991) Non-ICU systems (Philips et al., 1992) - Medicus	r = 0.75 r = 0.54–70

					Philips et al., 1992; Prescott et al., 1989)		- GRASP	
TOSS	1	+	Selection interventions with nurses (Iapichino, 1991)	NS	-	+/−	TISS (Iapichino, 1991)	r = 0.79
NCR-11	2	+	Selection interventions with nurses (Hjortso et al., 1992)	+	Coefficient of variation 5.9–10.4% (Walther et al., 2004)	+/−	TISS (Hjortso et al., 1992) NEMS (Walther et al., 2004)	r = 0.60 r not stated
WICSS	1	+/−	Selection interventions with nurses (van Aken and de Vries, 1993)	NS	-	NS	-	-
Acuity tool	1	+/−	Selection interventions with nurses, Content Validity Index 0.85 (Marsee et al., 1995)	+	Pearson's r2 = 0.84 (Marsee et al., 1995)	NS	-	-
CCPD	1	NS	-	+	Interrater reliability 74% Test-retest reliability 85% (Donoghue et al., 2001)	+/−	Comparing old and new system (Donoghue et al., 2001)	NS
TISS-28	9	+	Selection interventions with nurse (Miranda et al., 1996; Yee Kwok et al., 2005)	+	ICC: 0.93 – 1.00 (Yee Kwok et al., 2005; Moreno and Morais, 1997)	+/−	Time observations (Wysokinski et al., 2013) TISS-76 (Miranda et al., 1996; Yee Kwok et al., 2005; Moreno and Morais, 1997) NEMS (Rothen et al., 1999; Ducci et al., 2008; de Souza Urbanetto et al., 2014) SGI-Grading System (Rothen et al., 1999) NDS (Garfield et al., 2000) NAS (Ducci et al., 2008) SAPS (Yee Kwok et al., 2005)	Significant difference r = 0.78–0.96 r = 0.78 Good agreement r = 0.33 NAS>TISS, p = 0.001 R = 0.68

							Between ICUs and shifts (Wysokinski et al., 2010)	Kruskal Willis: H = 133.57
CritSlcore	1	+	Selection interventions with nurses (Scribante et al., 1996)	NS	-	NS	-	-
NEMS	11	+	Selection interventions from TISS with nurses (Reis Miranda et al., 1997)	+	ICC: 0.73–0.99 (Reis Miranda et al., 1997; Moreno and Reis Miranda, 1998; Junger et al., 2007; Perren et al., 2012)	+/−	TISS-28 (Reis Miranda et al., 1997; Rothen et al., 1999; de Souza Urbanetto et al., 2014) NAS (Ducci and Padilha, 2008; Stafseth et al., 2011; Carmona-Monge et al., 2013)NCR-11 (Walther et al., 2004) Case vignettes (Perren et al., 2014)	r = 0.78–0.88, r2 = 0.59–0.76 r = 0.16–0.93, r2 = 0.45–0.87 r not stated Accuracy 63.7–99.1%
Acuity System	1	NS	-	NS	-	+/−	Actual staff ICU (Lacovara, 1999)	r = 0.95
ICNSS	4	+	Selection interventions with nurses (Pyykko et al., 2000; Pyykko et al., 2001)	+	Kappa 0.81, Cronbach's alpha 0.91 (Pyykko et al., 2004a)	+/−	TISS (Pyykko et al., 2004a; 2004b)	r2 = 0.07–0.24
Perroca's instrument	3	NS	-	+	Cronbach's alpha 0.75–0.94 (de Fatima Salvini and Perroca, 2015; Cabral Ferreira et al., 2017)	+/−	1st and 2nd version (Perroca and Ek, 2007) NAS (de Fatima Salvini and Perroca, 2015) Non-ICU system: Beakta (Cabral Ferreira et al., 2017)	r = 0.60 r = 0.65 r = 0.60–0.83

NAS	12	+	Selection interventions with nurses (Miranda et al., 2003; Palese et al., 2016)	+/−	Kappa 0.02–0.69 (Stuedahl et al., 2015; Macedo et al., 2016) Cronbach's alpha 0.71 (Debergh et al., 2012)	+	Time observations (Miranda et al., 2003) Prospective/retrospective (Ducci et al., 2008) NAS/shift (Debergh et al., 2012; Armstrong et al., 2010) TISS (Miranda et al., 2003; Ducci and Padilha, 2008) NEMS (Ducci et al., 2008; Stafseth et al., 2011; Carmona-Monge et al., 2013; Hellin Gil et al., 2017) Perroca's instrument (Cabral Ferreira et al., 2017)	r = 0.81 r = 0.65 not stated r = 0.56 r = 0.16–0.93 r = 0.65
NWL- Patient Category Scoring System	1	NS	-	NS	-	+/-	Time observations (Adomat and Hicks, 2003)	NWL does not reflect nursing time, r not stated
CNIS	1	+	Selection interventions with nurses (Yamase, 2003)	+	Kappa 0.65 (Yamase, 2003)	+/−	NEMS (Yamase, 2003)	r = 0.75
WiN-score	2	+	Selection interventions with nurses (Sermeus et al., 2009)	NS	-	+/−	TISS (Sermeus et al., 2009) B-NMDS (Myny et al., 2014)	r = 0.74 r = 0.89
[No name]	1	+	Selection interventions with nurses (37)	NS	-	+/−	Time observations (Evans et al., 1980)	No significant differences, results not stated
SGI – Grading system	2	-	Based on dataset, no nurses involved (Jakob and Rothen, 1997)	NS	-	+/−	TISS, NEMS (Rothen et al., 1999)	r = 0.78, r2 = 0.62

AACN Synergy Model	4	+	Description ratio by nursing experts [(Australian College of Critical Care Nurses Ltd (ACCCN) 2002; Curley, 1998; Hartigan, 2000; Kohr et al., 2012)]	NS	-	+/−	Patient acuity-indicators (Brewer, 2006)	r = 0.30–0.60
Time weighted nursing demand	1	NS	-	NS	-	+/−	Actual staff (Volpatti et al., 2000)	r2 = 0.83
AUKUH-dependency tool	1	NS	-	+	Cronbach's alpha 0.99 (Hurst et al., 2008)	-	Non-ICU Leeds system (Hurst et al., 2008)	r = 0.77

Legend validity and reliability

Content validity		**Reliability**		**Validity**	
+	- Development with nurses AND - Content Validity Index at least 0.9 (if reported)	+	- Kappa > 0.60 OR - ICC > 0.70	+	- Validated with gold standard; time-observations and strong correlation (r ≥ 0.75)
+/-	- Development with nurses AND - Content Validity Index < 0.9 (if reported)	+/−	- Kappa < 0.60 OR - ICC < 0.70	+/−	- Validated with gold standard; time observations and moderate correlation (r ≥ 0.25, <0.75) OR - Validated with another ICU nursing workload system and moderate or strong correlation (r>0.25)
-	- Development without nurses -	-	Kappa < 0.20	-	- Validated with another system for ICU AND weak correlation (<0.25) - Validated with another system not for ICU AND weak correlation (<0.25)
NS	- Not stated	NS	Not stated	NS	- Not stated

Validity

Information about the validity was reported for 24 of the 27 (89%) scoring systems. A summary of the results of the validity of the scoring systems is presented in Table 2. The 'gold standard', observed time-measurement, was used in only 7 (26%) scoring systems. Although the TISS was originally (in 1974) developed without the use of continuous time-measurements, we found one study, published in1992, in which the TISS was retrospectively evaluated using continuous time-measurements (Dick et al., 1992). A strong correlation was shown between the time for nursing interventions and the TISS-76 ($r = 0.89$, $p < 0.0001$). The Classification System of the Jackson Memorial Medical Centre was developed and evaluated with continuous time-observations. It was concluded that the point-system was a good indicator of the actual care received (Sheppard and Garland, 1983). The PINI was validated with an observational time measurement study (Prescott et al., 1991). A strong correlation was found between the observed time and the rated hours of care ($r = 0.75$, $p < 0.001$). In 70% of the disagreements, nurses overestimated the hours of care. The NAS was validated with Multi Moment Recordings; 81% of the total time spent by nurses was explained by the NAS (Miranda et al., 2003). The NWL-Patient Category Scoring System was validated by comparing the results of the scoring system with time-measurements by video-observation. They concluded that this scoring system did not give an accurate reflection of the amount of nursing time (Adomat and Hicks, 2003). The system described by Evans et al. (No name) was validated with time observations; the expected needed hours per shift was compared with the observed hours per shift per category (Evans et al., 1980). They concluded that the expected and observed nursing care hours were equal, except for category II patients. This category expected 8 h nursing care per shift where 5.3 h nursing care were observed. The weaker method for validation, i.e. comparing the newly developed scoring system with an existing scoring system, was described for 16 scoring systems (59%). As we can see in Table 2, most studies ($n = 10$) used the TISS for this comparison. One study used case-vignettes for the evaluation of the validity (Perren et al., 2014).

Validity of the quantification of the nursing time needed

The way in which the workload systems quantify the needed nursing time and the validity of this quantification is described in Table 3. The most common way is classification of care into different categories of

Nurse:Patient-ratios. Any evaluation of the validity of these categories of Nurse:Patient-ratios was described for 15 out of 27 systems (56%).

In three cases (11%) the number of nurses needed according to the calculated Nurse:Patient-ratio was compared with actual time measurements. The calculated need for nursing staff according to TISS or PINI was higher than the measured nursing time (Prescott et al., 1991; Dick et al., 1992). Comparing the Nurse:Patient-ratio according to the NWL Patient Category Scoring System with the observed time-measurements, showed substantial differences. The time spent with a patient in the category with an expected Nurse:Patient-ratio of 0.5:1 was more than the time spent with a patient in the category with an expected Nurse:Patient-ratio of 1:1. The researchers concluded that the categorization according the NWL Patient Category Scoring System was not accurate (Adomat and Hicks, 2003).

In five systems (19%) the translation of scores into a N:P-ratio was evaluated by comparing different systems applied to the same patients. A good agreement was reported between the Nurse:Patient-ratio according to the TISS and NEMS; which is not surprising because the NEMS was developed based on the TISS (Dick et al., 1992). The need for nursing staff according to TOSS was up to 52% higher than with TISS. Where TISS indicated a Nurse:Patient-ratio of 2:1 TOSS indicated in the same patient a Nurse:Patient-ratio of 3:1 (Iapichino, 1991). Also in ICNSS the need for nurses was higher than in TISS for the same patients (Pyykko et al., 2004b). The need for nurses according to NAS was higher than the need for nurses according to NEMS (Carmona-Monge et al., 2013).

Table 3 shows a retrospective comparison of the Nurse:Patientratio with the actual or planned number of nurses in 11 systems (41%) (Hudson et al., 1979; Hjortso et al., 1992; Lacovara, 1999; Volpatti et al., 2000; Alexander et al., 1984; Ducci et al., 2008; Moreno and Reis Miranda, 1998; Chagnon et al., 1978b; Scribante and Bhagwanjee, 2007; Jie et al., 2013). The 'midnight census' on planning the actual number of nurses per unit was also described as a method for classification of care (Volpatti et al., 2000; Baernholdt et al., 2010). For five systems it was concluded that the need for nurses according to the system was higher than the actual present staff (Hudson et al., 1979; Iapichino, 1991; Adomat and Hicks, 2003; Prescott et al., 1991; Scribante and Bhagwanjee, 2007). However, none of these articles mentioned how the actual or planned number of nurses was calculated and on which assumptions these numbers were based.

Discussion

This review shows that over the years classification of nursing workload on an Intensive Care has been a topic of continuous attention. Our aim was to identify the existing scoring systems in literature. A high number of scoring systems has been developed and used for planning of care since the first system was published in 1974. In total we included 27 different systems for measuring nursing workload on an Intensive Care in this review. Remarkable is the continuous use of this first developed system, the TISS. Although developed in 1974, the TISS is still used in daily Intensive Care practice as well as for development and validation of other scoring systems. Table 1 shows an increasing number of new systems between 1980 and 2000. The continuous use of those systems since their development shows that quantification of nursing care is still actual and considered important.

The next important part of this review was the evaluation of the validity and the reliability of the scoring systems. Although we found many articles about validation and reliability of the different systems, none of the finally included 27 systems that claimed to quantify needed nursing staff satisfied all our pre-set validity and reliability criteria. Only a few satisfied a majority of our pre-set criteria. The content validity of almost all 27 included systems was good; most systems were developed by nurses or a multidisciplinary team of nurses and physicians. Only the items of the TISS concerned mainly medical interventions exclusively selected by physicians, which can be explained by the fact that the original aim of this system was to classify severity of illness and not nursing workload (Cullen et al., 1974). It is therefore remarkable that the TISS has become one of the most commonly used scoring systems to measure nursing workload. Moreover, he TISS itself or items of the TISS have also been used in the development of six other systems for measuring nursing workload, namely the NISS, TISS-28, CritScore, NEMS, NAS, and the NWL Patient Category Scoring System.

The inter- and/or intra-rater reliability was tested in less than half of the systems (44%). If described, the results were generally moderate to good. Only the results on the inter-rater reliability of the NAS evaluated in several different studies showed a large variability with weak to good results (Stuedahl et al., 2015). In particular the inter-rater reliability of nursing activities which included an estimate of the duration of that activity, such as monitoring and titration, hygiene procedures, support and care of patient and relatives, administrative tasks and mobilization, showed a large variability. For example, the inter-rater reliability of the item “Mobilization and positioning” resulted in an agreement of 49% (Kappa 0.16) if rated by a nurse and a manager. If rated by a nurse and a physician, the agreement was 39,6% (Kappa 0.020) (Stuedahl et al., 2015). On the other hand, a medical intervention like oxygen showed a 100% agreement. Because the duration of the activity has to be estimated, the assessment is partly subjective. This subjective estimation can lead to differences in NAS-scores and subsequently to differences in the calculated need for nursing staff.

We indicated the use of time-measurements or Multi Moment Recordings as the gold-standard for the development and validation of a system to quantify nursing time needed. This method was, however, only used in six of the 27 (22%) scoring systems. For all those systems the results showed a good validity. The most common method used for validation of a new system was the comparison with an already existing system. TISS was most frequently used for this purpose. Although the TISS-76 was developed without time measurements, a later version of TISS-76 was in fact validated with the gold standard (i.e. time-measurements), but not before 1992 (Cullen et al., 1974; Dick et al., 1992). Despite the lack of formal validation, the TISS-76 was already used as a reference in validation studies of other systems before 1992.

Overall the NAS performed best as it was developed by nurses, validated with time-measurements and explaining 80% of the nursing activities. The reliability varied between low to good. The studies which reported low reliability explicitly evaluated the reliability of scoring systems with categories of estimated time per intervention. This can be explained by the psychometric properties of these questions. The answers on subjective questions are more influenced by external factors as the involved observer self and knowledge of the definitions of those questions. Education and training in the use of the NAS is therefore necessary for a better use of this system. Furthermore, as more and more Intensive Cares are equipped with electronic patient records or patient data management systems, automatic bedside registration in an electronic patient record could also lead to more unambiguous scoring and improved reliability of the NAS.

Translation of a scoring system into another language is also known to influence the reliability (Hilton and Skrutkowski, 2002). This is important for the NAS, because the NAS is widely used, amongst others, in countries with Portuguese language. We found one study reporting about the psychometric properties of a translated Portuguese version of the NAS. This study concluded that the Portuguese version of the NAS was found to be a valid instrument (Macedo et al., 2016). One study in 7 different countries in Europe and

Brazil showed a large variation in NAS scores, ranging from a mean NAS per patient of 101.8 in a Norwegian Intensive Care to 44.5 in a Spanish Intensive Care (Padilha et al., 2015). We recognize this variation also in other studies included in our review. This could partly be explained by the fact that the studies are conducted in different countries with different organization structures of the Intensive Cares and different patient characteristics. Although all studies used the standardized NAS-scoring system, these differences still make reliable comparisons between the studies more difficult.

Finally we evaluated the ability of the workload scoring systems to quantify the nursing capacity in daily practice, mostly translated into a Nurse:Patient-ratio. The Nurse:Patient-ratio is important because this is the translation of 'abstract' points into nursing capacity in daily practice, and can so be helpful for Intensive Care management. It enables to plan the needed numbers of nurses' per shift and it enables a nurse to know at the beginning of his/her shift how many patients are under his/her responsibility. The Nurse:Patient-ratio was validated for only half of the systems (56%), of which only three systems used objective time-measurements (11%). Given the fact that the ultimate aim of a scoring system should be supporting the planning of nursing capacity, it is disappointing that the accuracy of translating the scores of a system into needed nursing time was only assessed to a limited extent and even then often in an inadequate way. Comparison with time-measurements is only described for the TISS, the PINI and the NWL Patient Category Scoring System. In all three studies the categorization to a Nurse:Patient-ratio led to an overestimation of needed time (Adomat and Hicks, 2003; Prescott et al., 1991; Dick et al., 1992). Comparison amongst different scoring systems also gave disappointing results with large differences in the reported Nurse:Patient-ratio, with examples of doubling the needed nursing staff (Pyykko et al., 2004b). In a number of articles, the calculated Nurse:Patient-ratio from a scoring system was compared to the actual available or planned number of nurses. However, a description of how the actual or planned number of nurses was determined was lacking. This information is crucial to interpret the results of the comparisons made. Without a validation by time-measurements it is impossible to assess the accuracy of both the actual planned staff as well as the planned staff according to a scoring system.

Implications for research

Regarding the validation of the systems, the low number of systems that were validated with the gold standard, i.e. time measurements, is striking. The implications of the absence of the gold standard becomes clear when interpreting the results of the second-best method for validation; comparing two different systems often show large variation (Iapichino, 1991; Miranda et al., 2003; Rothen et al., 1999; Ducci et al., 2008; de Souza Urbanetto et al., 2014). In these cases it is hard to tell which scoring system agrees with reality, due to the absence of time-measurements. Studies in which the systems-based nursing capacity was compared with the actual nursing staff show the same weakness. A higher indication for needed nursing staff by a system compared to the actual present staff would suggest that the workload of the nurses is too high and should be lowered. However, without information on the accuracy of both the actual planned and system-based calculated nursing staff, this conclusion cannot be made with certainty. Therefore, studies with time-measurements for both the systems and the Nurse:Patient-ratio should be performed before any implications for practice and actions to improve the practice can be made. New time-measurements should also be done for systems still in use but without any update in the last decades, i.e. the TISS, because daily care and treatment may have changed considerably over the years. Finally another way of assessing the accuracy of the planned staff is comparing the calculated workload with the subjective workload as experienced by nurses. Future studies should focus on how an objectively calculated workload with a workload-system correlates with the subjective workload as experienced by nurses.

Implications for practice

It is clear that the variety in calculated and overestimation of needed nursing staff could have large consequences for the actual planning of nursing staff. A scoring system should be able to quantify the need for nursing time as accurately as possible to be of any use as a tool for planning nursing staff. The conflicting results and lack of thorough validation make the scoring systems less useful for management decisions. Considering the results of the evaluation of the Nurse:Patient-ratio, the added value of a categorization into a Nurse:Patient-ratio with a system is debatable. If a system is able to measure the actual time needed for nursing care the needed number of nurses can be determined without such a calculated Nurse:Patient-ratio. An accurate calculation of the nursing time needed for certain patient categories should make it possible, on average, to plan the correct number of nurses. It could be that not the Nurse:Patient-ratio, but the workload per patient and therefor per nurse is important for management decisions. This adds to the value of the NAS; with the NAS-points it is possible to calculate the need for nursing time in minutes. The NAS is not calculating an Nurse:Patient-ratio. From a management perspective the balance between needed nursing time according NAS-points and available nursing time in NAS-points is sufficient to measure the workload and calculate the need for nursing staff.

Strengths and limitations

This study has some strengths and weaknesses. A strength of this study is that we used all relevant literature databases over a long period of time. The titles, abstracts and articles were independently assessed by two different reviewers and inclusion was based on consensus of both reviewers.

The included articles cover a period of more than 40 years. It is quite unusual for systematic reviews to include articles over such a long period as the relevance of this literature might become debatable. During a period of 40 years the nursing care on an Intensive Care is changed due to a changing patient population, development of new techniques and organizational changes. However, the first system developed in 1974, the TISS, including the translation of points into a Nurse:Patient-ratio, is still used in current practice. Therefore, it is important to update or validate systems, if still in use after such a period of time.

Because the limitation to articles in English, German or Dutch, we did exclude a relatively high number of articles for further analysis (n = 57). Amongst these excluded articles are a substantial number of articles on the NAS which were written in the Portuguese or Spanish language. Despite this high number of excluded NAS studies, the NAS is still well represented in the results of our review (n = 12). Therefore, we believe that the most important systems are represented in our review.

Conclusion

Scoring systems for measuring nursing workload and calculating the needed nursing staff on an Intensive Care received a lot of attention over the years. A range of systems has been developed and is still in use in daily practice of Intensive Care management. Overall, NAS performed best; it is the only system with good content-validity and Multi Moment Recordings showed that 81% of total time spent by nurses could actually be explained. The results of this review showed that the NAS is the most used system for measuring nursing workload. However, the intra- and inter-rater reliability evaluation of NAS showed a need for improvement.

Given the insufficient evaluation methods and results regarding the validity and reliability of most scoring systems we conclude that the value of these systems to plan nursing capacity in practice is debatable. Due to the important role of workload scoring systems for Intensive Care management, further research is needed to improve the reliability of scoring and the accuracy of the translation of the scores into the actual needed nursing time.

References

1. Adams, R., Johnson, B., 1986. Acuity and staffing under prospective payment. J. Nurs. Adm. 16, 21–25.
2. Adomat, R., Hicks, C., 2003. Measuring nursing workload in intensive care: an observational study using closed circuit video cameras. J. Adv. Nurs. 42, 402–412.
3. Adviescommissie Kwaliteit van het Zorginstituut, Kwaliteitsstandaard Organisatie van Intensive Care, 2016.
4. Alexander, R.H., Rumley, T.O., Blikken, W., 1984. Nursing intervention scoring system in management of a surgical intensive care unit. South Med. J. 77, 1248–1252.
5. Aragon Penoyer, D., 2010. Nurse staffing and patient outcomes in critical care: a concise review. Crit. Care Med. 38, 1521–1528.
6. Armstrong, E.K., Heymans, M.W., Girbes, A.R.J., Spijkstra, J.J., 2010. Using nursing activity scores (NAS) to assess nursing workload on medium care (MC) units. Intens. Care Med. 36, S395.
7. Australian College of Critical Care Nurses Ltd (ACCCN), 2002. Position statement on intensive care nursing staffing. Aust. Crit. Care 15, 6–7.
8. Baernholdt, M., Cox, K., Scully, K., 2010. Using clinical data to capture nurse workload: implications for staffing and safety. Cin. Comput. Inform. Nurs. 28, 229–234
9. Brewer, B.B., 2006. Is patient acuity a proxy for patient characteristics of the AACN synergy model for patient care? Nurs. Adm. Q 30, 351–357.
10. Cabral Ferreira, P., Machado, R.C., Soares Martins, Q.C., Sampaio, S.F., 2017. Classification of patients and nursing workload in intensive care: comparison between instruments. Rev. Gauch. Enferm. 38, 1–7.
11. Carmona-Monge, F.J., Rollan Rodriguez, G.M., Quiros Herranz, C., Garcia Gomez, S., Marin-Morales, D., 2013. Evaluation of the nursing workload through the nine equivalents for nursing manpower use scale and the nursing activities score: a prospective correlation study. Intens. Crit. Care Nurs. 29, 228–233.
12. Chagnon, A.G., Audette, L.M., Lebrun, L., Tilquin, C.A., 1978a. A patient classification system by level of nursing requirements. Nurs. Res. 27, 107–113.
13. Chagnon, M., Audette, L.M., Lebrun, L., Tilquin, C., 1978b. Validation of a patient classification system by level of nursing care requirements. Med. Care 16, 465–475.
14. Costa, D.K., 2016. Organizing critical care for the 21st century. JAMA: J. Am. Med. Assoc. 315, 751–752.
15. Cullen, D.J., Civetta, J.M., Briggs, B.A., Ferrara, L.D., 1974. Therapeutic intervention scoring system: a method for quantitative comparison of patient care. Crit. Care Med. 2, 57–61.
16. Curley, M.A.Q., 1998. Patient-nurse synergy; optimizing patients' outcome. Am. J. Crit. Care 7, 64–72.
17. de Fatima Salvini, M., Perroca, M.G., 2015. Comparative study between two instruments for classifying patients. Cienc. Cuid. Saude 14, 1498–1504.

18. de Souza Urbanetto, J., Travi Canabarro, S., Prado Lima Figueiredo, A.E., et al., 2014. Correlation between the TISS-28 and NEMS indicators in an intensive care unit. Int. J. Nurs. Pract. 20, 375–381.
19. Debergh, D.P., Myny, D., Van Herzeele, I., Van Maele, G., Reis Miranda, D., Colardyn, F., 2012. Measuring the nursing workload per shift in the ICU. Intens. Care Med. 38, 1438–1444.
20. DeVon, H.A., Block, M.E., Moyle-Wright, P.M., et al., 2007. A psychometric toolbox for testing validity and reliability. J. Nurs. Scholarsh. 39, 155–164.
21. Dick, W., Pehl, S., Tzanova, I., Heinrichs, W., Brost, F., Eich, P., 1992. Physician and nursing (personnel) requirements for ICUs. Therapeutic intervention scoring system (TISS) versus time requirements for patient care – a comparative study in an interdisciplinary surgical intensive care unit. Clin. Intens. Care 3, 116– 121.
22. Donoghue, J., Decker, V., Mitten-Lewis, S., Blay, N., 2001. Critical care dependency tool: monitoring the changes. Aust. Crit. Care 14, 56–63.
23. Driscoll, A., Grans, M.J., Caroll, D., et al., 2018. The effect of nurse-to-patient ratios on nurse-sensitive patient outcomes in acute specialist units: a systematic review and meta-analysis. Eur. J. Cardiovasc. Nurs. 17, 6–22.
24. Ducci, A.J., Padilha, K.G., 2008. Nursing activities score: a comparative study about retrospective and prospective applications in intensive care units. Acta Paul. Enferm. 21, 581–587.
25. Ducci, A.J., Zanei, S.S., Whitaker, I.Y., 2008. Nursing workload to verify nurse/patient ratio in a cardiology ICU. Rev. Esc. Enferm. USP 42, 673–680.
26. Evans, S.K., Laundon, T., Yamamoto, W.G., 1980. Projecting staffing requirements for intensive care units. J. Nurs. Adm. 10, 34–42.
27. Garfield, M., Jeffrey, R., Ridley, S., 2000. An assessment of the staffing level required for a high-dependency unit. Anaesthesia 55, 137–143.
28. Greenburg, A.G., McClure, K.D., Cheryl, A., Janus, C.A., Stubbs, J.A., 1978. Nursing intervention scoring system: a concept for management, research and communication. Med. Inf. Eur. 7, 738.
29. Hartigan, R.C., 2000. The synergy model. Establishing criteria for 1:1 staffing ratios. Crit. Care Nurse 20 (112), 114–116.
30. Hellin Gil, M.F., Lopez Montesinos, M.J., Seva Llor, A.M., Ferrer Bas, M.P., Macia Soler, M.L., 2017. Assessing the adequacy of workload measurement tools using a quality-based methodology. Rev. Brasil Enferm. 70, 39–46.
31. Hilton, A., Skrutkowski, M., 2002. Translating instruments into other languages: development and testing processes. Cancer Nurs. 15, 1–7.
32. Hjortso, E., Buch, T., Ryding, J., Lundstrom, K., Bartram, P., Dragsted, L., Qvist, J., 1992. The nursing care recording system. a preliminary study of a system for assessment of nursing care demands in the ICU. Acta Anaesthesiol. Scand. 36, 610–614.
33. Hudson, J., Caruthers, T.E., Lantiegne, K., 1979. Intensive care nursing requirements: resource allocation according to patient status. Crit. Care Med. 7, 69–75.

34. Hurst, K., Smith, A., Casey, A., Fenton, K., Scholefield, H., Smith, S., 2008. Calculating staffing requirements. Nurs. Manag. 15, 26–34.
35. Iapichino, G., 1991. Time oriented score system (TOSS): a method for direct and quantitative assessment of nursing workload for ICU patients. Intens. Care Med. 17, 340–345.
36. Jakob, S.M., Rothen, H.U., 1997. Intensive care 1980–1995: change in patient characteristics, nursing workload and outcome. Intens. Care Med. 23, 1165– 1170.
37. Jie, X., Weiquan, L., Minghao, F., Sufang, H., Hui, W., Xiuyun, L., 2013. Application value of ICU nursing scoring system in nurse allocation. Minerva Med. 104, 325–332.
38. Junger, A., Hartmann, B., Klasen, J., Brenck, F., Rohrig, R., Hempelmann, G., 2007. Impact of different sampling strategies on score results of the nine equivalents of nursing manpower use score (NEMS). Methods Inf. Med. 46, 410– 415.
39. Keene, A.R., Cullen, D.J., 1983. Therapeutic intervention scoring system: update 1983. Crit. Care Med. 11, 1–3.
40. Knapp, R., 2000. Basic Statistics for Nurses. Cengage Learning.
41. Kohr, L.M., Hickey, P.A., Curley, M.A., 2012. Building a nursing productivity measure based on the synergy model: first steps. Am. J. Crit. Care 21, 420–430.
42. Lacovara, J.E., 1999. Does your acuity system come up short? Nurs. Manag. 30 40A–40C.
43. Landis, J.R., Koch, C.G., 1977. The measurement of observer agreement for categorical data. Biometrics 33, 159–174.
44. Lee, A., Cheung, Y.S.L., Joynt, G.M., Leung, C.C.H., Wong, W.T., Gomersall, C.D., 2017. Are high nurse workload/staffing ratios associated with decreased survival in critically ill patients? A cohort study. Ann. Intensiv. Care 7, 1–9.
45. Liberati, A., Altman, D.G., Tetzlaff, J., Mulrow, C., Gøtzsche, P.C., Ioannidis, J.P., Clarke, M., Devereaux, P.J., Kleijnen, J., Moher, D., 2009. The PRISMA statement for reporting systematic reviews and meta analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 339, b2700. doi:10.1136/bmj.b2700, Jul 21.
46. Lundgren-Laine, H., Suominen, T., 2006. Nursing intensity and patient classification in a patient information system. Stud. Health Technol. Inform. 122, 894.
47. Marsee, V.D., Lovett, R.B., McMillan, S.C., 1995. Validity and reliability of an oncology critical care patient acuity tool. Oncol. Nurs. Forum 22, 967–971.
48. Miranda, D.R., de Rijk, A., Schaufeli, W., 1996. Simplified therapeutic intervention scoring system: the TISS-28 items – results from a multicenter study. Crit. Care Med. 24, 64–73.
49. Miranda, D.R., Nap, R., de Rijk, A., Schaufeli, W., Iapichino, G., 2003. TISS working group therapeutic intervention scoring system. Nursing activities score. Crit. Care Med. 31, 374–382.
50. Macedo, A.P., Mendes, C.M., Candeias, A.L., Sousa, M.P., Hoffmeister, L.V, Lage, M.I., 2016. Validation of the nursing activities score in Portuguese intensive care units. Rev. Bras. Enferm. 69, 826–832.

51. Moreno, R., Morais, P., 1997. Validation of the simplified therapeutic intervention scoring system on an independent database. Intens. Care Med. 23, 640–644.
52. Moreno, R., Reis Miranda, D., 1998. Nursing staff in intensive care in Europe: the mismatch between planning and practice. Chest 113, 752–758.
53. Myny, D., Bacquer, D., Hecke, A., Beeckman, D., Verhaeghe, S., Goubergen, D., 2014. Validation of standard times and influencing factors during the development of the workload indicator for nursing. J. Adv. Nurs. 70, 674–686.
54. Neuraz, A., Guerin, C., Payet, C., et al., 2015. Patient mortality is associated with staff resources and workload in the ICU: a multicenter observational study. Crit. Care Med. 43, 1587–1594.
55. O'Brien-Pallas, L., Leatt, P., Deber, R., Till, J., 1989. A comparison of workload estimates using three methods of patient classification. Can. J. Nurs. Adm. 2, 16–23.
56. Padilha, K.G., Stafseth, S., Solms, D., Hoogendoom, M., Monge, F.J., Gomaa, O.H., Giakoumidakis, K., Giannakopoulou, M., Gallani, M.C., Cudak, E., Nogueira Lde, S., Santoro, C., Sousa, R.C., Barbosa, R.L., Miranda Ddos, R., 2015. Nursing activities score: an updated guideline for its application in the intensive care unit. Rev. Esc. Enferm. USP 49 Spec No, 131–137.
57. Palese, A., Comisso, I., Burra, M., DiTaranto, P.P., Peressoni, L., Mattiussi, E., Lucchini, A., 2016. Nursing activity score for estimating nursing care need in intensive care units: findings from a face and content validity study. J. Nurs. Manag. 24, 549–559.
58. Perren, A., Previsdomini, M., Perren, I., Merlani, P., 2012. High accuracy of the nine equivalents of nursing manpower use score assessed by critical care nurses. Swiss Med. Wkly. 142, w13555.
59. Perren, A., Cerutti, B., Merlani, P., et al., 2014. SwissScoring – a nationwide survey of NEMS assessing practices and its accuracy. Acta Anaesthesiol. Scand. 58, 478–486.
60. Perroca, M.G., Ek, C., 2007. Assessing patients' care requirements: a comparison of instruments. Scand. J. Caring Sci. 21, 390–396.
61. Perroca, M.G., Gaidzinski, R.R., 2002. Perroca's instrument for patient classification: reliability test for evaluators agreement-correlation. Rev. Esc. Enferm. USP 36, 245–252.
62. Philips, C.Y., Castorr, A., Prescott, P.A., Soeken, K., 1992. Nursing intensity. Going beyond patient classification. J. Nurs. Adm. 22, 46–52.
63. Polit, D.F., Beck, C.T., 2012. Nursing Research: Generating and Assessing Evidence for Nursing Practice. Lippincott Williams & Wilkins, Philadelphia, USA.
64. Polit, D.F., Beck, C.T., 2016. The content validity index: are you sure you know what's being reported? Critique and recommendations. Res. Nurs. Health 29, 489–497.
65. Prescott, P.A., Philips, Y.C., 1988. Gauging nursing intensity to bring costs to light. Nurs. Health Care 9, 17–22.
66. Prescott, P.A., Soeken, K.L., Ryan, J.W., 1989. Measuring patient intensity. Eval. Health Prof. 12, 255–269.

67. Prescott, P.A., Ryan, J.W., Soeken, K.L., Castorr, A.H., Thompson, K.O., Phillips, C.Y., 1991. The patient intensity for nursing index: a validity assessment. Res. Nurs. Health 14, 213–221.
68. Pyykko, A.K., Laurila, J., Ala-Kokko, T.I., Hentinen, M., Janhonen, S.A., 2000. Intensive care nursing scoring system. Part 1: classification of nursing diagnoses. Intens. Crit. Care Nurs. 16, 345–356.
69. Pyykko, A.K., Laurila, J., Ala-Kokko, T.I., Hentinen, M., 2001. Intensive care nursing scoring system part 2: nursing interventions and nursing outcomes. Intens. Crit. Care Nurs. 17, 16–27.
70. Pyykko, A.K., Ala-Kokko, T.I., Laurila, J.J., Miettunen, J., Finnberg, M., Hentinen, M., 2004a. Validation of the new intensive care nursing scoring system (ICNSS). Intensive Care Med. 30, 254–259.
71. Pyykko, A.K., Ala-Kokko, T.I., Laurila, J.J., Miettunen, J., Finnberg, M., Hentinen, M., 2004b. Nursing staff resources in direct patient care: comparison of TISS and ICNSS. Acta Anaesthesiol. Scand. 48, 1003–1005.
72. Reis Miranda, D., Moreno, R., Iapichino, G., 1997. Nine equivalents of nursing manpower use score (NEMS). Intens. Care Med. 23, 760–765.
73. Rothen, H.U., Kung, V., Ryser, D.H., Zurcher, R., Regli, B., 1999. Validation of "nine equivalents of nursing manpower use score" on an independent data sample. Intens. Care Med. 25, 606–611.
74. Scribante, J., Bhagwanjee, S., 2007. ICU nurse allocation in a cardiothoracic ICU: how many hands do we need? S. Afr. J. Crit. Care 23, 66–69.
75. Scribante, J., Muller, M.E., Lipman, J., 1996. Development and validation of a critical care patient classification system. Am. J. Crit. Care 5, 282–288.
76. Sermeus, W., Gillet, P., Gillain, D., et al., 2009. Development and validation of nursing resource weights for the Belgian nursing minimum dataset in general hospitals: a Delphi questionnaire survey approach. Int. J. Nurs. Stud. 46, 256–267.
77. Sheppard, R.E., Garland, S., 1983. Staffing the intensive care unit. Hosp. Top 61, 16–18.
78. Stafseth, S.K., Solms, D., Bredal, I.S., 2011. The characterisation of workloads and nursing staff allocation in intensive care units: a descriptive study using the nursing activities score for the first time in Norway. Intens. Crit. Care Nurs. 27, 290–294.
79. Stuedahl, M., Vold, S., Klepstadt, P., Stafseth, S.K., 2015. Interrater reliability of nursing activity score among intensive care unit health professionals. Rev. Esc. Enferm. USP 49, 122.
80. Tan, S.S., Bakker, J., Hoogendoorn, M.E., et al., 2012. Direct cost analysis of intensive care unit stay in four European countries: applying a standardized costing methodology. Value Health 15, 81–86.
81. van Aken, G.H., de Vries, G., 1993. WICSS, or the weezenlanden intensive care scoring system. A reliable workload measuring tool for the ICU. TVZ 6, 318–322.
82. Volpatti, C., Leathley, M., Walley, K.R., Dodek, P.M., 2000. Time-weighted nursing demand is a better predictor than midnight census of nursing supply in an intensive care unit. J. Crit. Care 15, 147–150.

83. Walther, S.M., Jonasson, U., Karlsson, S., Nordlund, P., Johansson, A., Malstam, J., 2004. South-Eastern intensive care network of Sweden. Multicentre study of validity and interrater reliability of the modified nursing care recording system (NCR11) for assessment of workload in the ICU. Acta Anaesthesiol. Scand. 48, 690–696.
84. West, E., Barron, D.N., Harrison, D., Rafferty, A.M., Rowan, K., Sanderson, C., 2014. Nurse staffing, medical staffing and mortality in intensive care: an observational study. Int. J. Nurs. Stud. 51, 781–794.
85. Wysokinski, M., Ksykiewicz-Dorota, A., Fidecki, W., 2010. Demand for nursing care for patients in intensive care units in Southeast Poland. Am. J. Crit. Care 19, 149–155.
86. Wysokinski, M., Ksykiewicz-Dorota, A., Fidecki, W., 2013. Scope of nursing care in Polish intensive care units. Biomed. Res. Int. 2013, 463153.
87. Yamase, H., 2003. Development of a comprehensive scoring system to measure multifaceted nursing workloads in ICU. Nurs. Health Sci. 5, 299–308.
88. Yee Kwok, W.W., Chun Chau, J.P., Pau Le Low, L., Thompson, D.R., 2005. The reliability and validity of the therapeutic activity index. J. Crit. Care 20, 257–26

CHAPTER 3

VALIDATION OF THE NURSING ACTIVITIES SCORE (NAS) USING TIME-AND-MOTION MEASUREMENTS IN DUTCH INTENSIVE CARE UNITS

Charlotte C. Margadant, Marga E. Hoogendoorn, Rob J. Bosman, Jan Jaap Spijkstra, Sylvia Brinkman, Nicolette F. de Keizer

Under submission

Abstract

Background
The Nursing Activities Score (NAS) is widely used for workload measurement of Intensive Care Unit (ICU) nurses. However, the performance of the NAS to measure actual nursing time has not been comprehensively and externally validated. In order to validate the NAS with a more accurate measurement method compared to the work-sampling approach on which the NAS is developed, we performed time-and-motion techniques. The aim of this study is to validate the NAS using time-and-motion measurements in Dutch ICUs.

Methods
We measured nursing time for patients admitted to seven Dutch ICUs, between November 2016 and October 2017. The patient(s) that were under the care of a chosen nurse were followed by the observers during the entire shift and time-and-motion techniques were performed using an in-house developed web application. To validate the reliability of the NAS, we first converted NAS points per activity into minutes. Next, we compared the converted time per NAS item and the converted total nursing time per patient with the actual observed time. We used Wilcoxon signed-rank tests at nursing activity level and Pearon's R and R2 at patient level for these comparisons.

Results
A Pearson's correlation of R=0.59 (R2=0.35) was found between the total converted NAS time and the total observed time per patient. The median converted NAS time per patient (202.6 minutes) was higher compared to the observed time per patient (114.3 minutes). At NAS item level, we found significant differences between the converted NAS time and the observed time for all separate NAS items.

Conclusions
The development of the NAS was based on the work-sampling method which is not the most accurate method for this purpose. Therefore, we validated the NAS with time-and-motion technique which is more accurate, according to the literature. The NAS overestimates the needed nursing time for patients in Dutch ICUs. Therefore, we advise revisions of the time weights assigned to each NAS item to get better insight in the true nursing workload to enable the use of this information for more adequate nursing capacity planning.

Keywords: Nursing Activities Score (NAS), time-and-motion techniques, validation, nursing workload, patient acuity.

Contribution of the paper

What is already known about the topic?

- The Nursing Activities Score (NAS) is considered a valuable tool and is widely used for workload measurement in ICUs.
- The work sampling approach is not the most accurate method for measuring time.
- The performance of the NAS has not been comprehensively validated.

What this paper adds:

- Measurements with the most accurate time measurement: time-and-motion techniques.
- The NAS overestimates the needed nursing time for patients in the Dutch ICU setting.
- The low correlation of Pearson's R and R2 (0.59 and 0.35) implicates that the NAS is not accurate enough to estimate the nursing time at patient level.
- Revisions of the time weights assigned to each nursing activity of the NAS are needed to get better insight in the true nursing workload and to enable a more adequate nursing capacity planning.

Introduction

There are concerns regarding excessively high nursing workload, both in general- and ICU wards [1]. Excessive high nursing workload can lead to burnout and job dissatisfaction among nurses [2] and have a deleterious effect on patients [3]. Workload has risen due to increased turnover of patients, increased complexity of patient cases together with inadequate capacity of nurses due to shortages [4]. All this makes planning of nursing capacity important. In the last 30 years different instruments have been developed to measure the nursing workload to give insight in the needed nursing staff per shift and provide much needed input for capacity planning [5].

To assess nursing workload in the ICU Cullen et al. (1974) [6] created the Therapeutic Intervention Scoring System (TISS). The TISS was originally developed to classify nursing workload in relation to the severity of illness of ICU patients. The TISS exists of 76 therapeutic interventions that receive one to four points based on the severity of illness of the patient. It appeared that nursing workload is only partly related to the severity of illness, since less severely ill patients could also generate high nursing workload. For instance a patient recovering from a serious illness with an agitated delirium would not score high in severity of illness, but could demand very intensive nursing care, up to continuous bedside care throughout the day. This made the TISS less adequate to assess nursing workload. Therefore, the Nursing Activities Score (NAS) was developed in 2003 by Miranda et al. (2003) [7]. The NAS describes activities that largely represent the work actually performed by nurses at bedside in caring for the patients and has been developed to measure the nursing workload for each individual patient. The points assigned to the nursing activities provide an average time consumption in caring for the patients instead of representing the severity of illness. The NAS was created by using the work-sampling approach: at random moments per shift the nurse was asked what he or she was doing at that specific moment. For every activity there was a weight granted by the researchers. The total NAS for an individual patient is the sum NAS points of all activities, varying between 0 to 177 points. A score of 100 NAS points is equivalent to the amount of care which can be provided by one Full Time Equivalent nurse during either one shift or one day. A score above 100 points indicates that the needed care can only be provided by more than one nurse [7].

The NAS is considered a valuable tool and is widely used for workload measurement in ICUs [8,9]. However, the performance of the NAS has not been comprehensively validated. One study showed that the NAS might either under- or overestimate the actual nursing time required by patients and therefore recommended revision of the original NAS because of inadequate measurement of nursing activities [4].

Time-and-motion techniques are a widely used method for measuring the duration of nursing activities. In this method, continuous measurements are carried out by one observer to measure the occurrence and duration of all activities in a shift [10]. One-on-one observations are required in this approach and observers must follow nurses continuously for a large amount of time [11]. Since time-and-motion techniques are very labor- and cost intensive, researchers often choose to perform the work-sampling approach in which more, but less precise, measurements could be carried out in a short amount of time [11]. The authors of the NAS choose to work with the work-sampling method. However, research has demonstrated that this approach does not lead to an accurate representation of the true nursing workload. This is due to the fact that the weightings of the nursing activities are based on the probability that a particular nursing activity occurred [11]. The total amount of time in a shift is divided over the occurring nursing activities. When nursing activities frequently occur or take much time, they would also occur more frequent in the work-sampling approach. However, this approach will not lead to precise measurements, but will only approximate the time of the different activities. Thus, in contrast to time-and-motion techniques in which every minute of a nursing shift is measured, the work-sampling approach does not measure the real amount of time spent on the nursing activities which could lead to less accurate results [12]. Moreover, a major disadvantage of the self-reporting method in the work-sampling approach is that nurses might be able to be dishonest about the activities which they are carrying out at the specific moment of measurement. Nurses might be prone to show social desirable behavior when they need to report their work. In addition, nurses are burdened with extra registration and might compensate for this by overestimating their workload in the measurements [11]. Therefore, the time-and-motion technique is considered as the best technique for time measurement [13].

The aim of this study is to validate the NAS in the Dutch ICU setting using the time-and-motion technique, which give more precise approximations compared to the work-sampling approach on which the NAS was developed, and to identify which nursing activities are under- or overestimated in the NAS.

Methods

Setting

All 82 Dutch ICUs participate in the National Intensive Care Evaluation (NICE) quality registry. Fifteen of these ICUs are participating in the newly implemented voluntary nursing capacity module 14. Seven of these ICUs voluntarily participated in this study. Data on characteristics of the ICUs (such as number of ICU beds) and data on patient characteristics (such as age, BMI, admission type, and mortality) were extracted from the NICE registry.

Time-and-motion

The study involved time-and-motion measurements for patients admitted to the ICU. We measured in different types of hospitals (academic-, teaching-, and non-teaching hospitals) and in different shifts (day, evening, and night). At the start of a shift one nurse was chosen by the observer. The patient(s) that were under the responsibility of this nurse were followed by the observer during the entire shift. A patient admitted for a longer time could theoretically be observed on different dates during different shifts and therefore could possibly be followed during more than one measured shift. The measurements took place during different days of ICU admission of the patients (e.g. first ICU admission day through last ICU admission day) and with different type of nurses (registered and student nurses).

Observers were researchers CM and MH and ten independent student nurses. The students were trained in performing time-and-motion measurements by oral and written instructions and one day of measuring together with one of the researchers. The observers used an in-house developed web application to record start and stop times of each performed nursing activity (Appendix B). The application contained, among other, all activities occurring in the NAS [See Additional file 2]. If two nurses were simultaneously performing nursing activities for the same patient this was also registered, by pressing the 'two nurses button' and multiplying this time by two in the analysis. In case of two different activities carried out by two nurses, these activities could be measured simultaneously.

Measurements were conducted between November 1st 2016 and October 1st 2017. Participation of the hospitals was on a voluntary basis. Seven hospitals were willing to participate. Data was processed in an anonymous way.

Ethical approval

The Institutional Research Board of the Amsterdam University Medical Centre reviewed the research proposal and waived the need for informed consent (IRB protocol W17_366).

Data analysis

Nursing activities that occurred less than ten times in the total dataset were excluded from the analysis. All NAS items have a fixed number of NAS points but some items have different categories corresponding to different numbers of NAS points depending on the duration of that activity (e.g. bedside with hourly vital signs, bedside for two hours or more, or four hours or more). For these duration depended activities, we first used the measured time for that activity during the measurements to assign the correct number of points. For example, a nurse performed hygiene procedures on a patient for 1.2 hours during a shift according to our time measurements. This NAS item has three categories: performing hygiene procedures for less than two hours, for more than two hours, or for more than four hours. In above-mentioned example, the activity took 1.2 hours and would therefore be assigned to the category for less than two hours, which corresponds to 4.1 NAS points. Since we had the information of the occurrence and duration of all nursing activities per shift, we were able to convert all time-and-motion data into NAS scores. These obtained NAS scores give an accurate view of the NAS in each shift. To obtain these NAS scores, we first converted the originally assigned NAS points per activity into time. Based on Miranda et al. (2003) 6 100 NAS points correspond to 100% of care time provided by one nurse during a shift and hence 1 NAS point corresponds to 1% of care time provided by one nurse. Given the fact that a nurse is productive in 80% of the 8-hour shift, which has been described by the authors of the NAS, one NAS point corresponds to 3.84 minutes of nursing care during an 8-hour shift (((8 hours * 60mins)/100)*0.8) [7,15]. With this information we were able to convert the NAS scores into an estimated nursing time per patient and per nursing activity (from now on referred to as converted NAS time, see Appendix C). Next, we compared the time per NAS item and the total nursing time per patient, based on NAS scores according to the model, with observed times from the time-and-motion measurements. For the observed time, we took the sum of the times of all performed nursing activities per patient per shift in minutes (from now on referred to as observed time, see Additional file 3).

The median and interquartile ranges (IQR) of the converted NAS times and the observed times were calculated. First, the difference between the total converted NAS times and the total observed times per patient were visualized by scatterplots. Second, the correlation between the total converted NAS times and the total observed times per patient were assessed with the Pearson's correlation test. In addition, we also assessed the R^2, a measure for the proportion of the variance. For each nursing activity separately, medians and interquartile ranges (IQR) of the converted NAS times and observed times were calculated and differences were tested with the Wilcoxon signed-rank test. All statistical analyses were performed using R statistical software, version 3.3.2 16.

Results

Baseline results

Table 1 shows the ICU characteristics of the seven included ICUs compared to all Dutch ICUs. No significant differences were found in ICU characteristics between the included ICUs and all Dutch ICUs (Table 1). During our study, a total of 287 unique patients have been observed during 371 different shifts with time-and-motion measurements. In these patients, 46.319 nursing activities have been measured. In 60% of the measurements, nurses took care for two or three patients per shift. For the remaining 40%, nurses cared for one patient per shift. The patients in our study had a significant higher in-hospital mortality rate (22.3% versus 13.0%) and length of ICU stay (3.2 days versus 1.0 day) compared to all Dutch patients in the same period (Table 2). Furthermore, acute renal failure, chronic respiratory insufficiency, and cirrhosis differed between the groups, with a higher percentage in the patients in our study. For the other patient characteristics, the included patients and all Dutch ICU patients in this period were comparable.

Table 1. ICU characteristics

Variable	**Included ICUs (N=7)**	**All Dutch ICUs (N=84)**
Number of university hospitals (%)	1 (14%)	9 (11%)
Number of teaching hospitals (%)	4 (57%)	23 (27%)
Number of non-teaching hospitals (%)	2 (29%)	52 (62%)
Median number of ICU beds per ICU (IQR)	13.0 [9.0, 17.0]	12.0 [8.0, 16.0]

Table 2. Patient Characteristics

Variable	Included patients in measurements	All Dutch ICU patients
Number of unique patients, N	287	100.145
Age, median [IQR]	66.0 [56.0 – 76.0]	66.0 [55.0 – 75.0]
BMI, median [IQR]	26.0 [23.6 – 28.7]	25.9 [23.1 – 28.4]
Admission type		
-Medical, N (%)	121 (42.2)	51290 (52.7)
-Surgical: urgent and elective, N (%)	151 (52.6)	45905 (47.2)
In-hospital mortality, N (%)*	85 (22.3)	13017 (13.0)
ICU LOS (in days), median [IQR]*	3.2 [0.9, 14.8]	1 [0.7 – 4.0]
Comorbidities		
Acute renal failure, N (%)*	37 (12.9)	9211 (9.2)
Cardiovascular insufficiency, N (%)	16 (4.2)	4257 (4.3)
Chronic renal failure[1], N (%)	25 (6.7)	7976 (7.9)

Chronic respiratory insufficiency, N (%)*	7 (2.4)	4620 (4.6)
Cirrhosis, N (%)*	1 (3.5)	1751 (1.7)
COPD, N (%)	36 (12.5)	13304 (13.3)
Diabetes, N (%)	68 (17.8)	16273 (16.2)
Gastrointestinal bleeding, N (%)	2 (0.7)	2263 (2.3)
Hematologic malignancy, N (%)	6 (2.1)	2143 (2.1)
Immunological insufficiency, N (%)	16 (5.6)	8290 (8.3)

* Indicates a significant P-value of <0.05

COPD: Chronic Obstructive Pulmonary Disease; IQR: Interquartile Range

1 Chronic renal failure consists of chronic renal insufficiency and chronic dialysis

NAS validation

Excluded nursing activities

The following three NAS nursing activities occurred less than 10 times in all measurements and were therefore excluded from the analysis at activity level: activities at the pulmonary- or left atrium catheter, cardiopulmonary resuscitation (CPR), and specific interventions in the ICU (endotracheal intubation, insertion of pacemaker, cardioversion, endoscopies, emergency surgery in the previous 24 hours, gastric lavage). Furthermore, we did not measure specifically intravenous replacement of large fluid losses and treatment of metabolic acidosis/alkalosis, since these two nursing activities are usually administered under bedside activities.

Total patient time and times per NAS item

The median converted NAS time per patient (202.6 minutes; IQR 155.0 – 241.2 minutes) was significantly higher ($p<0.001$) compared to the observed time per patient (144.3 minutes; IQR 81.3 – 168.4 minutes), see Figure 1. A Pearson's correlation of $R=0.59$ ($R^2=0.35$) was found between the total converted NAS time and the total observed time per patient (Table 3).

For the time differences at NAS item level, we found significant differences between the converted NAS times and observed times for all items. These differences ranged from -54.6 minutes (support or care for patient or relatives for about 1 hour) to 79.2 minutes (mobilization and positioning with three nurses). For most (86%) nursing activities the median converted NAS time overestimated the observed time. For four activities (support or care for patient for about 1 hour, administrative tasks for less than 2 hours, administrative tasks for about 2 hours and specific interventions outside the ICU) the converted NAS time underestimated the observed time (Table 3).

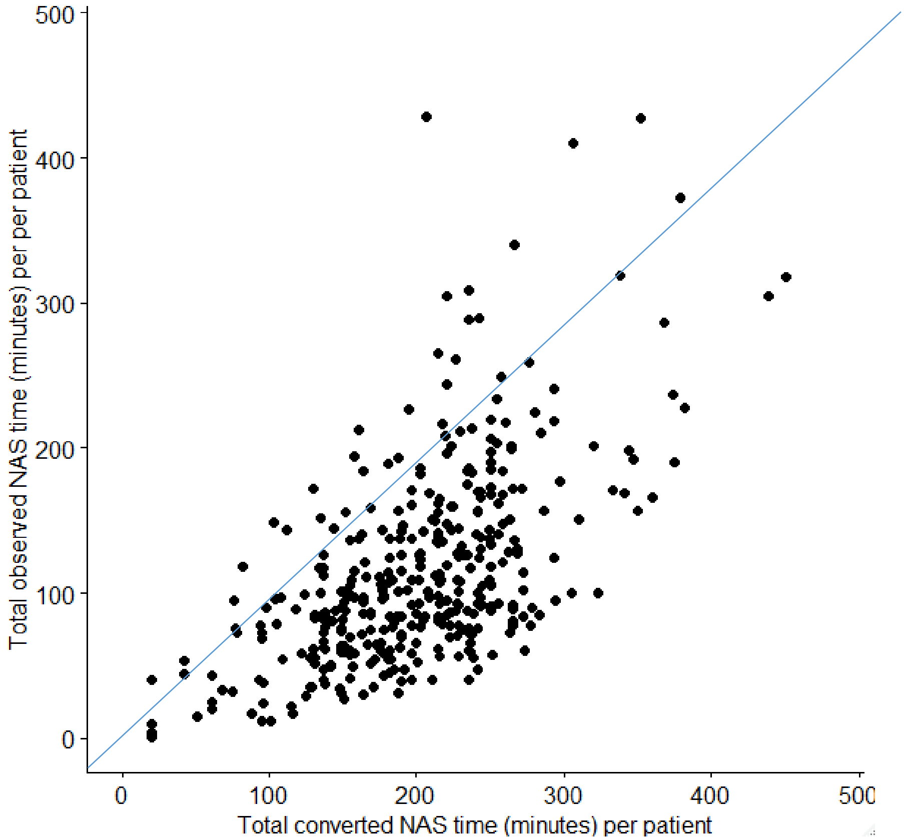


Figure 1. The correlation between the total converted NAS time in hours and the total observed time in minutes per patient. A full nursing shift is 480 minutes. Blue diagonal shows equal converted and observed time per patient.

Table 3. NAS activities with their points according to Miranda et al. (2003)[6], and the median converted NAS times and observed times per NAS item. N = 371 patients and 46.319 measured nursing activities.

NAS item	NAS points per activity	Median converted NAS time (minutes)	Median observed time (minutes) [IQR]	Difference in minutes, median [IQR]
1a Present at bedside and continuous observation or active for <2 hours [1]	4.5	21.6	14.22 [7.26-26.17]	7.38 [-4.57-14.35]*
1b Present at bedside and continuous observation or active for ≥2 hours	12.1	NA	NA	NA
1c Present at bedside and continuous observation or active for ≥4 hours	19.6	NA	NA	NA
2 Laboratory, biochemical and microbiological investigations	4.3	20.64	5.45 [3.13-8.81]	15.19 [11.83-17.51]*
3 Medication, vasoactive drugs excluded	5.6	26.88	2.24 [0.90-4.91]	24.64 [21.97-25.98]*
4a Performing hygiene procedures ≤2 hours	4.1	19.68	11.58 [3.95-27.8]	8.1 [-8.12-15.73]*
4b Performing hygiene procedures >2 hours	16.5	NA	NA	NA
4c Performing hygiene procedures >4 hours	20.0	NA	NA	NA
5 Care of drains	1.8	8.64	2.41 [0.92-4.64]	6.23 [4.0-7.72]*
6a Mobilization and positioning, performing procedure(s) up to 3 times per 24 hours	5.5	26.4	2.46 [0.91-4.88]	23.94 [21.52-25.49]*

6b Mobilization and positioning, performing procedure(s) >3 times per 24 hours, or with two nurses	12.4	59.52	4.82 [2.17-9.33]	54.69 [50.19-59.49]*
6c Mobilization and positioning, performing procedure(s) with 3 nurses	17.0	81.6	2.4 [0.89-6.16]	79.2 [75.44-80.71]*
7a Support or care for patient or relatives for about 1 hour	4.0	19.2	2.4 [0.89-6.16]	-54.58 [-65.18- -49.26]*
7b Support or care for patient or relatives for about 3 hours	32.0	NA	NA	NA
8a Administrative or managerial tasks for <2 hours	4.2	20.16	40.91 [28.53-60.33]	-20.74 [-40.17- -8.37]*
8b Administrative or managerial tasks for about 2 hours	23.2	111.4	130.0 [126.3-157.4]	-18.67 [-46.02- -14.92]*
8c Administrative or managerial tasks for about 4 hours	30.0	NA	NA	NA
9 Respiratory support	1.4	6.72	2.99 [1.42-5.9]	3.73 [0.82-5.30]*
10 Care of artifical airways	1.8	8.64	1.43 [0.5-4.77]	7.21 [3.87-8.14]*
11 Treatment for improving lung function	4.4	21.12	1.32 [0.64-2.79]	19.80 [18.33-20.48]*
12 Vasoactive medication	1.2	5.76	1.99 [0.95-4.99]	3.78 [-0.77-4.81]*
13 Intravenous replacement of large fluid losses	2.5	NA	NA	NA
14 Left atrium monitoring	1.7	NA	NA	NA

15 Cardiopulmonary resussctitation after arrest	7.1	NA	NA	NA
16 Hemofiltration techniques	7.7	36.96	18.76 [7.83-36.66]	18.20 [-1.67-28.78]*
17 Qualitative urine output measurement	7.0	33.6	1.35 [0.66-2.45]	32.25 [31.15-32.96]*
18 Measurement of intracranial pressure	1.6	7.68	0.91 [0.28-2.62]	6.77 [5.07-7.4]*
19 Treatment of complicated metabolic acidosis	1.3	NA	NA	NA
20 Intravenous hyperalimentation	2.8	13.44	2.64 [0.79-4.1]	10.80 [9.41-12.65]*
21 Enteral feeding trhourgh gastric tube	1.3	6.24	1.87 [0.81-4.64]	4.37 [1.6-5.43]*
22 Specific interventions in the ICU	2.8	NA	NA	NA
23 Specific interventions outside the ICU	1.9	9.12	18.18 [5.69-27.46]	-9.06 [-18.34-3.43]*
Total per patient	-	202.56 [155.04-241.2]	98.52 [71.86-127.72]	84.7 [50.31-127.72]*

* Indicates a significant P-value of <0.05 (Wilcoxon signed-rank test)
[1] Titles are abbreviated. For full activity names see Appendix D.
NA: not measured during measurement

Discussion

Our analysis showed that the NAS overestimates the needed nursing time for patients in the Dutch ICU setting. Times of most NAS items were overestimated by the NAS, except for four activities (support or care for patient for about 1 hours, administrative tasks for less than 2 hours, administrative tasks for about 2 hours, and specific interventions outside the ICU), where the NAS gives an underestimation of the observed time. This study showed that 35% of nursing time is explained by the NAS model (R2 = 0.35). The converted NAS time per patient (202.6 minutes per shift) in our study was comparable with the converted NAS times per patient in other studies. Bernet et al. (2005)[17] found 150 to 156 minutes per shift and Deberg et al. (2007)[18] found 180 to 228 minutes per shift. The different articles on the NAS give variable NAS times per shift. A full shift of work equals 480 minutes of nursing time.

The low correlation of Pearson's R and R^2 (0.59 and 0.35) implicates that the NAS is not accurate enough to estimate the nursing time at patient level. However, it is currently still the best nursing workload model for quantifying nursing workload in ICUs [5]. There is no clear cut-off point from which the model can be identified as 'good enough' based on the R2. However, since the NAS is used for capacity planning, a R2 closer to 1 would be more desirable.

Since in almost each shift ICU nurses also spend time on non-nursing duties, e.g. coaching a student or participating in an emergency team within the hospital, we performed a sensitivity analysis to determine whether these non-nursing duties were affecting the correlation. According to several studies nurses spend approximately 3 to 6% of their shift on non-nursing duties [19,20,21,22,23]. We therefore took the average of 4.5% and subtracted this from the 80% of productive nursing time, which we used in this study to calculate the converted time per NAS point. Using this approximation, the converted time would have changed from 3.84 to 3.62 minutes per NAS point. This change does not affect the results and we therefore conclude that non-nursing duties are not significantly influencing the performance of the NAS.

A strength of our study is that we validated the NAS with time-and-motion measurements which are considered to be the best technique for measuring nursing workload [13]. To our knowledge this has not been performed before in the context of NAS validation. Measurements for nursing activities by using time-and-motion measurements, are more accurate compared to the work-sampling approach as used for the development of the NAS [24]. Moreover, we measured the NAS precisely as was indicated by the authors, which makes the validation reliable. Furthermore, since measurements took place in all types of ICUs, we believe that results of this study are generalizable to all Dutch ICUs.

One of the limitations in our study are the excluded NAS activities due to their non- or limited occurrence of less than ten times. Two of these activities are mostly scored in other categories of activities: the activity 'intravenous replacement of large fluid losses' is mostly scored under NAS item 1 'bedside'. The activity 'treatment of complicated metabolic acidosis/alkalosis' is mostly scored in NAS item 3 'medication'. Since these activities could be scored in other categories, these activities can be excluded from the NAS. Three NAS activities (respectively left atrium monitoring, cardiopulmonary resuscitation after arrest, and specific interventions in the ICU) and six subcategories 1b, 1c, 4b, 4c, 7b, and 8c (the nurse activities that required dedication from the nurse for more than 2, 3 or 4 hours) did not happen often enough (>=10 times) during the measurements which makes the validation of the NAS incomplete. Given the fact that the median time of nursing care per patient is 2.4 hours (144.3 minutes), dedication of a nurse for more than 2, 3 or 4 hours to one activity is extremely high. As these nursing activities rarely occur in daily ICU practice it is not likely that our results have been affected by this situation.

Furthermore, the observed patients seem to have been more severely ill and consequently had a longer length of stay compared to all Dutch patients in the same time period, which is likely caused by our selection mechanism. In order to measure as many nursing activities as possible we probably choose more often nurses who took care of patients that were expected to stay the whole shift and these patients were probably more severely ill. This may have biased our results since our aim was to validate the NAS and check for under- or overestimations compared to time-and-motion measurements and it is possible that observed times in sicker patients differ from those in less sick patients. However, according to Armstrong et al. (2015) NAS scores in intermediate care patients did not differ from those in ICU patients [25].

At last, a limitation of the time-and-motion techniques are the change in behavior when nurses are aware that their occurrence and duration of activities are measured by an observer [26]. This bias is able to minimize over time but we measured the activities of many different nurses which makes our study prone to this type of bias. However, taking this bias into account, the time-and-motion techniques are still a more accurate measure for the measuring nursing time compared to the work-sampling approach.

Based on our results we believe there is room for improvement in the measurement of nursing workload. The NAS could be improved by adjusting the NAS points given to the different items. The developers of the NAS did not report the Pearson's R or R^2, but stated that the NAS is reflecting 81% of total nursing time. About 11% of the nurses' time is spent on personal activities. The remaining 8% comes from nursing activities derived from medical interventions, related exclusively to the severity of illness of the patient not measured by the NAS [7]. The TISS is taking these medical interventions into account, such as induced hypothermia, cardiac assist device, pacemaker, or ECG monitoring. For this reason, we suggest additional research towards the merging of the TISS-28 and the NAS. The models could be partly combined which could possibly improve the estimation of nursing workload. Our results on observed time per patient and per nursing activity could be taken into consideration when assigning weights to the activities in this new model. Moreover, we think that expressing nursing activities in minutes or hours would be more informative compared to points, since it is more straight forward for ICU managers to work with.

Conclusion

The development of the NAS was based on the work-sampling method which is not the most accurate method for this purpose. Therefore, we validated the NAS with time-and-motion techniques which are a more accurate measure, according to the literature. The NAS has been developed more than 15 years ago and significantly overestimates the needed nursing time for ICU patients in the current daily ICU practice. Therefore we recommend a revision of the time weights assigned to each nursing activity to get better insight in the true nursing workload and to enable a more adequate nursing capacity planning.

References

1. Hughes, R.: Patient safety and quality: An evidence-based handbook for nurses (Vol. 3). Rockville, MD: Agency for Healthcare Research and Quality 2008.
2. Aiken, L. H., Clarke, S. P., Sloane, D. M., et al: Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction. Jama 2002; 288(16), 1987-1993.
3. Margadant, C., Wortel, S., Hoogendoorn, et al.: The Nursing Activities Score per Nurse Ratio Is Associated With In-Hospital Mortality, Whereas the Patients Per Nurse Ratio Is Not. Critical Care Medicine 2019; 48(1), 3-9.
4. Palese, A., Comisso, I., Burra, M., et al.: Nursing Activity Score for estimating nursing care need in intensive care units: findings from a face and content validity study. Journal of nursing management, 2006; 24(4), 549-559.
5. Hoogendoorn, M. E., Margadant, C. C., Brinkman, S.: Workload scoring systems in the Intensive Care and their ability to quantify the need for nursing time: a systematic literature review. International Journal of Nursing Studies 2019; 103408.
6. Cullen, D. J., Civetta, J. M., Briggs, B. A., & Ferrara, L. C. (1974). Therapeutic intervention scoring system: a method for quantitative comparison of patient care. Critical care medicine 1974; 2(2), 57-60.
7. Miranda, D. R., Nap, R., de Rijk, A., et al.: Nursing activities score. Critical care medicine 2003; 31(2), 374-382.
8. Padilha, K. G., de Sousa, R. M. C., Garcia, P. C., et al.: Nursing workload and staff allocation in an intensive care unit: a pilot study according to Nursing Activities Score (NAS). Intensive and Critical Care Nursing 2010; 26(2), 108-113.
9. Gonçalves, L. A., Padilha, K. G., & Sousa, R. M. C. Nursing activities score (NAS): a proposal for practical application in intensive care units. Intensive and Critical Care Nursing 2007; 23(6), 355-361.
10. Roddy, P. C. (1987). Resource requirements of nursing home patients based on time and motion studies (No. 4). US Department of Health and Human Services, Public Health Service, National Center for Health Services Research and Health Care Technology Assessment.
11. Finkler, S. A., Knickman, J. R., Hendrickson, G. et al.: A comparison of work-sampling and time-and-motion techniques for studies in health services research. Health services research 1993;28(5), 577.
12. Urden, L. D., & Roode, J. L.: Work sampling: a decision-making tool for determining resources and work redesign. Journal of Nursing Administration 1997; 27(9), 34-41.
13. Burke, T. A., McKee, J. R., Wilson, et al.: A comparison of time-and-motion and self-reporting methods of work measurement. JONA: The Journal of Nursing Administration 2000; 30(3), 118-125.
14. Dutch National Intensive Care Evaluation (NICE) foundation. Available at: https://www.stichting-nice.nl/dd/#561 Accessed June 1, 2019.
15. Interview Reis Miranda, jaarboek NICE. Available at: https://www.stichting-nice.nl/focusIC.jsp Accessed September 1, 2019.

16. R Core Team. R: A language and Environment for Statistical Computing [computer program]. Vienna, Austria: R foundation. Available at: http://www.R-project.org Accessed May 1, 2019.
17. Bernat Adell A, Abizanda Campos R.: Nursing Activity Score (NAS). Our experience with a nursing load calculation system based on times. Enferm Intensiva 2005 Oct-Dec;16(4):164-73.
18. Debergh, D. P., Myny, D., Van Herzeele, I., et al.: Measuring the nursing workload per shift in the ICU. Intensive care medicine 2012; 38(9), 1438-1444.
19. Abbey Mphil, M., Chaboyer, W., & Mitchell, M.: Understanding the work of intensive care nurses: a time and motion study. Australian Critical Care 2012; 25(1), 13-22.
20. Bosman, R., Rood, E., Oudemans-van Straaten, H., Van der Spoel, J., et al.: Intensive care information system reduces documentation time of the nurses after cardiothoracic surgery. Intensive care medicine 2003; 29(1), 83-90.
21. Harrison, L., Nixon, G.: Nursing activity in general intensive care. Journal of Clinical Nursing 2002; 11(2), 158-167.
22. Kaya, H., Kaya, N., Turan, Y. et al.: Nursing activities in intensive care units in Turkey. International journal of nursing practice 2011; 17(3), 304-314.
23. Norrie, P.: Nurses' time management in intensive care. Nursing in critical care 1997; 2(3), 121-125.
24. Steyerberg, E.W.: Evaluation of Performance. Clinical Prediction models: a practical approach to development, validation, and updating. Springer 2009; p. 83-99.
25. Armstrong, E., de Waard, M. C., de Grooth, H. J. S., et al.: Using nursing activities score to assess nursing workload on a medium care unit. Anesthesia & Analgesia 2015; 121(5), 1274-1280.
26. Barker, K. N. (1980). Data collection techniques: observation. American Journal of Hospital Pharmacy, 37(9), 1235-1245.

Chapter 6 – Effect of patient characteristics and contextual factors on needed nursing time in intensive care units

CHAPTER 4

THE NURSING ACTIVITIES SCORE PER NURSE RATIO IS ASSOCIATED WITH IN-HOSPITAL MORTALITY, WHEREAS THE PATIENTS PER NURSE RATIO IS NOT

Charlotte C. Margadant, Safira A. Wortel, Marga E. Hoogendoorn, Rob J. Bosman, Jan Jaap Spijkstra, Sylvia Brinkman, Nicolette F. de Keizer

Critical Care Medicine 2020 48(1), 3-9. DOI: 10.1097/CCM.0000000000004005

Abstract

Objectives

Studies have shown contradicting results on the association of nursing workload and mortality. Most of these studies expressed workload as patients per nurse ratios; however, this does not take into account that some patients require more nursing time than others. Nursing time can be quantified by tools like the Nursing Activities Score. We investigated the association of the Nursing Activities Score per nurse ratio, respectively, the patients per nurse ratio with in-hospital mortality in ICUs.

Patients

All ICU patients admitted to and registered ICU nurses working at 15 Dutch ICUs between January 1, 2016, and January 1, 2018, were included. The association of mean or day 1 patients per nurse ratio and Nursing Activities Score per nurse ratio with in-hospital mortality was analyzed using logistic regression models.

Measurements and Main Results

Nursing Activities Score per nurse ratio greater than 41 for both mean Nursing Activities Score per nurse ratio as well as Nursing Activities Score per nurse ratio on day 1 were associated with a higher in-hospital mortality (odds ratios, 1.19 and 1.17, respectively). After case-mix adjustment the association between a Nursing Activities Score per nurse ratio greater than 61 for both mean Nursing Activities Score per nurse ratio as well as Nursing Activities Score per nurse ratio on day 1 and in-hospital mortality remained significant (odds ratios, 1.29 and 1.26, respectively). Patients per nurse ratio was not associated with in-hospital mortality.

Conclusions

A higher Nursing Activities Score per nurse ratio was associated with higher in-hospital mortality. In contrast, no association was found between patients per nurse ratios and in-hospital mortality in the Netherlands. Therefore, we conclude that it is more important to focus on the nursing workload that the patients generate rather than on the number of patients the nurse has to take care of in the ICU.

Introduction

Critically ill patients in ICUs require intensive nursing care and monitoring [1]. This is one of the main reasons why ICU care is expensive. In Europe, the costs of ICUs represent approximately 20% of total hospital costs [2]. On average, 50% of the budget of the ICU is spend on nursing staff [3]. Hence, taking a critical evaluation on how many nurses are actually needed is of great importance from an economical point of view. On the other hand, reducing the number of nurses working at the ICU due to budgetary cuts will increase nursing workload which may have a negative impact on patient safety and thereby patients' mortality risk. Additionally, reduction of the number of nurses can negatively impact nurses well-being [4]. Therefore, quantification of an optimal patient to nurse ratio (PNR) has been a topic of debate for years [5].

In a meta-analysis conducted by Numata et al. [6] a positive relationship between higher nurse staffing levels and lower in-hospital mortality among critical care patients was found, but after additional case-mix adjustment, the reported association became nonsignificant in four out of five included studies. A recent meta-analysis in ICUs found that higher PNRs are associated with a 14% increase in the in-hospital mortality risk [7]. In these studies, the fact that patients differ in their need for nursing care is neglected. This makes it difficult to compare studies in which nurses take care of different type of patients, thereby hampering meta-analyses. A recent editorial suggests that nurse staffing levels should reflect a combination of patient nursing need (acuity and dependency level), patient throughput, nursing competency, and availability of ancillary staff [8].

The Nursing Activities Score (NAS) is a well-known and frequently used scoring system to quantify the need of nursing care in ICUs [9]. This scoring system assigns a score per patient based on the executed nursing tasks, where each of them is weighted for the time spent per task. The scores are translated to full-time equivalent (FTE), where 100 points equals 1 nursing FTE. The score per patient can run from 0 to 170 points. So one nurse can take care of several patients, whereas some patients need to be taken care of by more than one nurse. To date, the NAS is the most commonly used instrument to measure nursing workload in ICUs. The NAS was validated with Multi Moment Recordings, where they showed that 81% of the total time spent by nurses was explained by the NAS [9]. Furthermore, Stafseth et al. [10] performed a study on the validity and reliability of the NAS and provided empirical support for its usefulness in the assessment of critical care nursing. The NAS per nurse ratio (NNR) represents the amount of workload per nurse executed per shift and may be a more reliable measure of workload than the PNR. Although the PNR does not consider the amount of workload that a patient requires, the NNR provides more elaborate information on the workload, since it takes into account the amount of time a nurse spends on a patient and is therefore more likely to be an efficient measure for capturing the workload of a nurse. However, to our knowledge, no studies have been published using the NNR to assess the association of workload with in-hospital mortality.

Therefore, the aim of this study was to investigate the association of nursing workload, expressed as NNRs and in-hospital mortality and to compare this with the association of PNR and in-hospital mortality in ICUs.

Materials and methods

Setting

We used data from the Dutch National Intensive Care Evaluation (NICE) quality registry in which all Dutch ICUs participate since 2016 [11]. The NICE minimal dataset consists of demographic, physiologic and diagnostic data, and in-hospital mortality of all admitted ICU patients in all Dutch hospitals. One of the optional modules in the NICE registry is the nursing workload module in which ICUs register per shift, among others, the present number of FTE nurses and all data items needed to calculate the NAS per patient per shift [12].

Data Collection

For this study, we performed a retrospective analysis of the NICE database. We included all ICU admissions between January 1, 2016, the start of the NAS data collection within the NICE registry, and January 1, 2018, from all 15 Dutch ICUs (out of 82 ICUs in the Netherlands) that voluntarily participated in the nursing workload module. We excluded patients that were readmitted during the same hospitalization period and patients that were admitted for less than 1 day. Almost all needed data for the nursing workload module (around 90%) can be automatically extracted from the electronic health record (EHR), but in some hospitals, the nurses have to manually register the data at the end of their shift into a specific form in the EHR. When nurses need to manually register the data, they are trained for this by following an e-learning on the nursing workload module of the NICE registry.

Ethical Approval

The Institutional Research Board of the Amsterdam University Medical Center reviewed the research proposal and waived the need for informed consent (Institutional Review Board protocol W17_366).

Calculation of NNRs and PNRs

NAS was developed to use for a 24 hours period but can also be used per shift as was done in the included ICUs. For the calculation of the daily NNRs, the number of assigned NAS points for each nursing intervention per shift has been converted into points at daily level per nursing intervention according to the original NAS model. For example, with the nursing intervention "drains," the maximum scored value over three shifts was taken, for the nursing intervention "bedside" the hours of the three shifts were summed. Finally, all these daily NAS points for the different nursing interventions were summed to reach the NAS score per day. Subsequently, the sum of the daily NAS score of all patients during a day were divided by the average number of registered ICU nurses present (excluding student or other types of nurses) during that day to assess the daily NNRs.

For the calculation of the daily PNRs, we converted the data collected per shift into a daily score. First, the PNRs of the day, evening, and night shifts were calculated by dividing the number of ICU patients (recovery unit and coronary care unit patients were excluded) present during the shift by the number of registered ICU nurses (student- and non-registered ICU nurses were excluded) present. Subsequently, the average of the PNRs of the day, evening, and night shift per day were used as daily PNRs. For the definitions of the NNR and PNR, see Appendix E.

As a sensitivity analysis, we also performed the NNRs and PNRs analysis including the recovery unit and coronary care unit patients admitted to the ICU and including the student nonregistered ICU nurses to check whether these changes in definition influenced the results.

Statistical Analysis

Categorical variables were presented using absolute and relative frequencies, and the continuous variables were presented using mean with SD's or median with interquartile ranges depending on their distribution. Patients with missing NAS values were excluded from the analysis.

Results

Baseline Results

The mean number of ICU beds of the included ICUs was slightly lower than ICUs that did not participate in the nursing workload module, although not significant. Table 2 shows that also the other characteristics of the ICUs are similar. In total, 34,524 patients were admitted for the first time during their hospitalization period to an ICU that participated with the NICE nurse workload module (Table 3). For 29,445 patients (85%) workload data was collected and could be included in our study. Because of missing NAS values, 15% of the patients were excluded from the analysis. The demographic characteristics of the included and excluded patients show some small differences of which we presume that they have no influence on our found results, as the in-hospital mortality rate among the excluded patients was similar to that of the included patients (Table 3). Of the 29,445 included patients, 11.8 % died during the hospital admission period. The median NNR on day 1 and the mean NNR during the ICU admission period were also almost similar, respectively, 61.1 and 61.2. The median PNR on day 1 was 1.38 and the PNR during the whole ICU admission period was 1.39.

Table 1. Overview of Logistic Regression Models

Model no.	Model Format
1	In-hospital mortality ~ (1|hospital of ICU admission) + PNR day 1
2	In-hospital mortality ~ (1|hospital of ICU admission) + PNR day 1 + comorbidities [a] + age + admission type [b]
3	In-hospital mortality ~ (1|hospital of ICU admission) + mean PNR during admission
4	In-hospital mortality ~ (1|hospital of ICU admission) + mean PNR during admission + comorbidities [a] + age + admission type [b]
5	In-hospital mortality ~ (1|hospital of ICU admission) + NNR day 1
6	In-hospital mortality ~ (1|hospital of ICU admission) + NNR day 1 + comorbidities [a] + age + admission type [b]
7	In-hospital mortality ~ (1|hospital of ICU admission) + mean NNR during admission
8	In-hospital mortality ~ (1|hospital of ICU admission) + mean NNR during admission + comorbidities [a] + age + admission type [b]

NNR = Nursing Activities Score per nurse ratio, PNR = patients per nurse ratio.
[a] *Comorbidities include immunological insufficiency, neoplasm, hematologic malignancy, chronic cardiovascular insufficiency, chronic dialysis, cirrhosis, chronic renal insufficiency, chronic obstructive pulmonary disease, and chronic respiratory insufficiency.*
[b] *Values for admission type: medical, urgent surgery, or elective surgery.*

Table 2. ICU Characteristics

Characteristics	Included ICUs (n=15)	All Dutch ICUs (n=82)
Number of university hospitals (%)	2 (13%)	8 (10%)
Number of teaching hospitals (%)	3 (20%)	22 (27%)
Number of non-teaching hospitals (%)	10 (67%)	52 (63%)
Median number of ICU beds per ICU (IQR)	9.0 [7.0, 17.0]	12.0 [8.0, 16.0]
Median number of registered ICU nurses working at the ICU (IQR)	38.0 [21.5, 62.8]	38.3 [21.8, 57.9]
Median number of ICU student nurses working at the ICU (IQR)	3.9 [2.3, 6.0]	4.0 [2.3, 7.1]

Table 3. Patient Characteristics

Characteristics	**Patients with workload information**	**Patients without workload information**	**P-value**
Number of patients, N	29,445	5,079	-
Age, median [IQR]	66.0 [55.0, 74.0]	65.0 [53.0, 74.0]	<0.001
Admission type			0.104
-Medical N (%)	11,977 (41.0)	2,616 (51.6)	-
-Surgical: urgent, N (%)	3,652 (12.5)	578 (11.4)	-
-Surgical: elective, N (%)	13,541 (46.4)	1,876 (37.0)	-
In-hospital mortality, N (%)	3,462 (11.8)	589 (11.8)	0.897
ICU LOS (in days), median [IQR]	1.00 [0.80, 2.68]	1.11 [0.76, 2.02]	<0.001
LOS prior to ICU admission, median [IQR]	1.56 [0.79, 2.61]	1.12 [0.63, 1.71]	<0.001
Comorbidities			
Acute renal failure, N (%)	2,085 (7.1)	354 (7.0)	0.603
Cardiovascular insufficiency, N (%)	1,350 (4.6)	68 (1.3)	<0.001
Chronic dialysis, N (%)	298 (1.0)	113 (2.2)	<0.001
Chronic renal insufficiency, N (%)	1,700 (5.8)	405 (8.0)	<0.001
Chronic respiratory insufficiency, N (%)	1,120 (3.8)	242 (4.8)	0.006
Cirrhosis, N (%)	295 (1.0)	92 (1.8)	<0.001
COPD, N (%)	3,493 (11.9)	624 (12.3)	0.815
Diabetes, N (%)	4,975 (16.9)	896 (17.6)	0.288
Gastrointestinal bleeding, N (%)	434 (1.5)	85 (1.7)	0.582
Hematologic malignancy, N (%)	489 (1.7)	128 (2.5)	0.001
Immunological insufficiency, N (%)	2,868 (9.8)	239 (4.7)	<0.001
Neoplasm, N (%)	1,902 (6.5)	157 (3.1)	<0.001
NNR and PNR[a]			
NNR on day 1 Median [IQR]	61.1 [41.2, 77.2]	-	-
NNR mean[b], Median [IQR]	61.2 [41.2, 76.3]	-	-
PNR on day 1, Median [IQR]	1.38 [0.96, 1.93]	-	-
PNR mean[b], Median [IQR]	1.39 [0.97, 1.91]	-	-

IQR = interquartile range, LOS = length of stay, NNR = Nursing Activities Score per nurse ratio, PNR = patients per nurse ratio.
[a] PNR and NNR are calculated on only registered ICU nurses and ICU patients.
[b] Calculated over the patients whole ICU LOS.
- Indicates there is no added value to show these differences / data is not available.

Model Results

For the NNR on day 1 and the mean NNR during the whole ICU admission period, the unadjusted in-hospital mortality significantly increased when NNR exceeded 41 NAS points per nurse. After case-mix adjustment, this association remained significant when NNR exceeded 61 NAS points per nurse (Table 4). After including all type of nurses and all type of patients at the ICU the unadjusted in-hospital mortality significantly increased just when NNR exceeded 61 NAS points per nurse but after case-mix adjustment results were similar to the analyses with only ICU certified nurses and ICU patients (Appendix F). For the PNR on day 1 and the mean PNR during the whole ICU admission period, there were no significant associations found before and after case-mix correction (Table 5). These findings did not change after including all type of nurses (i.e., student- and non-registered ICU nurses) and all type of patients at the ICU (i.e., ICU as well as recovery unit and coronary care unit patients) as can be seen in Table S3 (Appendix G).

Table 4. Association Nursing Activities Score per Nurse Ratios and In-Hospital Mortality (*n* = 29,445).

		[c] Model: NNR		[d] Model: NNR +adjustment	
Covariate	**Range**	**OR**	**(95% CI)**	**OR**	**(95% CI)**
NNR day 1	< 41.2	-	(reference)	-	(reference)
	41.2 - < 61.2	1.174	(1.026 – 1.344)	1.127	(0.972 - 1.306)
	61.2 - < 76.3	1.295	(1.120 – 1.496)	1.257	(1.074 - 1.471)
	≥ 76.3	1.180	(1.014 – 1.373)	1.180	(1.001 - 1.390)
Mean NNR	< 41.2	-	(reference)	-	(reference)
	41.2 - < 61.2	1.187	(1.035 – 1.361)	1.072	(0.922 – 1.245)
	61.2 - < 76.3	1.433	(1.238 – 1.660)	1.285	(1.095 – 1.509)
	≥ 76.3	1.307	(1.120 – 1.525)	1.237	(1.045 – 1.462)

CI = Confidence Interval; NNR = NAS per Nurse Ratio; OR = Odds Ratio
Association models:
[c]Hospital mortality ~ (1|hospital of ICU admission) + NNR
[d]Hospital mortality ~ (1|hospital of ICU admission) + NNR + comorbidities + age + admission type
NNR calculated including only registered ICU nurses and ICU patients.

Table 5. Association Patients per Nurse Ratios and In-Hospital Mortality (n = 29,445)

		[a]Model: PNR		[b]Model: PNR + adjustment	
Covariate	**Range**	**OR**	**(95% CI)**	**OR**	**(95% CI)**
PNR day 1	< 0.96	-	(reference)	-	(reference)
	0.96 - < 1.38	1.132	(0.983 – 1.304)	1.128	(0.996 – 1.316)
	1.38 - < 1.93	1.057	(0.999 – 1.475)	1.089	(0.970 – 1.434)
	≥ 1.93	1.037	(0.852 – 1.263)	1.109	(0.894 – 1.376)
Mean PNR	< 0.97	-	(reference)	-	(reference)
	0.97 - < 1.39	1.119	(0.966 – 1.297)	1.079	(0.918 – 1.268)
	1.39 - < 1.91	1.032	(0.988 – 1.573)	1.021	(0.998 – 1.465)
	≥ 1.91	1.001	(0.789 – 1.204)	1.034	(0.821 – 1.304)

CI = Confidence Interval; PNR = Patients per Nurse Ratio; OR = Odds Ratio
Association models:
[a]Hospital mortality ~ (1|hospital of ICU admission) + PNR
[b]Hospital mortality ~ (1|hospital of ICU admission) + PNR + comorbidities + age + admission type
PNR calculated including only registered ICU nurses and ICU patients.

Discussion

A higher workload per patient per nurse, expressed as NNR greater than 61 was related with a higher in-hospital mortality risk in the Netherlands, whereas PNRs are not associated with in-hospital mortality of ICU patients. This association suggests that it is not the actual number of patients treated by a nurse, but the overall workload during the treatment of these patients that is essential for their outcome.

Previous studies on the association of nursing workload, expressed as PNR, with in-hospital mortality show contradictory results. A systematic review concluded that increased nurse staffing is associated with a lower in-hospital mortality [8], but this review included studies performed on different types of wards and not exclusively on ICUs, which could be an explanation for the different finding. Sakr et al. [15] found that a PNR of greater than 1.5 was independently associated with a lower risk of in-hospital mortality in ICUs (OR, 0.69; 95% CI, 0.53–0.9; $p < 0.001$). In contrast, the literature review and meta-analysis of Numata et al. [6] did not show an association between PNR and in-hospital mortality in ICUs, which confirms our findings. We believe that the PNR is not an appropriate measure for in-hospital mortality in ICUs and that it is better to use the NNR since NAS measures all the work that is being done in caring for each individual patient, both at the bedside and on the ward (e.g., administrative tasks and caring for relatives). To the best of our knowledge, this is the first study to assess the association of workload per nurse, instead of crude number of patients per nurse, with in-hospital mortality. Some studies did investigate the association between nursing workload and in-hospital mortality where nursing workload was measured as the occupancy rate per shift or with the Therapeutic Intervention Scoring System 76 (Keene 1983) [16, 17, 18]. In both studies, a distinction was made between high and low nursing workload and both studies found that patients with a higher nursing workload had a higher risk of in-hospital mortality. Similar results were found by Padilha et al. [18] , who concluded that high NAS values were associated with increased in-hospital mortality. However, all these studies did not take the available nursing capacity into account.

Our primary analysis included ICU patients (recovery unit and coronary care unit patients were excluded) and registered ICU nurses (student- and non-registered ICU nurses were excluded) because the NAS was originally developed to be used for this type of patients and nurses. Other previous studies often did not explicitly describe their inclusion criteria, hampering comparison of the results. Although student nurses also partly take care of the present workload, and therefore can lower the average workload, guiding these student nurses can also be time consuming for the registered ICU and thus increase the workload. Because we were not sure how these effects might influence the results, we therefore performed a sensitivity analysis which showed similar results. This result might partly be explained by the relative low number of student nurses compared with the registered nurses in Dutch ICUs. In our data, we did not have information available on the years of nursing experience. Although our sensitivity analyses (Appendix F and Appendix G) did not suggest that the outcome is associated with the type of nurse, further research should point out whether the years of experience of the ICU nurse are of influence on the outcome.

A limitation of our study is that our data contained information of only 15 out of 82 Dutch hospitals. However, these 15 ICUs form a representative sample of Dutch ICUs regarding size, type, and geographical location of the hospitals. Another limitation of the study is the retrospective design. Due to this retrospective design, we can only assess associations and no causalities. Furthermore, because we use routinely collected data for retrospective analyses, we were not able to correct the missing data, for example, the missing NAS scores, in a later phase of the study. Finally, a limitation of the study is the percentage of patients (15%) with missing NAS values which were excluded from the analysis. This potentially could lead to some biased results, however, as patients with and without NAS scores showed similar patient characteristics (Table 3) and in-hospital mortality rates we assume that the missing values appear randomly and will not influence our results.

Miranda et al. [9] suggest that one FTE nurse corresponds approximately to the work of 100 NAS points. However, this score was not empirically justified and they did not explicitly recommend this ratio as optimum in the context of in-hospital mortality. Until now, it is unclear which NAS score per ICU nurse is optimal. Based on our results, we suggest that one registered ICU nurse should not provide more than 61 NAS points per day. However, more research with more ICUs, preferably in multiple countries, is needed to develop a strong evidence-based recommendation and a clear cutoff point. Furthermore, the NAS is developed specifically for critical care settings, which makes the NNR not a useful indicator for in-hospital mortality in noncritical care settings. Therefore, results of this study are not applicable to other clinical care settings. This study gives a good indication for the usefulness of the NNR compared with the PNR when relating nursing workload to in-hospital mortality. The results could be used for staffing nurses more adequately due to a broader knowledge on the impacts of nurse staffing on patient outcome. ICU managers could use these results to benchmark their current NNRs and to reconsider whether the nurses on their ICUs have a workload which is too high to provide care without a higher chance on in-hospital mortality for their patients. Consequently, ICU nurses could be deployed more efficiently while supporting higher quality of care in terms of survival among their patients. Further research should be performed on the generalizability of these results to ICUs in other countries.

Conclusions

The ratio of ICU patients per nurse is not associated with in-hospital mortality in ICUs while the nursing workload, expressed as the NNR, is positively associated with in-hospital mortality. Therefore, we conclude that it is more important to focus on the workload that the patients generate instead of the actual number of patients.

Acknowledgements

We would like to thank Jorge Salluh for his valuable discussions and remarks during the process of performing analysis. We thank the National Intensive Care Evaluation registry and its participating ICUs for their continuous efforts to collect data for quality improvement and providing us data for this motivating use case.

References

1. Castellan C, Sluga S, Spina E, et al: Nursing diagnoses, outcomes and interventions as measures of patient complexity and nursing care requirement in intensive care unit. J Adv Nurs 2016; 72:1273–1286
2. Tan SS, Bakker J, Hoogendoorn ME, et al: Direct cost analysis of intensive care unit stay in four European countries: Applying a standardized costing methodology. Value Health 2012; 15:81–86
3. Bittner MI, Donnelly M, van Zanten AR, et al: How is intensive care reimbursed? A review of eight European countries. Ann Intensive Care 2013; 3:37
4. Carayon P, Alvarado CJ; Systems Engineering Initiative for Patient Safety: Workload and patient safety among critical care nurses. Crit Care Nurs Clin North Am 2007; 19:121–129
5. Costa DK, Kahn JM: Organizing critical care for the 21st century. JAMA 2016; 315:751–752
6. Numata Y, Schulzer M, van der Wal R, et al: Nurse staffing levels and hospital mortality in critical care settings: Literature review and metaanalysis. J Adv Nurs 2006; 55:435–448
7. Driscoll A, Grant MJ, Carroll D, et al: The effect of nurse-to-patient ratios on nurse-sensitive patient outcomes in acute specialist units: A systematic review and meta-analysis. Eur J Cardiovasc Nurs 2018; 17:6–22
8. Scruth E: Nurse-patient staffing relationship in the ICU: More than just numbers. Crit Care Med 2018; 46:1681–1682
9. Miranda DR, Nap R, de Rijk A, et al; TISS Working Group. Therapeutic Intervention Scoring System: Nursing activities score. Crit Care Med 2003; 31:374–382
10. Stafseth SK, Tønnessen TI, My Diep L, et al: Testing the reliability and validity of the nursing activities score in critical care nursing. J Nurs Meas 2018; 26:142–162
11. Van De Klundert N, Holman R, Dongelmans DA, et al: Data resource profile: The Dutch National Intensive Care Evaluation (NICE) registry of admissions to adult intensive care units. Int J Epidemiol 2015; 44:1850
12. National Intensive Care Evaluation: Dutch National Intensive Care Evaluation (NICE) Foundation, 1995. Available at: http://www.stichting-nice. nl. Accessed December 3, 2018
13. Aiken LH, Sloane DM, Bruyneel L, et al; RN4CAST consortium: Nurse staffing and education and hospital mortality in nine European countries: A retrospective observational study. Lancet 2014; 383:1824–1830
14. Sasichay-Akkadechanunt T, Scalzi CC, Jawad AF: The relationship between nurse staffing and patient outcomes. J Nurs Adm 2003; 33:478–485
15. Sakr Y, Moreira CL, Rhodes A, et al; Extended Prevalence of Infection in Intensive Care Study Investigators: The impact of hospital and ICU organizational factors on outcome in critically ill patients: Results from the extended prevalence of infection in intensive care study. Crit Care Med 2015; 43:519–526

16. Cullen DJ, Civetta JM, Briggs BA, et al: Therapeutic intervention scoring system: A method for quantitative comparison of patient care. Crit Care Med 1974; 2:57–60
17. Kane RL, Shamliyan TA, Mueller C, et al: The association of registered nurse staffing levels and patient outcomes: Systematic review and meta-analysis. Med Care 2007; 45:1195–1204
18. Keene, A. R., & Cullen, D. J. (1983). Therapeutic intervention scoring system: update 1983. Critical care medicine, 11(1), 1-3.
19. Padilha KG, de Sousa RM, Garcia PC, et al: Nursing workload and staff allocation in an intensive care unit: A pilot study according to Nursing Activities Score (NAS). Intensive Crit Care Nurs 2010; 26:108–113

CHAPTER 5

NURSE OPERATION WORKLOAD (NOW): A NEW NURSING WORKLOAD MODEL FOR INTENSIVE CARE UNITS BASED ON TIME MEASUREMENTS

Charlotte C. Margadant, Nicolette F. de Keizer, Marga E. Hoogendoorn, Rob J. Bosman, Jan Jaap Spijkstra, Sylvia Brinkman

Under submission

Abstract

Purpose

Several instruments have been developed to measure patient related nursing workload to support ICU nurses and managers in capacity planning. The commonly used Nursing Activity Score (NAS) and Therapeutic Intervention Scoring System (TISS) are applied to all types of ICU patients. Former research showed that NAS explained 59 to 81% of actual nursing time, whereas the Therapeutic Intervention Scoring System (TISS) described only 43% of the actual nursing time. In both models the development was not based on time measurements. The aim of this study was to develop and validate a new time-based model which can assess patient related nursing workload more accurately and to evaluate whether patient characteristics influence the time spent per activity and therefore should be included in the model.

Methods

Time spent per nursing activity per patient was measured in different shifts in seven Dutch ICUs. Activities from the TISS and NAS scoring systems and five additional activities were measured by using an in-house developed web application. Three different models of varying complexity (1. nursing activities only; 2. nursing activities and case-mix correction; 3. complex model with case-mix correction per nursing activity) were developed to explain the total amount of nursing time per patient. The performance of the three models was assessed in 1000 bootstrap samples using the squared Pearson correlation coefficient (R2), Root Mean Squared Prediction Error (RMSPE), Mean Absolute Prediction Error (MAPE), and prediction bias.

Results

In total 287 unique patients have been observed in 371 different shifts with in total 46.319 nursing activities. Model one's Pearson's R was 0.89 (95%CI 0.86-0.92), model two with case-mix correction 0.90 (95%CI 0.88 – 0.93), and the third complex model 0.64 (95%CI 0.56 – 0.72) compared with the actual patient related nursing workload. Based on the Pearson's R, R2, RMSPE, MAPE, and prediction bias of the three models, there was no significant difference between model one and two, but these two models outperform the more complex model three.

Conclusions

Our model one, which includes a selection of NAS and TISS nursing activities, outperforms existing models in measuring patient related nursing workload, while it includes a lower number of activities and therewith lowers the registration burden. Case-mix correction does not further improve the performance of this model. We called the new model the Nurse Operation Workload (NOW). The patient related nursing workload measured by the NOW gives insight in the actual nursing time needed by patients and can therefore be used to evaluate the average workload per patient per nurse.

Keywords

Nursing workload, time measurements, intensive care, nursing activity score

Introduction

Healthcare costs increased extremely in the last two decades in the Netherlands as well as in other western countries. In the Netherlands, the total costs amounted to € 40.3 billion in 1998 and increased to € 100 billion in 2018 [1]. Intensive care units (ICUs) are responsible for approximately 20% of total hospital costs and on average 50% of this sum is spent on nursing staff [2,3]. There is high urgency to account for and to limit these rising costs. At the same time, we are facing an aging population which may lead to a decline in nursing workforce while at the same time the need for care may increase. Given these facts, a careful analysis on the deployment of nurse staffing is needed [1].

Over the last 50 years several instruments have been developed to measure nursing workload to support ICU nurses and managers in capacity planning [4]. In 1974 Cullen et al. developed the Therapeutic Intervention Scoring System (TISS-76) which quantifies nursing workload based on the complexity and severity of the performed (medical) interventions at the ICU [5]. Two decades later, the Nursing Activities Score (NAS) was developed which quantifies nursing workload based on the performed nursing activities. The weights assigned to each nursing activity were assessed using the work-sampling approach which is based on the probability of performing a particular nursing activity rather than the actual time spent on an nursing activity [6]. This means that the weights assigned to each activity are an estimation of the actual time spent. The NAS scoring system is widely used for clinical and research purposes.

Research showed that the TISS explains only 43% of actual nursing time [3], and the NAS only explains 59 to 81% of the actual nursing time [6,7]. Palese et al. (2016) re-evaluated the face- and content validity of the NAS and showed that the NAS was either under- or overestimating the actual patient related nursing workload, tested with a panel of experts of three intensivists and two researchers with a background in nursing. They concluded that the various weights assigned to each nursing activity need to be reconsidered and therefore they recommended to revise the original NAS [8]. Therefore revision of the workload models is needed.

To support ICU nurses and managers, we believe that a more accurate model to measure the amount of nursing workload and therewith the required nursing time per patient is needed. The work-sampling approach used for the development of the NAS might provide inaccurate estimation of the actual patient related nursing workload whereas time-and-motion techniques are considered more accurate 6. Time-and-motion techniques involve continuous time observations of a single person during a single shift whereby an observer precisely records how much time is being spent on each task. These techniques have not been applied before, probably because they are more labor-intensive and costly than the work-sampling approach [9,10].

The aim of this study was to develop and validate a new time-based model which can more accurately measure patient related nursing workload by assigning new weights to activities from the TISS and NAS scoring systems and if necessary by adding some additional activities, which did not exist during the development of the TISS and NAS, and therefore were not captured in the current models. Both TISS and NAS do not apply for case-mix correction while patient characteristics may influence the time spent per nursing activity, e.g. hygiene procedures could take more time with severely obese patients or communication with the patients' family might consume more time with more severely ill patients. Therefore, the need for additional case-mix correction was also investigated.

Methods

Data collection

Seven randomly chosen Dutch ICUs were invited. Using a time-and-motion technique the time spent per nursing activity, the total time spent per patient, and the frequency of the performed activities were measured in different shifts. We defined patient related nursing workload as the total time of all nursing activities for a patient. We logged the start- and stop times of each nursing activity performed at bedside, using an in-house developed web application. The application included all activities occurring in the NAS, activities from the TISS-28 which do not occur within the NAS, and five additional activities. The included activities from the TISS comprise of: performing neuro checks, proceedings at the intra-aortic balloon pump, performing electrocardiography (ECG), pacemaker, arterial pressure measurement, central venous pressure measurement, heart minute volume measurement, administer blood products, induced hypothermia, and obtain normothermia. The additional activities were added by the researchers since they did not regularly occur when NAS and TISS were developed while they do regularly occur in modern ICU situation and therewith influence the total patient related nursing workload. These variables comprise of: continuous electroencephalography (EEG) monitoring, other vasoactive medication, extracorporeal membrane oxygenation (ECMO) ventilation support, building up Renal Replacement Therapy (RRT), and liver dialysis. All activities are described in Appendix H. Not predefined nursing activities could be registered in the "other category" when necessary. Patients of selected nurses (ICU- and student nurses) were followed by observers during the entire shift. The selected nurses comprised nurses who were taking care of patients that were expected to stay the entire shift, in order to measure as many nursing activities as possible. Observers were researchers of our group (CM and MH) and ten nursing students, who were trained in scoring the nursing activities.

We linked the time-and-motion data to the National Intensive Care Evaluation (NICE) registry to analyze the influence of patient characteristics. The NICE registry contains demographic, physiological, and clinical data of all admitted ICU patients in all Dutch hospitals 11. In the context of the NICE registry ICUs may voluntarily participate in the nursing capacity module in which all TISS and NAS variables and the five additional nursing activities of all patients are collected. Four out of the seven included ICUs for this study participate in the nursing capacity module.

We evaluated the nursing activities which were measured less than 10 times during the measurements, to divide them into new sub-items so they will be included in the model with a reliable weight.

Ethical approval

The Institutional Research Board of the Amsterdam University Medical Centre reviewed the research proposal and waived the need for informed consent (IRB protocol W17_366). We acquired an approval from the Medical Ethical Committee (METC) for this study.

Statistical analyses

Three different linear regression models varying in complexity were developed using the nursing activities of the TISS, NAS, and the five additional nursing activities to measure the total amount of nursing time. We used total time per patient as outcome measure in all models and filled the models with all measured nursing activities as dichotomous variables (1 = occurred; 0 = not occurred).

Model one gives an estimation of the nursing workload per patient without case-mix correction by adding up the coefficients per occurred nursing activity. Model two gives an estimation of the nursing workload per patient with additional case-mix correction by adding up the coefficients per occurred nursing activity. For the third more complex model we first developed separate “sub” models for each nursing activity to estimate the time spent on that particular activity based on the case-mix of the patient. Subsequently, the third model estimated nursing workload per patient by adding up all separate “sub” model estimations of time spent per nursing activity.

The case-mix variables included in model two and the "sub" models of model three were: admission type (medical/urgent surgical/elective surgical), body mass index (BMI), age, severity of illness expressed as the APACHE III score, type of shift (day/evening/night), day of ICU admission (1st day, 2nd day, etc.), renal insufficiency, respiratory failure, immunological insufficiency, neoplasm, hematological malignancy, cardio vascular insufficiency, and cirrhosis. Furthermore we included the interaction between severity of illness at ICU admission and day of ICU admission. All continuous covariates (age, BMI, APACHE III score, and day of ICU admission) were included as splines in the model.

We used the Least Absolute Shrinkage and Selection Operator (Lasso) procedure to shrink the models by selecting only the covariates that significantly influence the patient related nursing workload 12. The selected covariates and found interaction terms were used in the final models. All statistical analyses were performed using R statistical software, version 3.3.2.7 13.

Performance assessment

To evaluate the ability of the different models to accurately measure patient related nursing workload, the performance was assessed using the Pearson's R correlation coefficient, the R squared (R2) 14; Root Mean Squared Prediction Error (RMSPE); Mean Absolute Prediction Error (MAPE); and prediction bias. The R^2 represents the fraction of the variance in observed time which is explained by a model, the RSMPE estimates the mean residual, the MAPE expresses average model prediction error for the observed time, or unexplained standard error of predictions obtained by using the model 15, and the prediction bias estimates the magnitude and direction of the average model compared to the observed time. Interpretation, formula and range of these measures are described in more detail in Appendix I. The performance of the models was analyzed using the bootstrap method in which the performance measurements and the associated 95% confidence intervals were assessed in 1000 bootstrap samples 16. We used overlap of confidence intervals to test whether the models were significantly different from each other.

Results

Baseline results

Measurements took place between November 1st 2016 and October 1st 2017 in mixed medical-surgical ICUs located in university hospitals (n=1), teaching hospitals (n=4), and non-teaching hospitals (n=2). In total 287 unique patients were observed in 371 different shifts with in total 46.319 nursing activities. Of these 371 shifts 133 (36%) were day shifts, 135 (36%) evening shifts, and 103 (28%) night shifts.

The 287 patients included in our study had a higher in-hospital mortality (22.3% versus 13.0%), higher severity of illness (median APACHE III 68.0 versus 50.0), and longer ICU length of stay (median of 3.2 days versus 1.0) compared to all Dutch patients. The comorbidities acute renal failure, cirrhosis, and respiratory insufficiency occurred more frequent in the included patients than in the total Dutch ICU population (see Table 1).

Table 1. Patient Characteristics

	Patients in measurements	**All Dutch patients during the months of measurements**
Number of patients, N	287	100145
Age, median [IQR]	66.0 [56.0 – 76.0]	66.0 [55.0 – 75.0]
BMI, median [IQR]	26.0 [23.6 – 28.7]	25.9 [23.1 – 28.4]
Admission type		
-Medical, N (%)	121 (42.2)	51290 (52.7)
-Surgical: elective and urgent, N (%)	151 (52.6)	45905 (47.2)
In-hospital mortality, N (%)	85 (22.3)	13017 (13.0)
APACHE III score, median [IQR]	68.0 [47.3 – 96.5]	50.0 [35.0 – 71.0]
ICU LOS (in days), median [IQR]*	3.2 [0.9, 14.8]	1 [0.7 – 4.0]
Comorbidities		
Acute renal failure, N (%)*	37 (12.9)	9211 (9.2)
Cardiovascular insufficiency, N (%)	16 (4.2)	4257 (4.3)
Chronic renal failure[1], N (%)	25 (6.7)	7976 (7.9)
Chronic respiratory insufficiency, N (%)*	7 (2.4)	4620 (4.6)
Cirrhosis, N (%)*	1 (3.5)	1751 (1.7)
COPD, N (%)	36 (12.5)	13304 (13.3)
Diabetes, N (%)	68 (17.8)	16273 (16.2)
Gastrointestinal bleeding, N (%)	2 (0.7)	2263 (2.3)
Hematologic malignancy, N (%)	6 (2.1)	2143 (2.1)
Immunological insufficiency, N (%)	16 (5.6)	8290 (8.3)
Neoplasm, N (%)	9 (3.1)	4506 (4.5)

COPD: Chronic Obstructive Pulmonary Disease; IQR: Interquartile Range; LOS: length of stay

[1] Chronic renal failure consists of chronic renal insufficiency and chronic dialysis

* Indicates a significant P-value of <0.05

Results time-and-motion techniques

Several activities were measured less than 10 times. An expert panel of two ICU physicians (JJS, RB) and an ICU nurse (MH) have indicated that most of these activities very rarely occur in current daily ICU practice, do not generate any patient related nursing workload, or are covered by another nursing activity. However, for three activities the experts indicated that these activities do sometimes occur in daily ICU practice and are of value to include in the models. Therefore the experts were asked to validate or estimate the mean time spent on these activities: cardioversion, gastro- or coloscopy, and extracorporeal membrane oxygenation (ECMO). Moreover, to take all necessary activities into account in the models, we created five new aggregated variables of which the experts estimated that the original variables were very similar as concerned the type of the activities, as well as the time to perform them. These five aggregated variables comprise of: other medication (thrombolysis, other vasoactive medication, continuous epileptic infusion), isolation (barrier isolation, reversed isolation, cytostatic barrier, and droplet isolation), renal replacement therapy (carried out by an ICU nurse or by a dialysis nurse and liver dialysis), thermoregulation (thermoregulation and induced hypothermia), and monitoring (ECG monitoring, heart minute volume measurement, intracranial pressure measurement, and continuous Electroencephalography). The following nursing activities were measured less than 10 times in total and could not be divided over the aggregated categories: cardioversion, gastro- or coloscopy, balloon tamponade, cardiac assist device, cardiopulmonary resuscitation, pacemaker, pulmonary- or left atrium catheter, patient in prone position, placing/replacing artificial airway, and placing new drain at ICU. Total time per patient (the outcome measure of the models) had a skewed distribution and was therefore taken as log transformation during model development. To provide easy to use coefficients for the users of this model, we did not log transform the outcome measure in the final model. The log transformed coefficients are shown in Appendix J. These coefficients can be add up if the activity occurred and this sum needs to be back transformed to obtain the total time per patient.

Excluded variables after Lasso regression

After performing Lasso regression on the models, the following nursing activities or categories of activities were excluded from the models, since they did not significantly influence the measured patient related nursing workload: administer inotropes (1 time/shift), proceedings at the central venous catheter, proceedings at the central line, central venous pressure measurement, spontaneous respiration, tracheal suction, parenteral feeding, collecting blood (1 or 2 times/shift), patient in sitting position, administration (between 1 and 2 hours), and monitoring (ECG monitoring, heart minute volume measurement and intracranial pressure measurement).

Results performance

The performance measures and corresponding 95% confidence intervals (CI) obtained in the bootstrap samples are showed in Table 2. The three models' Pearson's R correlation coefficient varied between 0.64 and 0.90, the R2 between 0.41 and 0.81, the RMSPE between 0.68 and 0.82, the MAPE between 0.3 and 0.37, and the prediction bias between -0.33 and -0.02. Based on the RMSPE, MAPE, and prediction bias of the three models, there was no significant difference between model one and two. Based on the R2 and prediction bias, model one and two performed significantly better compared to the more complex model three. Model one explains 0.89 (0.86 – 0.92), model two explains 0.90 (0.88 – 0.93) and the third model explains 0.64 (0.56 – 0.72) of patient related nursing workload.

All included nursing activities (which remained after Lasso regression) are shown in Table 2. This Table shows the final model. The corresponding definitions of the several nursing activities are described in Appendix H. By adding up the intercept and the coefficients per occurred nursing activity from this table, the patient related nursing workload can be calculated.

Table 2. Coefficients per nursing activity in the final model[1]

Model item	**Coefficients, in minutes**
Intercept	83.38
Central nervous system	
Neurological checks, 1 time /shift	7.07
Neurological checks, 2-4 times /shift	4.19
Neurological checks, >=4 times /shift	30.61
Cardiovascular	
Administer inotropes, >1 times/shift	20.21
Cardioversion*	30
Administer antiarrhythmics	14.6
Administer anticoagulation	-1.35
Medication via arterial line	17.47
Arterial pressure measurement	1.12
Respiratory	
Oxygen	15.16
Pressure support	8.07
Controlled ventilation	15.24
ECMO*	60
Tractus digestive	
Enteral feeding	18.63
Gastro- or coloscopy*	60
Renal	
Building up renal replacement therapy	62.4
RRT by an ICU or dialysis nurse	
Urinary tract catheter	-3.7
Blood	
Administer blood products, 1 or more times/shift	24.71
Administer blood products, 3 or more times/shift	58.58
Administer blood products, 5 or more times/shift	62.08
Blood sampling, 3 times/shift	4.92
Blood sampling, 4 times/shift	3.28
Blood sampling, 5 or more times/shift	3.11
Monitoring, titration, mobilization and positioning	
Bedside, <1 hour	-24.74
Bedside, >1 hour	28.61

Positioning by one nurse	8.64
Positioning by two nurses	22.78
Positioning by three or more nurses	37.55
Infectology	
Administer antibiotics, 1 or 2 times/ shift	1.07
Administer antibiotics, 3 or more times/ shift	8.35
Isolation (barrier isolation, droplet isolation, cytostatic barrier and reversed isolation)	-2.41
Hygiene	
Hygiene activities, <1 hour/shift	19.11
Hygiene activities, >1 hour/shift	93.81
Administration and support	
Care for patient or relatives, <1 hour/shift	32.57
Care for patient or relatives, >=1 hour/shift	123.61
Administration, <1 hour/shift	-50.15
Administration, >2 hours/shift	64.33
Administration, >3 hours/shift	150.82
Other	
Other medication (other vasoactive medication and epileptic medication via IV)	6.77
Thermoregulation	11.92
Care for drains	2.4

* *For these activities an expert opinion is used, since these activities occurred less than 10 times during the measurements, but do sometimes occur in daily practice.*

[1] See Appendix H for the definitions of the items used in the model.
[2] These items form together the final model (after Lasso regression).

Table 3. Performance of models expressed by Pearson's R, RMSPE, MAPE and bias.
Means and confidence intervals were obtained with bootstrap sampling.

Model	**Pearson's R (CI)**	**R squared (CI)**	**Root mean squared prediction error (RMSPE) (CI)**	**Mean absolute prediction error (MAPE) (CI)**	**Bias (CI)**
Model one: Nurse Operating Workload (NOW) (nursing activities only)	0.89 (0.86 – 0.92)	0.79 (0.76 - 0.86)	0.68 (0.4 – 1.0)	0.31 (0.24 – 0.38)	-0.02 (-3.59 – 3.64)
Model two (nursing activities and case-mix)	0.90 (0.88 – 0.93)	0.81 (0.77 – 0.86)	0.82 (0.35 – 1.39)	0.30 (0.22 – 0.38)	-0.04 (-3.26 – 3.17)
Model three (complex: case mix correction per nursing activity)	0.64 (0.56 – 0.72)	0.41 (0.31 – 0.52)	0.71 (0.53 – 0.91)	0.37 (0.33 – 0.42)	-0.33 (-0.41 - -0.26)

Discussion

In this study we developed and validated three new models to measure patient related nursing workload using nursing activities of the TISS, NAS, and five additional modern nursing activities. Case-mix adjustments, which were included in two of the three models, did not significantly improve the performance of the models and thus are not necessary to include in the models. We recommend to use model one, which we call Nurse Operating Workload (NOW) as this model has a comparable performance as model two, outperforms model three, and uses less covariates making it a more simple model to implement in practice. The NOW model explains 89% of the actual needed nursing time for patient care. Besides the patient-related time, nurses spend time on other tasks, such as participating in an emergency team, taking a course, or supervising a student. Our models do not take these activities into account and this should be considered when using the model for nursing capacity planning.

Our hypothesis was that severity of illness was probably of influence on nursing workload. However, our results showed that this is not the case, which is in line with earlier studies [17,18]. For instance, the NAS in a Medium Care Unit is not lower compared to the NAS at an ICU [19,20]. Our findings strengthen the suggestion that case-mix adjustment is not needed in nursing workload models.

The performance of the NAS was tested in earlier studies which advised revision of the NAS [7,8,21]. The validity study of Stafseth et al. (2018) showed a Spearman's' correlation of 39% for the NAS with the actual nursing workload [22]. In the study of Margadant et al. (2019) the validity of the NAS was tested on the same data as in the current study, which makes comparisons more reliable [7]. This study gives a Pearson's correlation of 59% (R^2=0.35) between the NAS and the actual nursing time measured with time-and-motion techniques. When considering the Pearson's R and the R^2, the performance of the newly developed models in our study are significantly better compared to the NAS and TISS [5,22]. To our knowledge, no other studies have been performed on the evaluation of a patient related nursing workload model based on time measurements.

One of the limitations of this study is the zero or low frequency of a few nursing activities during the measurements. For three of these nursing activities, experts think that despite their low frequency they should be included in the models since when they occur, these activities will generate high nursing workload. For this reason, an expert panel of two ICU physicians and an ICU nurse have estimated times for these activities. The other zero or low frequency nursing activities are excluded from model development. To determine whether it was a coincidence that these activities did not occur during our measurements, we assessed their occurrence in the nursing capacity module of the NICE registry containing information on all nursing shifts of 15 Dutch hospitals over a three year time period. The occurrence of the activities during our measurements were comparable with the occurrence in the NICE database, e.g. the nursing proceeding 'induced hypothermia' occurred in 0.5% of the patients in this study and in 0.6% of all patients in the NICE registry. If in practice the excluded activities will occur, the corresponding patient related nursing workload will partly be represented in other activities that are included in the model.

Furthermore, patients included in our time and motion measurements seem to be more severely ill and consequently had a longer length of stay compared to all Dutch patients in the same time period. This is probably caused by our selection mechanism. In order to measure as many nursing activities as possible we have chosen to randomly select nurses who took care of patients that were expected to stay the whole shift and these patients are probably more severely ill. We don't think this has biased our results since our analyses took case-mix into account. If case-mix was of any influence, it is likely that it was noticed in these more severe ill patients, which was not the case. Moreover, severity of illness was shown not to be related to nursing time [3,23].

Another limitation of our study was the large proportion of data derived from one hospital, which could have influenced the results. We were not able to perform a sensitivity analysis without this hospital since the remaining data has a lack of power to run the models. Finally, the results of the performance tests should be interpreted with some caution because we tested the performance of the models on the same data as from which the models were developed. However, we used bootstrapping with 1000 samples, a robust method [16], to make the results more generalizable. Furthermore, the validation of the NAS was also tested on this same dataset in an earlier study [7].

A strength of the performed study is the large sample size, considering the measurements with time-and-motion techniques. Furthermore, our study did not interfere with the working routine of nurses as in other (work-sampling) studies because the measurements in our study were carried out by independent observers. Another strength of this study is the research question related to the effect of case-mix for each nursing activity, which has not been done in other nursing workload studies. This correction enabled us to find out whether case-mix correction per nursing activity led to more accurate measures for specific types of patients, which was not the case. We believe this is important knowledge as this suggests that benchmarking of nurse to workload ratios across different hospitals is possible due to the fact that the time spent on nursing activities is not influenced by the patient population but by the nursing activities themselves. Finally, the patient population in our measurements is mainly comparable to patients from all Dutch ICUs, which make the models generalizable for Dutch ICUs and probably for other comparable Western ICUs. In non-Western countries the working pace and/or the nursing workload might differ, which could increase the total patient related nursing workload. This is not accounted in the model.

By using Lasso regression we were able to develop a model that includes less covariates while performing well compared to existing patient related nursing workload models. The advantage of including less activities in our model is a reduction in registration burden among ICU nurses.

Conclusion

We were able to develop a model to accurately estimate patient related nursing workload and which is simpler in use than existing TISS and NAS models while it's performance is significantly better. Patients' case-mix correction does not seem to improve the model. Our recommended Nurse Operation Workload (NOW) model can easily be used to assess the patient related nursing workload by adding up the intercept and the coefficients per nursing activity that occurred during an 8 hour shift. The patient related nursing workload calculated by this model could be used to evaluate whether nurses were deployed according to their workload. More research is needed to validate the NOW model in other hospital settings (for instance other hospital wards or ICUs in non-Western countries) with time-and-motion techniques to assess the generalizability.

References

1. Central Bureau for Statistics (CBS), 18th of May 2017 Zorguitgaven in drie benaderingen; aanbieders van zorg. Derived from: https://opendata.cbs.nl/#/CBS/nl/dataset/84054NED/table?ts=1569919225724.
2. Oostenbrink, J. B., Buijs-Van der Woude, T., van Agthoven, M., Koopmanschap, M. A., & Rutten, F. F. (2003). Unit costs of inpatient hospital days. Pharmacoeconomics, 21(4), 263-271.
3. 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.
4. Hoogendoorn, M. E., Margadant, C. C., Brinkman, S., Haringman, J. J., Spijkstra, J. J., & de Keizer, N. F. (2019). Workload scoring systems in the Intensive Care and their ability to quantify the need for nursing time: a systematic literature review. International Journal of Nursing Studies, 103408.
5. Cullen, D. J., Civetta, J. M., Briggs, B. A., & Ferrara, L. C. (1974). Therapeutic intervention scoring system: a method for quantitative comparison of patient care. Critical care medicine, 2(2), 57-60.
6. Urden, L. D., & Roode, J. L. (1997). Work sampling: a decision-making tool for determining resources and work redesign. Journal of Nursing Administration,27(9), 34-41.
7. Margadant, C.C., Brinkman, S., Hoogendoorn, M.E., Bosman, R.J., Spijkstra, J.J., de Keizer, N.F. (2019). Validation of the Nursing Acitivities Score (NAS) using time-and-motion techniques in Dutch intensive Care Units. [under submission].
8. Palese, A., Comisso, I., Burra, M., DiTaranto, P. P., Peressoni, L., Mattiussi, E., & Lucchini, A. (2016). Nursing Activity Score for estimating nursing care need in intensive care units: findings from a face and content validity study. Journal of nursing management, 24(4), 549-559.
9. Finkler, S. A., Knickman, J. R., Hendrickson, G., Lipkin Jr, M., & Thompson, W. G. (1993). A comparison of work-sampling and time-and-motion techniques for studies in health services research. Health services research, 28(5), 577.
10. Burke, T. A., McKee, J. R., Wilson, H. C., Donahue, R. M., Batenhorst, A. S., & Pathak, D. S. (2000). A comparison of time-and-motion and self-reporting methods of work measurement. Journal of Nursing Administration, 30(3), 118-125.
11. Dutch National Intensive Care Evaluation (NICE) foundation (2017). Derived from: http://www.stichting-nice.nl.
12. Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Methodological), 58(1), 267-288.
13. R Core Team (2015). R: A language and Environment for Statistical Computing [computer program]. Vienna, Austria: R foundation. Derived from: http://www.r-project.org/.

14. Steyerberg, E.W. (2009). Evaluation of Performance. Clinical Prediction models: a practical approach to development, validation, and updating. Springer; p. 83-99.
15. Barnston, A. G. (1992). Correspondence among the correlation, RMSE, and Heidke forecast verification measures; refinement of the Heidke score. Weather and Forecasting, 7(4), 699-709.
16. Parke, J., Holford, N. H., & Charles, B. G. (1999). A procedure for generating bootstrap samples for the validation of nonlinear mixed-effects population models. Computer methods and programs in biomedicine, 59(1), 19-29.
17. Debergh, D. P., Myny, D., Van Herzeele, I., Van Maele, G., Miranda, D. R., & Colardyn, F. (2012). Measuring the nursing workload per shift in the ICU. Intensive care medicine, 38(9), 1438-1444.
18. Guccione, A., Morena, A., Pezzi, A., & Iapichino, G. (2004). The assessment of nursing workload. Minerva anestesiologica, 70(5), 411-416.
19. Swiger, P. A., Vance, D. E., & Patrician, P. A. (2016). Nursing workload in the acute-care setting: A concept analysis of nursing workload. Nursing outlook, 64(3), 244-254.
20. Armstrong, E., de Waard, M. C., de Grooth, H. J. S., Heymans, M. W., Miranda, D. R., Girbes, A. R., & Spijkstra, J. J. (2015). Using nursing activities score to assess nursing workload on a medium care unit. Anesthesia & Analgesia, 121(5), 1274-1280.
21. Stafseth, S. K., Tønnessen, T. I., Diep, L. M., & Fagerstrøm, L. (2018). Testing the Reliability and Validity of the Nursing Activities Score in Critical Care Nursing. Journal of nursing measurement, 26(1), 142-162.
22. Dick W, Pehl S, Tzanova I, Heinrichs W, Brost F, Eich P. Physician and nursing (personnel) requirements for ICUs. Therapeutic Intervention Scoring System (TISS) versus time requirements for patient care--a comparative study in an interdisciplinary surgical intensive care unit. Clin Intensive Care, 1992, 3:116-121.
23. Miranda, D. R., de Rijk, A., & Schaufeli, W. (1996). Simplified Therapeutic Intervention Scoring System: the TISS-28 items--results from a multicenter study. Critical care medicine, 24(1), 64-73.

CHAPTER 6

EFFECT OF PATIENT CHARACTERISTICS AND CONTEXTUAL FACTORS ON NEEDED NURSING TIME IN INTENSIVE CARE UNITS

Charlotte C. Margadant, Sylvia Brinkman, Marga E. Hoogendoorn, Rob J. Bosman, Jan Jaap Spijkstra, Nicolette F. de Keizer

Under submission

Abstract

Background

Workforce planning in ICUs is complex due to differences in needed nursing care per patient and adventitious emergencies leading to unplanned admissions. Needed nursing time might depend on patient characteristics or contextual factors such as type of shift or size of ICU. The objective of this study was to get insight into the variability of time spent by ICU nurses on their patients for different patient- and contextual characteristics.

Methods

To assess the time spend on nursing activities in the ICU, we used data from the nursing capacity module of the National Intensive Care Evaluation (NICE) registry including the occurrence of nursing activities to calculate the Nursing Activities Score (NAS) and the total nursing time in minutes per shift based on the Nurse Operation Workload (NOW) model.

Results

We analyzed 19.688 unique patients from 9 ICUs during 159.733 different shifts. Significant differences in nursing time per shift spend on patients were found for all included patient characteristics and contextual factors, however most of these differences were not clinical relevant. Among others, we found that medical patients required more nursing care compared to elective- and surgical admitted patients (2.16 hours vs. 2.12 and 1.9 hours, $p<0.001$). Patients required less care from a nurse during the night shift compared to patients in the day- or evening shift (1.97 hours vs. 2.08 and 2.14 hours, $p<0.001$).

Conclusion

The results of this study showed that the time ICU nurses spent on their patient depends on patient characteristics and contextual factors. It is therefore not recommended to plan nursing capacity on a fixed ratio of patients per nurse but to include patient characteristics and contextual factors while planning nursing capacity and during assignment of patients to nurses per shift.

Keywords

ICU, nursing workload, nursing time, nurse capacity, Nursing Activities Score (NAS), Nurse Operation Workload (NOW) .

Introduction

A high nursing workload can lead to adverse events for patients in intensive care units (ICUs), such as pneumonia, infections, and in-hospital mortality [1-4]. Moreover, a high nursing workload can also lead to negative events for ICU nurses themselves, such as burn-out, stress, and job dissatisfaction [5]. When information on needed nursing time is used to assign patients to ICU nurses, nurses could be allocated more efficiently and workforce could be planned with more caution [6].

Workforce planning in ICUs is complex due to differences in needed nursing time per patient and adventitious emergency leading to unplanned admissions. Healthcare managers need to make decisions to deliver efficient and effective workforce in the context of fixed budget constraints [7]. When scheduling ICU care, most ICUs apply a fixed ratio of patients per nurses with some subjective adjustments based on the patients' severity of illness. However, if the current Corona crisis learned us anything, it is most likely how scarce nursing capacity, especially on the ICU, really is. An objective and evidence-based capacity planning should consider which differences in needed nursing time per patient group or diagnosis need to be taken into account. Possible differences in needed nursing time could be found in patient characteristics, such as age, BMI, admission type, and mortality risk or in contextual factors, such as shift type and size of ICU. Until now, no studies have been published on determinants of needed nursing time.

The objective of this study is to get insight into the variability of time spent by ICU nurses on their patients per shift (expressed as needed nursing time) among different patient- and contextual characteristics and thereby contributing to evidence-based nurse capacity planning.

Methods

Data

We used data from the Dutch National Intensive Care Evaluation (NICE) quality registry to describe demographic, diagnostic, and severity of illness data of the included ICU patients 8. All Dutch ICUs collect these data for all their patients. We used data of all patients admitted between January 1st of 2017 until July 1st of 2019 from the nursing capacity module of the NICE, which is a voluntary module in which nine ICUs participate. The nursing capacity module of the NICE includes the occurrence of all nursing activities to calculate the Nursing Activities Score (NAS) 9 and to calculate the total nursing time in minutes per shift based on a model we recently developed, the Nurse Operation Workload (NOW) model [10].

Recent validation of the NAS showed that the model only explained 59% of needed nursing time per shift, therefore we developed a more accurate nursing workload model called the NOW. The NOW consists of a selection of Therapeutic Intervention Scoring System (TISS)[11] and NAS variables. In the NOW, minutes are assigned to all nursing activities which can be add up to obtain the total nursing time per shift per patient. The NAS, NAS time, and NOW time will be reported in our results for this cohort of patients.

Data analyses

We transformed the NAS per patient into time by multiplying the number of NAS points with 3.84 minutes. According to Miranda et al. (2003) 100 NAS points correspond to 100% of care time provided by one nurse during a shift and hence one NAS point corresponds to 1% of care time provided by one nurse during an 8-hour shift [12]. The authors of the NAS described that a nurse performs patient-related activities for 80% of the 8-hours shift. For this reason one NAS point corresponds to 3.84 minutes of nursing time during an 8-hour shift (((8 hours * 60mins)/100)*0.8) [13].

Based on the availability of data in the NICE registry we analyzed the following patient characteristics of which we thought that they would influence total nursing time per shift: type of admission (medical, emergency surgical, and elective surgical), APACHE IV mortality risk (low = less than 30%, medium = 30-69%, and high 70% or more), Body Mass Index (BMI) (underweight = BMI less than 18.5, normal weight = BMI between 18.5 and 24.9, overweight = BMI between 25.0 and 29.9, and obese = BMI of 30.0 or more), age (<50 years, 50-79 years, and >=80 years), and three common and large groups defined on reason for ICU admission (sepsis, Community Acquired Pneumonia, and Out of Hospital Cardiac Arrest). For the contextual factors we analyzed type of shift (day, evening, and night) and ICU size (<12 ICU beds and >= 12 ICU beds). This last cut-off point is based on the Dutch quality standard for ICUs [14].

To observe the differences in needed nursing time per shift across different patient- and contextual groups, we calculated the median and interquartile ranges (IQR) for each group as needed nursing time per shift is not normally distributed. To check differences between the groups, we performed Wilcoxon Rank Tests and considered groups as significantly different when a p-value lower than 0.05 was found. All statistical analyses were performed using R statistical software, version 3.3.3 15.

Ethical approval

The Institutional Research Board of the Amsterdam University Medical Centre reviewed the research proposal and waived the need for informed consent (IRB protocol W17_366).

Results

We included 19.688 unique patients in 159.733 different shifts in nine ICUs between January 1st 2017 and July 1st 2019. The median age of the patients was 67 years and the in-hospital mortality was 20.6% (Table 1). Three of the included ICUs were university hospitals, three were teaching hospitals, and three were non-teaching hospitals (Table 2). In general, the NOW gives a lower needed nursing time per shift compared to the NAS time in all patient- and contextual categories.

Table 1. Patient Characteristics.

Variable	**Patients in nursing capacity module NICE**
Number of nursing shifts, N	159.733
Number of patients, N	19.688
Age, median [IQR]	67.0 [57.0 – 74.0]
BMI, median [IQR]	26.3 [23.7 – 30.9]
Admission type	
-Medical, N (%)	7852 (39.9)
-Surgical: urgent and elective, N (%)	5485 (27.9)
In-hospital mortality, N (%)	2749 (20.6)
ICU LOS (in days), median [IQR]	0.8 [0.7 – 0.9]
Comorbidities	
Acute renal failure, N (%)	2073 (10.5)
Cardiovascular insufficiency, N (%)	769 (4.0)
Chronic renal failure[1], N (%)	753 (3.8)
Chronic respiratory insufficiency, N (%)	320 (1.6)
Cirrhosis, N (%)	75 (0.4)
COPD, N (%)	1560 (7.9)
Diabetes, N (%)	2464 (12.5)
Gastrointestinal bleeding, N (%)	214 (1.1)
Hematologic malignancy, N (%)	201 (1.0)
Immunological insufficiency, N (%)	929 (4.7)
Neoplasm, N (%)	367 (1.9)

Chronic Obstructive Pulmonary Disease; IQR: Interquartile Range
[1] Chronic renal failure consists of chronic renal insufficiency and chronic dialysis

Table 2. ICU characteristics.

	ICUs in nursing capacity module NICE
Number of university hospitals (%)	3 (33%)
Number of teaching hospitals (%)	3 (33%)
Number of non-teaching hospitals (%)	3 (33%)
Median number of ICU beds per ICU (IQR)	10 [9 – 18]

For the patient characteristics, both the NAS and NOW showed that medical patients needed significantly more nursing time per shift compared to emergency- and elective surgical admitted patients (1.9% and 13.7% respectively, $p<0.001$). Emergency surgical patients needed 12% more nursing time per shift compared to elective surgical patients ($p<0.001$). Patients with a high APACHE IV mortality risk needed significantly (19%) more nursing time per shift compared to patients with a low APACHE IV morality risk (<0.001). For the NOW, all BMI categories were statistically different from each other ($p<0.001$). Obese patients required significantly more nursing time per shift compared to patients with normal weight and underweight (both 3.5%) for the NAS ($p<0.001$). Very small differences were measured for the different age groups with NOW. The needed nursing time per shift increased with the age of the patient for the NOW. Patients between 50-79 years needed 1.5% more nursing time per shift and patients of 80 years and older needed 2.0% more nursing time per shift compared to patients below 50 years ($p<0.001$). For the NAS, there was no difference in needed nursing time per shift between patients between 50-79 years and patients of 80 years and older. Finally nursing time per shift of patients with Sepsis, CAP, and OHCA was comparable with the nursing time per shift of patients with a medium APACHE IV mortality risk and did not clinically differ from less severe ill patients. All found statistically significant differences were small and probably not clinically relevant.

For the contextual characteristic 'type of shift', both NOW and NAS showed that the patients in the day- and evening shift required significantly more nursing time per shift compared to patients in the night shift (respectively 5.6% and 8.6% more NOW time, p<0.001). We found no relevant differences between patients in ICUs with less than 12 beds compared to ICUs with 12 or more beds for the NOW and the NAS.

Table 3: Patient characteristics. N = 159.733 nursing shifts.

Variable		*Categories (median observed nursing time/points per patient, IQR)*			
Total time	NOW time (h)	2.06 [1.55 – 2.78]			
	NAS time (h)	2.60 [1.98 – 3.38]			
	NAS points	40.60 [31.00 – 52.8]			
Admission type		**Medical (N=90.198)**	**Emergency surgical (N=21.994)**	**Elective surgical (N=47.186)**	
	NOW time (h) *	2.16 [1.60 – 3.03]	2.12 [1.59 – 2.74]	1.90 [1.44 – 2.40]	
	NAS time (h) *	2.75 [2.16 – 3.59]	2.52 [1.98 – 3.34]	2.34 [1.90 – 2.96]	
	NAS points *	42.90 [33.7 – 56.10]	39.40 [31.00 – 52.20]	36.6 [29.70 – 46.30]	
APACHE IV mortality risk		**Low (<0.30) (N=94.075)**	**Medium (0.30 – 0.69) (N=39.724)**	**High (>=0.70) (N=18.493)**	
	NOW time (h) [a]	1.94 [1.45 – 2.62]	2.27 [1.74 – 3.11]	2.31 [1.81 – 3.08]	
	NAS time (h) *	2.50 [1.90 – 3.22]	2.83 [2.27 – 3.67]	2.73 [2.26 – 3.67]	
	NAS points *	39.10 [29.70 – 50.30]	44.30 [35.40 – 57.30]	42.60 [35.30 – 57.30]	

BMI group		Underweight (BMI < 18.5) (N=3.407)	Normal weight (BMI 18.5 – 24.9) (N=60.131)	Overweight (BMI >=25 – 29.9) (N=55.468)	Obesity (BMI >=30) (N=38.158)
	NOW time (h) *	1.99 [1.39 – 2.93]	2.06 [1.55 – 2.70]	2.06 [1.57 – 2.76]	2.13 [1.62 – 3.03]
	NAS time (h) [b]	2.68 [2.06 – 3.46]	2.56 [1.98 – 3.32]	2.57 [1.98 – 3.33]	2.71 [2.10 – 3.57]
	NAS points [b]	41.90 [32.20 – 54.00]	40.00 [31.00 – 51.80]	40.20 [31.00 – 52.10]	42.40 [32.80 – 55.80]
Age		**<50 years (N=22.978)**	**50-79 years (N=119.566)**	**>=80 years (N=17.189)**	
	NOW time (h) *	2.04 [1.45 – 2.71]	2.07 [1.56 – 2.79]	2.08 [1.60 – 3.81]	
	NAS time (h) [c]	2.45 [1.90 – 3.33]	2.61 [1.99 - 3.39]	2.60 [1.99 – 3.39]	
	NAS points [c]	38.30 [29.70 – 52.00]	40.80 [31.10 – 52.90]	40.60 [31.10 – 52.90]	
Patient subgroups		**Sepsis (N=12.376)**	**CAP (N=10.420)**	**OHCA (N=8.067)**	
	NOW time (h) *	2.29 [1.66 – 3.41]	2.20 [1.61 – 3.21]	2.38 [1.83 – 3.25]	
	NAS time (h) *	3.03 [2.34 – 3.82]	2.90 [2.28 – 3.64]	3.00 [2.28 – 3.68]	
	NAS points *	47.30 [36.60 – 59.70]	45.30 [35.60 – 56.90]	43.2 [35.70 – 57.50]	

OHCA: Out of Hospital Cardiac Arrest; CAP: Community Acquired Pneumonia
* Significantly different between all categories (p = <0.001)
[a] Significantly different between low and high APACHE IV mortality risk and between low and medium APACHE IV mortality risk (p = <0.001)
[b] Statistically different between patients with underweight and normal weight, underweight and obesity; normal weight and obesity (p = <0.001)
[c] Statistically different between patient <50 years and patient of 50 years and older (p = <0.001)

Table 4: contextual factors. N = 159.733 nursing shifts.

Variable		*Categories (median total observed nursing time/points per patient, IQR)*		
Shift type		**Day shift (N=56.478)**	**Evening shift (N=52.122)**	**Night shift (N=51.133)**
	NOW time (h) *	2.08 [1.63 – 2.93]	2.14 [1.59 – 2.99]	1.97 [1.43 – 2.50]
	NAS time (h) *	2.61 [1.93 – 3.45]	2.73 [2.09 – 3.55]	2.43 [1.98 – 3.09]
	NAS points *	40.80 [30.10 – 53.90]	42.60 [32.60 – 55.50]	38.00 [31.00 – 48.30]
Number of beds per ICU		**ICUs with less than 12 beds (N=34.346)**	**ICUs with 12 beds or more (N=125.387)**	
	NOW time (h) *	1.59 [1.16 – 2.22]	2.15 [1.73 – 2.94]	
	NAS time (h) *	2.80 [2.25 – 3.35]	2.52 [1.98 – 3.39]	
	NAS points *	43.80 [35.20 – 52.30]	39.40 [31.00 – 52.90]	

* Significantly different between all categories

Discussion

Our study showed that needed nursing time significantly differs between several patient subgroups and contextual characteristics. Much to our surprise however, most differences were small and clinically irrelevant. Medical ICU patients and emergency surgical patients had a significantly higher need for nursing time per shift compared to elective surgical patients. Medical and emergency surgical admissions are normally acute and unexpected and take a lot of time, especially in the first few shifts, where elective surgical patients are transferred from the operation room, usually after they have been stabilized and are judged to be well enough. Although it can be expected that the average nursing time levels out during longer admissions, the difference will be visible provided that the length of stay of the elective surgical patients is normally short.

Patients with a higher mortality risk according to APACHE IV require more nursing time per shift compared to patients with a lower mortality risk. These patients need more and more complex nursing interventions which take more nursing time. There are only a few articles available on the differences in needed nursing time but they mostly are in line with our results: patients who were most severely ill and consequently died during ICU admission had a significantly higher NAS compared to patients who survived [16]. The same study did not find a significant difference in NAS between age groups and medical- and surgical admitted patients. However they divided NAS only in two groups: low values (<66.4%) and high values (>66.4%) which might explain the deviation from our results [16]. On the other hand the study of Armstrong [17] found no differences in NAS for patients on an ICU and on an intermediate care unit, which seems to contrast with our findings, although they did not provide information on the APACHE scores of the patients on the intermediate care.

Carrara et al. (2016) observed that higher numbers of nursing staff are needed since obese patients require more time for instance for mobilization, hygiene procedures, and performing of change positions [18]. It is remarkable that we found only small (significantly, but not clinically relevant) differences between patients in different BMI subgroups, which is not in line with the study of Carrera et al. A possible explanation could be the timing of the measurements. Where mobilization can take a lot of workload in obese patients, most of these procedures would take place while the patients are recovering.

The differences in needed nursing time between the day/evening shift and the night shift, although small, could be explained by the fact that nursing activities such as mobilization, hygiene procedures, and family conversations are not performed during the night shift, which therefore decreases nursing workload. Furthermore, there is more nursing staff available during the day- and evening shifts which results in a lower needed nursing time to be delivered per nurse.

We speculated that nursing workload on ICUs of different size might differ because of contextual aspects. For instance bigger ICUs could admit patients with more complex problems than smaller ICUs. However, we were not able to find any difference. If there are differences because of the type of patients it probably takes more detailed measurements to find them.

In a previous study we performed time-and-motion techniques to measure the exact needed nursing time per shift per patient among 287 patients. We applied the analyses of this study to that small dataset with observed needed nursing time per shift from the time-and-motion techniques so we can compare the results of the NOW and NAS time with this exact needed nursing time data which we consider as golden standard. In general, the nursing time that the NOW model predicts is more consistent with the observed needed nursing time per shift compared to the NAS time (Appendix K and L). This finding is in line with our former study [14], in which we showed that the NAS overestimated the observed nursing workload. Therefore, we believe that the results of the NOW time are more accurate and should preferably be used for nursing capacity planning.

To our knowledge this is the first study which identifies the variability in needed nursing time across different patient- and contextual subgroups. Another strength of our study is the large dataset of 159.733 nursing shifts in different types of ICUs. As the nine ICUs are a representative sample of all Dutch ICUs we believe that the results are generalizable for Dutch ICUs but further research in a larger group of ICUs is needed to prove this. Further research is also needed to investigate how generalizable these results are to other countries, especially non-Western countries with traditionally other patient to nurse ratios.

Our study has several limitations. We only measured average nursing time per shift for all patient groups. It is very likely that nursing time varies in patients according to the period of their illness. A medical patient could take more time immediately after admission than after several days and an obese or elder patient could take considerable more time for mobilization while revalidating. As we showed in earlier work that the amount of NAS points is associated with mortality, this aspect could also be of considerable importance. Another limitation of our study is that we only measured patient related nursing activities. According to the literature, nurses spend approximately 3% to 6% of nursing time on non-nursing activities in ICUs [7,19-21]. This time is dedicated to e.g. participating in an emergency team or supervising of a student. Therefore, this time is also important to take into account when considering the entire nursing workload.

Conclusion

Needed nursing time per shift varies between different patient- and contextual groups but most differences were clinically irrelevant. In general, patients who are more severely ill (medical admitted patients or patients with a high mortality risk) need more time from a nurse in comparison with less severely ill patients (elective admitted patients or patients with a low mortality risk). Similarly, patients in day and evening shifts need more nursing time compared to patients in night shifts. It is therefore not recommended to plan nursing capacity on a fixed ratio of patients per nurse but to include patient characteristics and contextual factors while planning nursing capacity and during assignment of patients to nurses per shift.

References

1. Duffield, C., Diers, D., O'Brien-Pallas, et al. (2011). Nursing staffing, nursing workload, the work environment and patient outcomes. Applied nursing research, 24(4), 244-255.
2. Driscoll, A., Grant, M. J., Carroll, D., et al. (2018). The effect of nurse-to-patient ratios on nurse-sensitive patient outcomes in acute specialist units: a systematic review and meta-analysis. European Journal of Cardiovascular Nursing, 17(1), 6-22.
3. Margadant, C., Wortel, S., Hoogendoorn, M., et al. (2019). The Nursing Activities Score per Nurse Ratio Is Associated With In-Hospital Mortality, Whereas the Patients Per Nurse Ratio Is Not. Critical Care Medicine.
4. Hugonnet, S., Chevrolet, J. C., & Pittet, D. (2007). The effect of workload on infection risk in critically ill patients. Critical care medicine, 35(1), 76-81.
5. Carayon, P., & Alvarado, C. J. (2007). Workload and patient safety among critical care nurses. Critical care nursing clinics of North America, 19(2), 121-129.
6. Brennan, C. W., Krumlauf, M., Feigenbaum, K., et al. (2018). Patient Acuity Related to Clinical Research: Concept Clarification and Literature Review. Western journal of nursing research, 0193945918804545.
7. Cartmill, L., Comans, T. A., Clark, M. J., et al. (2012). Using staffing ratios for workforce planning: evidence on nine allied health professions. Human resources for Health, 10(1), 2.
8. Dutch National Intensive Care Evaluation (NICE) foundation [Internet]. Available at: http://www.stichting-nice.nl
9. Margadant, C.C., Brinkman, S., Hoogendoorn et al. (2020). Validation of the Nursing Acitivities Score (NAS) using time-and-motion techniques in Dutch intensive Care Units. (under submission).
10. Margadant, C., de Keizer, N., Hoogendoorn, M., Bosman, R., Spijkstra, J., Brinkman, S. (2020). Nurse Operation Workload (NOW): a new nursing workload model for Intensive Care Units based on time measurements. (under submission).
11. Cullen, D. J., Civetta, J. M., Briggs, B. A., & Ferrara, L. C. (1974). Therapeutic intervention scoring system: a method for quantitative comparison of patient care. Critical care medicine, 2(2), 57-60.
12. Miranda, D. R., Nap, R., de Rijk, A., et al.: Nursing activities score. Critical care medicine 2003; 31(2), 374-382.
13. Interview Reis Miranda, jaarboek NICE. Available at: https://www.stichting-nice.nl/focusIC.jsp. Accessed September 1, 2019.
14. Quality standard Dutch ICUs [Internet]. Available at: https://www.zorginzicht.nl//
15. R Core Team (2015). R: A language and Environment for Statistical Computing [computer program]. Vienna, Austria: R foundation. Available at: http://www.r-project.org/.

16. Padilha, K. G., de Sousa, R. M. C., Queijo, A. F., Mendes, A. M., & Miranda, D. R. (2008). Nursing Activities Score in the intensive care unit: analysis of the related factors. Intensive and Critical Care Nursing, 24(3), 197-204.
17. Armstrong, E., de Waard, M. C., de Grooth, H. J. S., Heymans, M. W., Miranda, D. R., Girbes, A. R., & Spijkstra, J. J. (2015). Using nursing activities score to assess nursing workload on a medium care unit. Anesthesia & Analgesia, 121(5), 1274-1280.
18. Carrara, F. S. A., Zanei, S. S. V., Cremasco, M. F., & Whitaker, I. Y. (2016). Outcomes and nursing workload related to obese patients in the intensive care unit. Intensive and Critical Care Nursing, 35, 45-51.
19. Bosman, R., Rood, E., Oudemans-van Straaten, H., et al. (2003). Intensive care information system reduces documentation time of the nurses after cardiothoracic surgery. Intensive care medicine, 29(1), 83-90.
20. Kaya, H., Kaya, N., Turan, Y., et al. (2011). Nursing activities in intensive care units in Turkey. International journal of nursing practice, 17(3), 304-314.
21. Norrie, P. (1997). Nurses' time management in intensive care. Nursing in critical care, 2(3), 121-125.

CHAPTER 7

GENERAL DISCUSSION

7.1 Introduction

The quantification of nursing workload is essential to efficiently distribute the workload among the available ICU nurses, to ensure the health of the nursing staff, to optimize the quality of care, and to limit personnel costs in the ICU. The aim of this thesis was to enhance our knowledge of the quantification of nursing workload, for which a profound understanding of the concept and its potential for improving the quality of care in ICUs was required.

To realize the objective of this thesis, four research questions were raised and addressed:

1. Which nursing workload models are used in ICUs and how valid are they to estimate nursing workload?
2. Is nursing workload associated with in-hospital mortality in the Dutch ICU population?
3. How could nursing workload be quantified in the most accurate way?
4. How important are patient characteristics and contextual factors for the amount of nursing workload generated?

Different workload scoring systems have been developed to quantify the need for nursing time and to serve as input for the daily planning of nursing staff in ICUs. We compared the validity and reliability of these different nursing workload models and validated the best performing and most widely used nursing workload model, the Nursing Activities Score (NAS). With the NAS, we assessed whether in-hospital mortality was influenced by the number of patients per ICU nurse or the amount of workload per nurse. Next, we developed a new model and checked whether case-mix was of influence in this model. Finally, we gave an overview of differences of nursing workload in different patient- and ICU subgroups. This chapter gives an overview of the overall discussion of the different studies as performed in this thesis.

7.2 Reliability and validity of currently used nursing workload models

The first research question was "Which nursing workload models are used in ICUs and how valid are they to estimate nursing workload?".

This research question was addressed in Chapter 2 and Chapter 3, in which we identified and compared nursing workload models in ICUs.

In Chapter 2 we identified and compared the performance of published nursing workload models in ICUs. A systematic literature review was carried out to evaluate the validity and reliability of the nursing workload models. In our literature review, we included 71 articles. We identified 27 different scoring systems with a translation into needed nursing time. We excluded (n=57) articles which were written in other languages than English. Among them, a substantial number of Spanish or Portuguese articles revolved around the NAS. Despite this seemingly high number of excluded NAS studies, the NAS is still well represented in the results of our review (n=12). Therefore, we believe that the most important systems are represented in the English publications and hence in our review.

The review showed a variety of models which intended to measure nursing workload. These models were not or not effectively validated for their purpose. This lack of thorough validation makes most of the scoring systems less reliable for making management decisions and capacity planning. We concluded that the NAS performed best since the authors of the NAS validated the model through Multi Moment Recordings. However, the intra- and interrater reliability of the NAS showed need for improvements. The reliability of the NAS could be improved among ICU nurses with more education or training in the use of the NAS given the psychometric properties of the questions. The answers on subjective questions, for example the time spent on a particular nursing activity, are more influenced by the external factors when the nurse lacks knowledge on the definitions of those questions. For instance, when a nurse is helping the patient with hygiene procedures for more than two hours during her shift, she might think she could score this time in the bedside category of the NAS instead of the hygiene category, since she was present at the bedside of the patient during this activity. In this example, she fills in the NAS incorrectly which results in an underestimation of the actual workload because less NAS points are appointed to bedside time compared to hygienic procedures.

7.3 Validation of the NAS

To evaluate the validity of the NAS, we compared the NAS with the optimal measure for time measurements, time-and-motion techniques, in Chapter 3. In time-and-motion techniques an observer constantly measures how much time is devoted to each activity. Each activity will be recorded together with the time spend on it. This technique is labor intensive but yields precise results for measuring time of nursing activities [1]. We refer to the results out of the time-and-motion techniques as 'observed time'. Since the NAS is working with points instead of time, we converted NAS points per activity to time and compared this with the observed time per activity. A correlation of 35% (R^2) was found whereby the total converted NAS time was higher compared to the total observed time. For the validation at activity level, the NAS overestimated most nursing activities when compared to the observed times. Despite the low performance of the NAS, this model is frequently used in daily practice with the assumption that a nurse can take care for 100 NAS points. In reality, the number of NAS points for which a nurse can take care for will be higher since the NAS overestimates nursing workload. However, in the assumption on how much NAS points a nurse can work on per shift, non-patient related tasks are not included. These tasks vary in type and duration between hospitals and it is therefore hard to define an absolute number of NAS points per shift a nurse should take care of.

The NAS has been developed more than 15 years ago and significantly overestimates the needed nursing time for ICU patients in the current daily ICU practice. In conclusion of this study, we advise to revise the time weights assigned to nursing activities in the NAS to get more accurate insights in the true nursing workload and to enable the use of this information for capacity planning. In this study we were not able to measure all activities occurring in the NAS. However, it is not likely that our results have been influenced by this, since these non-frequently measured activities also rarely occur in daily ICU care.

7.4 Nursing workload is associated with in-hospital mortality

The second research question was: *"Is nursing workload associated with in-hospital mortality in the Dutch ICU population?",* which we addressed in Chapter 4.

A very actual question is whether there is an association between workload and outcome of patients. Several studies have assessed this question with conflicting results. These studies used the number of patients for which a nurse cared for per shift: the Patients per Nurse Ratio (PNR). However this starting point does not take into account that one patient may need far more nursing time than another, which makes a Patients per Nurse Ratio (PNR) less useful. In contrast with this, NAS measures the nursing time needed for each individual patient to quantify the nursing workload. In our study we compared the PNR with the NAS to Nurse Ratio (NNR) on their association with patient outcome. We showed that the PNR does not have an association with in-hospital mortality, but the NNR does. This study revealed that it is more important to focus on the nursing workload the patients generate, than on the number of patients the nurse has to take care of in the ICU. Furthermore, the nursing workload should not exceed 61 NAS points per nurse, since this is associated with an increase in in-hospital mortality. This information is important in the context of quality of care. ICU managers could use these results to benchmark their current NNRs and to consider whether the nurses on their ICUs have an acceptable workload in the context of increased risk of in-hospital mortality for their patients. Consequently, ICU nurses could be deployed more efficiently while supporting higher quality of care in terms of survival among their patients.

The results of this study are of great importance since current guidelines are based on the PNR. This could negatively influence daily ICU practice, especially in hospitals with many patients with a high nursing workload. We therefore advise to revise the guidelines and to use the NNR instead of the PNR.

As concluded in Chapter 3, the NAS overestimates nursing workload. We therefore suggested that an ICU nurse should be able to carry out more workload than the 100 NAS points per shift as stated in the original study. However, our study showed that the workload of a nurse higher than 61 NAS points per shift could have negative consequences for the in-hospital mortality of ICU patients. The workload per nurse is clearly a very important tool to provide good quality of care, but we think it is still prematurely to define an exact and safe limit. More research is needed before one can state an optimum nursing workload per ICU nurse.

7.5 Quantifying nursing workload

The third research question was: *"How could nursing workload be quantified in the most accurate way?"* which we addressed in Chapter 5.

As chapter 2 and 3 showed that existing nursing workload models do not perform perfectly, we wanted to develop a better model. For this study we developed three different nursing workload models varying in complexity. These models were based on the time spent per nursing activity per patient. Model one and two give an estimate of the nursing workload per patient by adding up the coefficients per occurred nursing activity. For model two this was extended with case-mix correction. The third more complex model was developed with separate "sub" models for each nursing activity to estimate the time spent on that particular activity based on the case-mix of the patient. Subsequently, the third model estimated the nursing workload per patient by adding up all separate "sub" model estimations of time spent per nursing activity. We tested the performance of all models by several tests. Model one explains 79% (R^2) of patient related nursing workload, model two 81% (R^2), and model three 41% (R^2). Even though model two has a slightly better performance than model one, we recommend to use the simpler model one, which we called Nurse Operation Workload (NOW). The NOW can easily be used to assess the workload per patient by adding up the coefficients per nursing activity that occurred during an 8-hour shift. The NOW accurately estimates patient related nursing workload and therewith, outperforms existing models while it's simpler to use, since less variables need to be registered in comparison to existing models. Moreover, the nursing workload calculated with the NOW is given in minutes instead of points, like the NAS does, and this makes the outcome easier to interpret. The patient related nursing workload measured by the NOW gives insight in the actual nursing time needed per patient and could therefore be used to evaluate whether nurses are deployed in the most efficient way. The NOW could be used in daily ICU practice since it looks like the most accurate way to quantify nursing workload. However, we advise more validation in other settings before it can be implemented in the Dutch guideline for ICUs.

7.6 Variation in nursing workload

The final research question addressed in this thesis was: '*How important are patient characteristics and contextual factors for the amount of nursing workload generated?'.* This research question was addressed in Chapter 6, in which we determined the differences in nursing workload for several patient characteristics and contextual factors.

When scheduling ICU care, differences in nursing workload need to be taken into account for an optimal use of available nursing staff. Our study showed that patients who are more severely ill (medical patients or patients with a high mortality risk) need more nursing time in comparison with less severely ill patients (patients admitted after elective surgery or patients with a low mortality risk). Furthermore, in night shift the patients require significantly less care from a nurse than patients in day- or evening shifts. A strength of this study is the large dataset and the use of different type of ICUs, which makes the results generalizable to other Dutch ICUs. A limitation of our study is that we measured the average nursing time per shift over the complete IC stay for all types of patients groups. It is likely that nursing time varies in patients according to the period of their illness. It is conceivable that patients take significantly more nursing time in the first shift(s) than once they are stabilized.

In Chapter 5, we focused on the performed nursing activities to examine whether case-mix corrections are required to revise the weight or duration of each executed activity by nursing staff. For example, do hygienic procedures take more time for patients with a high BMI compared to patients with a lower BMI? Overall, we found no evidence that such case-mix corrections are required. In Chapter 6, we set out to study whether specific patient groups need more nursing time compared to other groups. In some instances, we found that this is indeed the case. For example, we found that obese patients need more nursing time in total per shift compared to patients with normal BMI. This is likely because an obese patient requires more and/or additional time consuming nursing activities, such as the positioning in a prone position.

We believe that the results of the study of Chapter 6 could be used for capacity decisions on ICU nurses in Western countries with a similar organization as Dutch ICUs. In another context the number of nurses that should be allocated to different types of patients or shifts can be very different [2,3]. This chapter contributes to nursing workforce knowledge and therefore could help to allocate nurses in the most efficient way.

7.7 Strengths and limitations

All scientific research is subject to various strengths and limitations, so is ours. In the performed studies for this thesis we addressed the following limitations.

7.7.1 Limitations

Data used during the studies

A limitation of our study was the large proportion of time-and-motion data derived from one hospital, which could have influenced the results of the NAS validation and model development. We were not able to perform a sensitivity analysis without this hospital since the remaining data had a lack of power to run the models. Since we were measuring the duration of nursing activities with the time-and-motion techniques, we don’t think this has substantially biased our results.

When we compared the patients admitted to the nine ICUs included in our measurements with all Dutch ICU patients in the same period, it turned out that the measured patients had a higher length of stay and were more severe ill. This was probably caused by our selection mechanism. In order to measure as many nursing activities as possible we have chosen to randomly select nurses who took care of patients that were expected to stay during the entire shift and not discharged. It is likely that these patients were sicker compared to patients that remained under care for a shorter time. However, we don’t think this has biased our results since our analysis took case-mix into account. If case-mix was of any influence, it is likely that it was noticed in these more severe ill patients, which was not the case.

NAS validation and model development

In the performed studies there were non-occurring or low frequent measurements of some nursing activities. For the development of a nursing workload model, an expert panel (of three intensivists and two researchers with a background in nursing) carefully evaluated these activities and determined whether some of these activities are important to include in the models. They determined that three out of 18 of these activities rarely occur in daily ICU practice, but when they occur, they will generate a high nursing workload. Therefore, the experts have estimated times for these activities based on their experience in ICU care. The other 15 non-occurring or low frequent measured nursing activities, which according to the experts do not generate a lot of workload, were divided over new aggregated variables of which the experts estimated that the time and type of the activity is similar, so they will be allocated to nursing time when they occur in the future. The absence of these non- or low frequent occurred nursing activities made the validation of the NAS incomplete. To check whether it was a coincidence or selection bias that these activities did not occur during our measurements, we assessed their occurrence in the nursing capacity module of the NICE registry containing information on all nursing shifts of 15 Dutch hospitals over a three-year time period. The occurrence of the activities during our measurements were comparable with the occurrence in the NICE database, e.g. the nursing proceeding ‘induced hypothermia’ occurred in 0.5% of the patients in our study and in 0.6% of all patients in the NICE registry. As these nursing activities rarely occur in daily ICU practice it is not likely that our NAS validation results have been affected by this situation.

Since we evaluated the performance of the NAS and the developed patient related nursing workload models, among which the NOW, on the same data as from which the models were developed, results should be interpreted with some caution. However, for the performance tests of the developed models, we used bootstrapping with 1000 samples to make the results more generalizable [4]. Considering that we compared the performance of the NAS with the performance of the NOW model on the same data, we believe this did not influence the conclusion of our comparisons. However, it would still be important to validate the NOW model in a new external setting, in order to optimally validate the NOW model.

7.7.2 Strengths

Use of time-and-motion techniques

An important strength of our studies is that they are based on the time-and-motion technique, which is considered as the optimum measure for time measurement [5]. Measurements for nursing activities with the use of time-and-motion techniques, are more accurate in comparison with the work-sampling approach as used for the development of the NAS [6]. In the context of validation of the NAS and the development of a patient related nursing workload model, this technique has not been used before. Using time-and-motion techniques is labor intensive and very time consuming. However, due to the commitment of several nursing students and scientists in measuring nursing activities in ICUs, we achieved a large sample size of 287 unique patients in which we measured 46.319 nursing activities.

Case-mix correction

Since we were able to use the large dataset of the NICE registry [7], we were able to study the contribution of case-mix adjustment in the development of nursing workload models and while studying the association between nursing workload and in-hospital mortality. For the development of a new nursing workload model this correction enabled us to find out whether case-mix correction per nursing activity led to more accurate predictions for specific types of patients, which was not the case. When comparing nursing workload with in-hospital mortality, these case-mix adjustments helped to accurately determine the number of NAS points a nurse can maximally take care for. These studies are the first in their subject in which case-mix adjustment was taken into account.

7.7 Implications and future work

The currently most used nursing workload models, the TISS and NAS, have both limitations and are somewhat outdated. The newest nursing workload model of these two, the NAS, has been shown to overestimate the workload. Furthermore, the recommendation of 100 NAS points per nurse may lead to an increase in in-hospital mortality. For this reason we recommend to use our developed NOW model which outperforms earlier developed models on nursing workload, combines the NAS and TISS-28 items, and decreases registration burden. Although our internal validation showed good performance, it would be important to validate the model in an external setting to get insight in its reliability. The model is generalizable to other Dutch ICUs and probably to ICUs in countries where ICU care is organized in a similar way. However, more research is needed to validate the model in other settings.

We believe the NOW model is a good basis to study more specifically how high the workload is in different patient- and contextual subgroups. Further research on the association between NOW and patient outcomes is needed. With a nursing workload instrument such as NOW it would also be possible to determine how large the mean nursing capacity should be over a year, comparing different ICU wards and different categories of patients. The information of this instrument could be used to evaluate capacity numbers afterwards. Therewith, a new national benchmark could be developed on the size of nursing capacity in the ICU. This benchmark could be processed in national guidelines and replace the currently used Patient to Nurse Ratios (PNR) by NOW to nurse ratios.

The PNR could lead to a negative effect of daily ICU practice, especially for hospitals with a lot of patients with a high nursing workload. For these ICUs the use of the current guidelines, which are intended for all ICUs, could lead to a continuous high workload for their nurses. Updating guidelines with NNR instead of PNR might have a positive effect on quality of care and hence patient outcomes, but also on the workload balance for nurses themselves.

As the NOW model showed to outperform NAS, the NOW model will be implemented in the capacity module of the NICE registry. In this way, more information of the NOW model is gathered and ICUs which participate in the capacity module of NICE can use the results of the NOW model to benchmark the NNRs and evaluate whether the nurses in their ICUs have a workload which is too high or low, and adjust the workload per nurse if needed.

The NOW model could be used to evaluate the nursing workload for the past shift(s). For future work, it is tempting to speculate on the development of a prediction model which can assess the nursing workload for the next shift(s) in order to plan the nursing capacity more accurately. This would be a challenging task since several patient subgroups have different needs of nursing time and hence you need to model what kind of patients are expected to be admitted to and discharged from the ICU in the coming shifts. More detailed data is needed for this purpose. Furthermore, more information is needed on the trend in nursing workload during the admission of ICU patients. For these questions data from the NICE registry could be useful. The use of more frequent and detailed data on nursing activities and patient flows in the ICU, in combination with machine learning techniques have a large potential to develop capacity planning models.

7.9 Conclusions

We demonstrated that the NAS is among the currently used nursing workload models in practice the most reliable model for the quantification of nursing workload. However, careful validation of the NAS showed that the instrument overestimates the workload necessary for a patient. Furthermore, we demonstrated that nursing workload is associated with in-hospital mortality where the number of patients per nurse is not, which makes an optimal quantification of nursing workload of great importance. Therewith, we demonstrated that the workload should not be higher than a certain threshold to avoid unnecessary negative patient outcomes. We developed an easy-to-use model which accurately measures nursing workload and which outperforms the NAS. This model (the NOW) is the first nursing workload model which is based on time measurements. In the future, it may be used to develop a new national guideline based on the NOW to nurse ratio as well as to develop methods to predict needed nursing capacity for upcoming shifts.

References

1. Finkler, S. A., Knickman, J. R., Hendrickson, G., Lipkin Jr, M., & Thompson, W. G. (1993). A comparison of work-sampling and time-and-motion techniques for studies in health services research. Health services research, 28(5), 577.
2. Panunto, M. R., & Guirardello, E. D. B. (2012). Nursing workload in an intensive care unit of a teaching hospital. Acta Paul Enferm, 25(1), 96-101.
3. Padilha, K. G., de Sousa, R. M. C., Garcia, P. C., Bento, S. T., Finardi, E. M., & Hatarashi, R. H. (2010). Nursing workload and staff allocation in an intensive care unit: a pilot study according to Nursing Activities Score (NAS). Intensive and Critical Care Nursing, 26(2), 108-113.
4. Parke, J., Holford, N. H., & Charles, B. G. (1999). A procedure for generating bootstrap samples for the validation of nonlinear mixed-effects population models. Computer methods and programs in biomedicine, 59(1), 19-29.
5. Burke, T. A., McKee, J. R., Wilson, H. C., Donahue, R. M., Batenhorst, A. S., & Pathak, D. S. (2000). A comparison of time-and-motion and self-reporting methods of work measurement. JONA: The Journal of Nursing Administration, 30(3), 118-125.
6. Steyerberg, E.W. (2009). Evaluation of Performance. Clinical Prediction models: a practical approach to development, validation, and updating. Springer; p. 83-99.
7. Dutch National Intensive Care Evaluation (NICE) foundation (2017). URL https://www.stichting-nice.nl/dd/#561

SUMMARY

An average of 50% of an Intensive Care Units (ICU) budget is spend on nursing staff [1]. From an economical point of view, it is therefore important to understand how many nurses are minimally required to perform care of good quality. Excessive budgetary cuts targeted to reduce the nursing staff are likely to increase the nursing workload This will have a negative impact on patient safety and survival chances [2,3]. Moreover, reducing the number of nursing staff may also negatively impact the nurses well-being if the workload becomes too high [4]. Without an accurate workload measurement method, ICUs are at risk of under- or overstaffing nurses. Additionally, the lack of a validated measurement method hampers the development of a national standard for the allocation of nurses per ICU bed [5]. In order to take the capacity planning of nurses beyond rough estimates and assumptions, an accurate method for measuring nursing workload is essential.

This thesis anticipates on these topics by assessing different nursing workload models and their validity and reliability, with special attention for the NAS (Chapter 2 and 3), by assessing the association between nursing workload and in-hospital mortality (Chapter 4), by developing a new nursing workload model (Chapter 5), and finally, by providing an overview of differences in nursing workload over different patient- and contextual groups (Chapter 6).

Chapter 2 focusses on nursing workload models currently used in ICU care and their ability to quantify needed nursing time. A systematic literature review was carried out to evaluate the content validity, reliability, and validity of the several nursing workload models. MEDLINE, Embase, and Cinahl were searched on English written articles published before March 2018. This resulted in 71 articles identifying 34 different scoring systems of which 27 were included for further analysis, as these described a translation of workload into needed nursing time. We identified the Nursing Activities Score (NAS) as best performing model. The NAS explains 80% of actual nursing time, which was validated by Multi Moment Recordings. The results of the reliability of the NAS, i.e. inter- and/or intra-rater reliability, varied between low to good. The conflicting results of the different studies performed on nursing workload models and lack of thorough validation make the scoring systems less useful for management decisions.

Chapter 3 describes a study in which we performed a validation on the currently most used nursing workload model, the NAS. The aim of this study was to find out whether this model is in need of revisions after more than 10 years since its development. The study used time-and-motion techniques in different hospital types and with different types of nurses. Nurses were followed during their shift and start- and stop times of all nursing activities were logged using an in-house developed web-application. The original NAS points assigned to the nursing activities were converted to time and compared with the observed time per activity during the time-and-motion techniques. A correlation of 59% was found, indicating that the NAS only explains 59% of patient related nursing workload. The NAS overestimated the needed nursing time for most nursing activities. We conclude that the NAS overestimates the total needed nursing time for patients in Dutch ICUs. Therefore, revisions of the time weights assigned to each nursing activity are advised to get better insight in the true nursing workload before using it for capacity planning.

Chapter 4 describes the development of a new nursing workload model based on the nursing activities in the NAS and Therapeutic Intervention Scoring System (TISS), using time techniques. Time spent per nursing activity per patient was used for the development of three nursing workload models, varying in complexity. Model one and two give an estimation of the patient related nursing workload per patient by adding up the coefficients per occurred nursing activity, for model two this was extended with case-mix correction. The third, more complex, model was developed with separate “sub” models for each nursing activity to estimate the time spent on that particular activity based on the case-mix of the patient. Subsequently, the third model estimated the patient related nursing workload per patient by adding up all separate “sub” model estimations of time spent per nursing activity. Performance of the models were assessed in 1000 bootstrap samples with the Pearson correlation coefficient (R^2), Root Mean Squared Prediction Error (RMSPE), Mean Absolute Prediction Error (MAPE), and prediction bias. Model one explains 83% of patient related nursing workload, model two 85%, and model three 64%. We conclude that case-mix correction is not necessary and therefore recommend to use the most simple model one, which we call the Nurse Operation Workload (NOW) model. The NOW can easily be used to assess the patient related nursing workload by adding up the assigned time for each nursing activity that occurred during an 8 hour shift. The patient related nursing workload derived from this model could be used to evaluate whether nurses were deployed efficiently without over demanding.

Chapter 5 evaluates the association between nursing workload and in-hospital mortality. To determine nursing workload we used two measures: the currently most used Patients per Nurse ratios (PNR) which does not take into account that some patients require more nursing time than others, and the NAS per nurse ratio (NNR) which quantifies nursing workload. We investigated the association between the in-hospital mortality and the PNR and NNR respectively in fifteen Dutch ICUs. We tested all associations with and without case-mix correction, which included the patients' comorbidities, age, and admission type. The PNR is not associated with in-hospital mortality. The NNR is significantly associated with the in-hospital mortality, before and after case-mix correction. A NNR higher than 61 NAS points per nurse leads to significantly more in-hospital mortality among ICU patients. This study showed that it is more important to focus on the nursing workload that the patients generate rather than on the number of patients the nurse has to take care for in the ICU.

Chapter 6 focusses on the differences in needed nursing time over several patient characteristics and contextual factors, to determine whether these are relevant in the process of nursing workforce planning. We used data from the nursing capacity module of the National Intensive Care Evaluation (NICE) registry including the occurrence of nursing activities to calculate the NAS and the total nursing time in minutes per shift based on the in our team developed NOW model. We found that patients who were more severely ill (medical admitted patients or patients with a high mortality risk) need more time from a nurse in comparison with less severely ill patients (elective admitted patients or patients with a low mortality risk). Furthermore, patients in the night shift required significantly less care from a nurse compared to patients in the day- or evening shift. This information contributes to nursing workforce knowledge and therefore could help to allocate nurses in the most efficient way. This study confirms again that it is more important to focus on the NOW per Nurse Ratio (NNR) instead of the Patients per Nurse Ratio (PNR) in nursing capacity planning while the NNR is a constant guideline which automatically leads to less nurses in the night shift while the PNR does not correct for this.

The currently most often used nursing workload models, the TISS and NAS, have both limitations and are somewhat outdated. The newest of these two, the NAS, overestimates the workload. Furthermore the recommendation of 100 NAS points per nurse may lead to an increase of in-hospital mortality. For this reason we recommend to use our developed NOW model which outperforms earlier developed models on nursing workload, combines the NAS and TISS-28 items, and decreases registration burden. Although our internal validation showed good performance, it would be important to validate the model in an external setting. We believe the NOW model is a good basis to study more specific how high the workload is in different patient- and contextual subgroups. Further research on the association between NOW and patient outcomes is needed. With a nursing workload instrument such as NOW it would also be possible to determine how large the mean nursing capacity should be over a year, comparing different ICU wards and different categories of patients. The information of this instrument could be used to evaluate capacity numbers afterwards. Therewith, a new national benchmark could be developed on the size of nursing capacity in the ICU. This benchmark could be processed in national guidelines and replace the currently used patient to nurse ratios by NOW to nurse ratios.

References

1. Bittner, M. I., Donnelly, M., Van Zanten, A. R., Andersen, J. S., Guidet, B., Cabello, J. J. T., ... & Schmutz, A. (2013). How is intensive care reimbursed? A review of eight European countries. Annals of intensive care, 3(1), 37.
2. Boev, C., & Xia, Y. (2015). Nurse-physician collaboration and hospital-acquired infections in critical care. Critical Care Nurse, 35(2), 66-72.
3. Penoyer, D. A. (2010). Nurse staffing and patient outcomes in critical care: a concise review. Critical care medicine, 38(7), 1521-1528.
4. Carayon, P., Alvarado, C. et al. Workload and Patient Safety Among Critical Care Nurses. Crit Care Nurs Clin North Am. 2007;19(2):121–9.
5. Adviescommissie Kwaliteit van het Zorginstituut, Kwaliteitsstandaard Organisatie van Intensive Care, 2016.

NEDERLANDSE SAMENVATTING

Gemiddeld wordt 50% van het budget van een Intensive Care (IC) uitgegeven aan verpleegkundige bezetting [1]. Vanuit een economisch standpunt is het daarom belangrijk om te begrijpen hoeveel verpleegkundigen er nodig zijn op de IC om een goede kwaliteit van zorg te waarborgen. Bezuinigingen om de verpleegkundige bezetting te verminderen leiden waarschijnlijk tot een toename van verpleegkundige werklast. Deze hoge werklast kan leiden tot negatieve gevolgen voor de patiëntveiligheid en overlevingskansen van patiënten [2,3]. Daarbij heeft het verminderen van de verpleegkundige bezetting ook negatieve gevolgen voor de gezondheid van verpleegkundigen [4]. Bij gebrek aan een accuraat meetinstrument voor verpleegkundige werklast lopen IC's het risico om te veel of te weinig verpleegkundigen in te zetten. Ook zorgt het gebrek aan een dergelijk meetinstrument ervoor dat er geen nationale normen ontwikkeld kunnen worden voor het aantal patiënten waar een verpleegkundige per dienst zorg voor kan dragen [5]. Om capaciteitsplanning van verpleegkundigen niet alleen te baseren op aannames is het van belang dat er een accurate methode komt om verpleegkundige werklast te kunnen meten.

Dit onderzoek anticipeert op deze onderwerpen door in kaart te brengen welke verpleegkundige werklastmodellen er worden gebruikt op de IC en hoe betrouwbaar en valide deze zijn, specifiek voor de NAS (Hoofdstuk 2 en 3), het bepalen van de associatie tussen verpleegkundige werklast en ziekenhuismortaliteit (Hoofdstuk 4), het ontwikkelen van een nieuw werklast model voor de IC (hoofdstuk 5) en tot slot het bieden van een overzicht in verschillen in verpleegkundige werklast tussen verschillende patiënt- en contextuele subgroepen (Hoofdstuk 6).

Hoofstuk 2 richt zich op modellen die verpleegkundige werklast in kaart brengen op de IC en op het vermogen van deze modellen om verpleegkundige werklast te kwantificeren. Voor dit onderzoek is een systematische literatuurreview uitgevoerd om de inhoudsvaliditeit, betrouwbaarheid en validiteit te evalueren van de verschillende werklastmodellen. Er is gezocht in de databases MEDLINE, Embase en Cinahl op Engels geschreven artikelen gepubliceerd voor Maart 2018. Dit resulteerde in 71 artikelen die 34 verschillende scoringssystemen beschreven. Van deze artikelen werden er 27 geïncludeerd voor verdere analyse, omdat deze een vertaling van werklast naar benodigde verpleegkundige tijd beschreven. De Nursing Activities Score (NAS) presteerde het best doordat het 80% van de verpleegkundige tijd voorspelt, wat is gevalideerd met Multi Moment

Recordings. De resultaten van de betrouwbaarheid van de NAS, zoals de inter- en/of intra-rater betrouwbaarheid, varieerden van laag tot goed. De tegenstrijdige resultaten van de verschillende uitgevoerde onderzoeken naar werklast modellen en het gebrek aan grondige validatie maken de werklast modellen minder geschikt voor managementbeslissingen. Daarbij kan overschatting van tijd door verpleegkundigen leiden tot overbezetting en een toename in de verpleegkundige kosten.

Hoofstuk 3 evalueert de validiteit van de NAS, wat op dit moment het meest gebruikte model is om verpleegkundige werklast te voorspellen. Het doel van deze studie was om te onderzoeken of dit model, na meer dan tien jaar sinds zijn ontwikkeling, toe is aan revisie. In dit onderzoek gebruikten we tijdsmetingen in verschillende ziekenhuistypes en met verschillende typen verpleegkundigen. Random gekozen verpleegkundigen werden gevolgd gedurende hun dienst en start- en stoptijden van alle verpleegkundige activiteiten werden bijgehouden door gebruik van een in-huis ontwikkelde webapplicatie. NAS-punten per verpleegkundige activiteit werden omgezet naar tijd en vergeleken met de geobserveerde tijd per activiteit. Er werd een correlatie gevonden van 59% waarbij de totale omgezette NAS tijd hoger is in vergelijking met de totale geobserveerde tijd. Voor de validatie per activiteit bleek dat de meeste activiteiten werden overschat door de NAS. We concluderen dat de NAS de verpleegkundige werklast op Nederlandse IC's overschat. De NAS is toe aan revisie van de gewichten die per activiteit zijn gegeven, opdat er beter inzicht kan worden verkregen in de daadwerkelijke verpleegkundige werklast. Daarbij kan de gereviseerde NAS bijdragen aan het gebruik van werklastinformatie voor capaciteitsplanning.

Hoofstuk 4 beschrijft de ontwikkeling van een nieuw werklast model gebaseerd op tijdsmetingen van de verpleegkundige activiteiten zoals in de NAS en Therapeutic Intervention Scoring System (TISS). De tijd gespendeerd per verpleegkundige activiteit per patiënt werd gebruikt voor de ontwikkeling van drie verschillende werklast modellen, variërend in complexiteit. Model één en twee geven een inschatting van de verpleegkundige werklast per patiënt door het optellen van de coëfficiënten van de geobserveerde tijd per voorgekomen verpleegkundige activiteit. Voor model twee is dit uitgebreid met correctie voor patiënt karakteristieken. Het derde, meer complexe, model is ontwikkeld met verschillende "sub" modellen voor elke verpleegkundige activiteit, gebaseerd op patiënt karakteristieken om de gespendeerde tijd voor elke activiteit te kunnen voorspellen.

Vervolgens werden alle predicties uit de “sub” modellen bij elkaar opgeteld om de verpleegkundige werklast tijd per patiënt te verkrijgen. De performance van de verschillende modellen werd getest in 1000 bootstrap samples met de Pearsons correlatiecoëfficiënt (R^2), Root Mean Squared Prediction Error (RMSPE), Mean Absolute Prediction Error (MAPE), en predictie bias. Model één verklaart 89% van de verpleegkundige werklast, model twee 85% en model drie 64%. We concludeerden dat patiënt karakteristieken er niet toe doen wanneer verpleegkundige werklast in kaart wordt gebracht en raden daarom aan om model één te gebruiken, welke we het Nurse Operation Workload (NOW) model hebben genoemd. Dit model is gemakkelijk te gebruiken om verpleegkundige werklast in kaart te brengen door de toegewezen tijden per verpleegkundige activiteit die voorkwam, in een achturige dienst, bij elkaar op te tellen. De verpleegkundige werklast die voortkomt uit dit model kan gebruikt worden om te evalueren of verpleegkundigen juist zijn ingezet aan de hand van de verpleegkundige werklast.

Hoofstuk 5 onderzoekt of er een associatie is tussen verpleegkundige werklast en ziekenhuismortaliteit. Om verpleegkundige werklast vast te stellen hebben we gebruik gemaakt van twee maten: de patiënt per verpleegkundige ratio (PNR), welke geen rekening houdt met het feit dat sommige patiënten meer verpleegkundige werklast genereren dan anderen. Daarnaast gebruikten we de NAS per verpleegkundige ratio (NNR), welke verpleegkundige werklast kwantificeert. We onderzochten de associatie tussen de PNR, en respectievelijk de NNR, op 15 Nederlandse IC’s. We testten alle associaties met en zonder correctie voor de volgende patiëntkarakteristieken: comorbiditeiten, leeftijd en opnametype. De PNR bleek geen associatie te hebben met ziekenhuismortaliteit. De NNR bleek wel significant geassocieerd te zijn met ziekenhuismortaliteit, voor en na correctie voor patiëntkarakteristieken. Dit onderzoek toont aan dat het belangrijker is om te focussen op de verpleegkundige werklast die een patiënt genereert dan op het aantal patiënten waarvoor een verpleegkundige moet zorgen op de IC.

Hoofdstuk 6 beschrijft hoe verpleegkundige werklast is verdeeld over verschillende patiëntkarakteristieken en contextuele factoren om vast te kunnen stellen of deze relevant zijn om mee te nemen in het proces van capaciteitsplanning. We gebruikten data van de capaciteitsmodule van de Nationale Intensive Care Evaluatie (NICE) voor het aantal en type verpleegkundige handelingen per patiënt. Daarbij gebruikten we de NAS uit de capaciteitsmodule en berekenden we de totale benodigde verpleegkundige tijd per dienst op basis van het NOW model. Uit de resultaten bleek dat patiënten die ernstiger ziek zijn (medisch opgenomen patiënten of patiënten met een hoge sterftekans) meer tijd nodig hebben van de verpleegkundige in vergelijking met patiënten die minder ernstig ziek zijn (electief opgenomen patiënten of patiënten met een lage sterftekans). Patiënten in de dag- en avonddienst hadden significant meer zorg nodig dan patiënten in de nachtdienst. Ook dit onderzoek bevestigt dat het belangrijker is om te focussen op de NOW per verpleegkundige ratio in plaats van de patiënt per verpleegkundige ratio bij capaciteitsplanning. Dit is van belang omdat de NOW per verpleegkundige ratio een constante richtlijn geeft, welke bijvoorbeeld automatisch leidt tot minder verpleegkundigen in de nachtdienst, terwijl de patiënt per verpleegkundige ratio hier niet voor corrigeert.

De huidige meest gebruikte verpleegkundige werklastmodellen, de TISS en NAS, hebben beiden zwakke punten en zijn aan de verouderde kant. Het nieuwste werklast model, de NAS, overschat de verpleegkundige werklast en de aanbeveling van 100 NAS punten per verpleegkundige, wat leidt tot onnodige ziekenhuissterfte onder IC patiënten. Om deze reden raadden wij specifieke revisies aan voor de NAS. Echter mist de NAS na deze revisies nog steeds enkele medische interventies welke de TISS wel omvat. Daarom raadden we aan om het door ons ontwikkelde NOW-model te gebruiken. Het NOW model presteert beter dan de eerder ontworpen modellen die als doel hebben om verpleegkundige werklast in kaart te brengen. De NOW bestaat uit een combinatie van handelingen uit de NAS en TISS en verlaagt de registratielast onder IC verpleegkundigen. Om dit model optimaal te valideren, is het van belang om een validatie te laten plaatsvinden in een nieuwe setting. Met een volledig gevalideerd werklastmodel, kan er meer specifiek onderzocht worden hoe verschillende patiënt- en contextuele factoren de verpleegkundige werklast beïnvloeden en op wat voor manier verpleegkundige werklast is geassocieerd met patiënten uitkomsten. Met een betrouwbaar instrument wordt het mogelijk om vast te stellen hoe hoog de gemiddelde verpleegkundige capaciteit zal moeten zijn over een jaar, specifiek voor elke ziekenhuisafdeling en voor verschillende categorieën van patiënten. Deze informatie kan naderhand gebruikt worden om de verpleegkundige capaciteit te evalueren. Daarbij kan er een nieuwe benchmark ontwikkeld worden voor de grootte van de verpleegkundige capaciteit op de IC. Deze benchmark kan verwerkt worden in nationale richtlijnen, die op dit moment de patiënt per verpleegkundige ratio gebruiken in plaats van de NOW per verpleegkundige ratio.

Referenties

1. Bittner, M. I., Donnelly, M., Van Zanten, A. R., Andersen, J. S., Guidet, B., Cabello, J. J. T., ... & Schmutz, A. (2013). How is intensive care reimbursed? A review of eight European countries. Annals of intensive care, 3(1), 37.
2. Boev, C., & Xia, Y. (2015). Nurse-physician collaboration and hospital-acquired infections in critical care. Critical Care Nurse, 35(2), 66-72.
3. Penoyer, D. A. (2010). Nurse staffing and patient outcomes in critical care: a concise review. Critical care medicine, 38(7), 1521-1528.
4. Carayon, P., Alvarado, C. et al. Workload and Patient Safety Among Critical Care Nurses. Crit Care Nurs Clin North Am. 2007;19(2):121–9.
5. Adviescommissie Kwaliteit van het Zorginstituut, Kwaliteitsstandaard Organisatie van Intensive Care, 2016.

APPENDICES

Author contributions

Appendix A. Search-rules in the used databases

Database	Search rule
MEDLINE	("Nursing"[Mesh] OR "nursing" [Subheading] OR "Critical Care Nursing"[Mesh] OR "Nursing Stations"[Mesh] OR "Models, Nursing"[Mesh] OR "Students, Nursing"[Mesh] OR "Nursing, Team"[Mesh] OR "Nursing, Practical"[Mesh] OR "Nursing Staff, Hospital"[Mesh] OR "Nursing Staff"[Mesh] OR "Nursing Care"[Mesh] OR "Nursing Assessment"[Mesh] OR "Nurses"[Mesh]) AND ("Critical Care"[Mesh] OR "Intensive Care Units"[Mesh] OR "Step down unit"[tiab] OR "high dependency unit" [tiab] OR "Critically ill patient" [Mesh]) AND (patient classification[tiab] OR "Classification"[Mesh] OR "classification" [Subheading] OR classification systems[tiab] OR quantification [tiab] OR quantificate [tiab] OR nursing score [tiab] OR scoring system [tiab] OR workload [tiab] OR "Personnel Staffing and Scheduling"[Mesh]) NOT (Neonatal OR Burn unit [Mesh])
Cinahl	(AB ("Nursing" OR "Nurse")) AND (AB ((MH "Intensive Care Units+") OR (MH "Critical Care+") OR (MH "Critical Care Nursing+") OR "Intensive Care")) AND (AB ((MH "Workload") OR "workload" OR (MH "Workload Measurement") OR (MH "Nurse-Patient Ratio") OR (MH "Classification+") OR "classification" OR (MH "Classification (Library)") OR (MH "Patient Classification") OR (MH "Nursing Classification+"))) NOT (AB ((MH "Burn Units") OR (MH "Burn Patients") OR "Burn unit" OR (MH "Intensive Care Units, Neonatal") OR (MH "Neonatal Intensive Care Nursing") OR (MH "Intensive Care, Neonatal+") OR "Neonatal"))
Embase	('nursing staff'/exp OR 'nurse'/exp) AND 'intensive care'/exp AND ('workload'/exp OR 'nurse patient ratio'/exp) AND 'classification'/exp NOT ('burn unit'/exp OR 'newborn intensive care'/exp)

Appendix B. Interface of the web application used by observers during measurements.

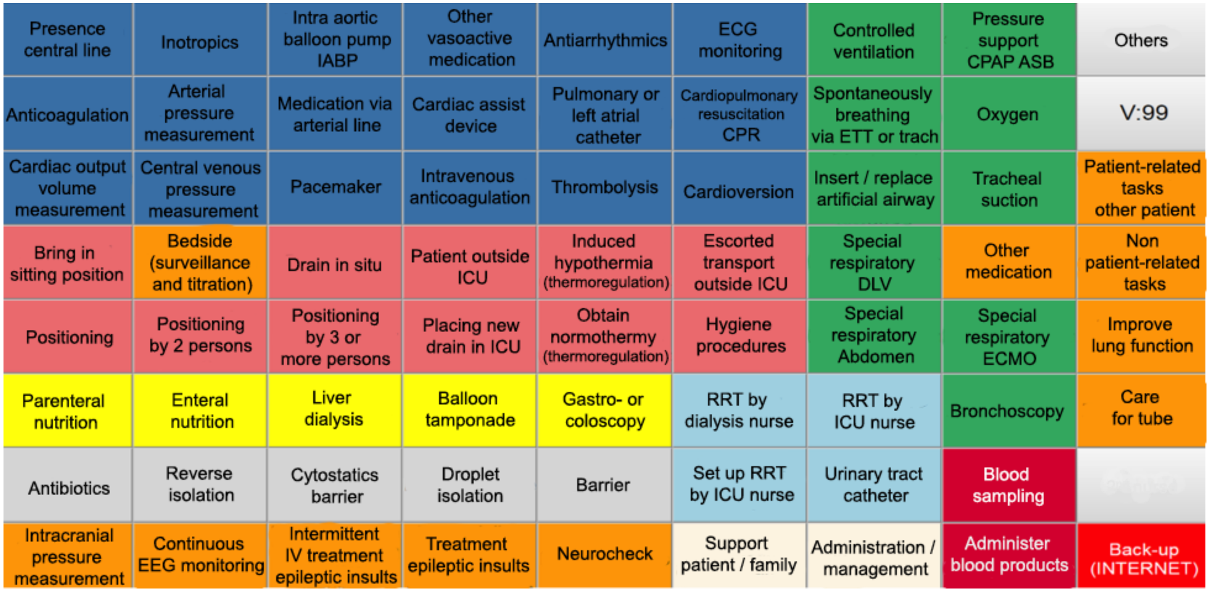


Figure 1. Interface of the web application used by observers during measurements*.* For each time a nurse started a task, the observer pressed the corresponding button for recording the start time. When a button is switched on, the color changes, which indicates the time of the specific task is recorded. When the nurse ends the activity, the button is switched off and records the time of ending.

Appendix C. Description of main variables used in Chapter 3.

Variable	**Explanation**
Observed time	Observed time in minutes, derived from the measurements
Converted NAS time	Observed times per activity are used to match the correct number of NAS points. A nurse is productive in 80% of the shift and 1 % of nursing care during an 8-hour shift corresponds to 1 NAS point which gives the following equation: (8 hours * 60mins)/100)*0.8 = 3.84. Subsequently, the NAS points are converted to time by multiplying the number of NAS points with 3.84 (1 NAS point = 3.84 minutes nursing time).

Appendix D. Nursing activities according to the Nursing Activities Score (Miranda et al., 2003).

	Basic activities	Score
1	Monitoring and titration	
1a	Hourly vital signs, regular registration and calculation of fluid balance	4.5
1b	Present at bedside *and* continuous observation *or* active for 2 hrs or more in any shift, for reasons of safety, severity, or therapy such as noninvasive mechanical ventilation, weaning procedures, restlessness, mental disorientation, prone position, donation procedures, preparation and administration of fluids or medication, assisting specific procedures	12.1
1c	Present at bedside *and* active for 4 hrs or more in any shift for reasons of safety, severity, or therapy such as those examples above (1b)	19.6
2	Laboratory, biochemical and microbiological investigations	4.3
3	Medication, vasoactive drugs excluded	5.6
4	Hygiene procedures	
4a	Performing hygiene procedures such as dressing of wounds and intravascular catheters, changing linen, washing patient, incontinence, vomiting, burns, leaking wounds, complex surgical dressing with irrigation, and special procedures (e.g. barrier nursing, cross-infection related, room cleaning following infections, staff hygiene)	4.1
4b	The performance of hygiene procedures took >2 hrs in any shift	16.5
4c	The performance of hygiene procedures took >4 hrs in any shift	20.0
5	Care of drains, all (except gastric tube)	1.8
6	Mobilization and positioning, including procedures such as: turning the patient; mobilization of the patient; moving from bed to chair; team lifting (e.g. immobile patient, traction, prone position)	
6a	Performing procedure(s) up to three times per 24 hrs	5.5
6b	Performing procedure(s) more frequently than 3 times per 24 hrs, or with two nurses, any frequency	12.4
6c	Performing procedure with three or more nurses, any frequency	17.0
7	Support and care of relatives and patient, including procedures such as telephone calls, interviews, counseling; often, the support and care of either relatives or patient allow staff to continue with other nursing activities (e.g., communication with patients during hygiene procedures, communication with relatives while present at bedside, and observing patient)	
7a	Support and care of either relatives or patient requiring *full dedication* for about 1 hr in any shift such as to explain clinical condition, dealing with pain and distress, difficult family circumstances	4.0
7b	Support and care of either relatives or patient requiring *full dedication* for 3 hrs or more in any shift such as death, demanding circumstances (e.g., large number of relatives, language problems, hostile relatives)	32.0
8	Administrative and managerial tasks	
8a	Performing routine tasks such as processing of clinical data, ordering examinations, professional exchange of information (e.g., ward rounds)	4.2
8b	Performing administrative and managerial tasks requiring *full dedication* for about 2 hrs in any shift such as research activities, protocols in use, admission and discharge procedures	23.2
8c	Performing administrative and managerial tasks requiring *full dedication* for about 4 hrs or more of the time in any shift such as death and organ donation procedures, coordination with other disciplines	30.0
Ventilatory support		
9	Respiratory support: any form of mechanical ventilation/assisted ventilation with or without positive end-expiratory pressure, with or without muscle relaxants, spontaneous breathing with or without positive end-expiratory pressure with or without endotracheal tube supplementary oxygen by any method	1.4
10	Care of artificial airways: endotracheal tube or tracheostomy cannula	1.8
11	Treatment for improving lung function: thorax physiotherapy, incentive spirometry, inhalation therapy, intratracheal suctioning	4.4
Cardiovascular support		
12	Vasoactive medication, disregard type and dose	1.2
13	Intravenous replacement of large fluid losses. Fluid administration >3 L/m^2/day, irrespective of type of fluid administered	2.5
14	Left atrium monitoring: pulmonary artery catheter with or without cardiac output measurement	1.7
15	Cardiopulmonary resuscitation after arrest, in the past period of 24 hrs (single precordial thump not included)	7.1
Renal support		
16	Hemofiltration techniques, dialysis techniques	7.7
17	Quantitative urine output measurement (e.g., by indwelling urinary catheter)	7.0
Neurologic support		
18	Measurement of intracranial pressure	1.6
Metabolic support		
19	Treatment of complicated metabolic acidosis/alkalosis	1.3
20	Intravenous hyperalimentation	2.8
21	Enteral feeding through gastric tube or other gastrointestinal route (e.g., jejunostomy)	1.3
Specific interventions		
22	Specific intervention(s) in the intensive care unit: endotracheal intubation, insertion of pacemaker, cardioversion, endoscopies, emergency surgery in the previous 24 hrs, gastric lavage; routine interventions without direct consequences to the clinical condition of the patient, such as: radiographs, echography, electrocardiogram, dressings, or insertion of venous or arterial catheters, are not included	2.8
23	Specific interventions outside the intensive care unit: surgery or diagnostic procedures	1.9

In the items 1, 4, 6, 7, and 8, only one subitem (a, b, or c) can be scored; the weights represent the percentage of time spent by one nurse on the activity mentioned in the item, if performed.

Appendix E. Nursing workload measures and their definitions as used in this study in Chapter 4.

Term	**Definition**
Nursing Activity Score (NAS)	A system specifically developed for the ICU which assigns a score per executed nursing task per patient based on the time spent per nursing task. A high NAS score of a patient indicates a high patient acuity.
Patients per nurse ratio (PNR)	The ratio of the number of ICU patients present at the ICU divided by the number of registered ICU nurses present at the ICU per nursing shift
NAS per nurse ratio (NNR)	The ratio of the sum of the NAS score of all ICU patients present at the ICU divided by the number of registered ICU nurses present at the ICU per nursing shift

Appendix F. Association NNRs and in-hospital mortality (*n* = 29,445).

			[a] Model: NNR		[b] Model: NNR + adjustment
Covariate	**Range**	**OR**	**(95% CI)**	**OR**	**(95% CI)**
NNR day 1	< 39.0	-	(reference)	-	(reference)
	39.0 - < 54.7	1.171	(1.027 – 1.335)	1.133	(0.982- 1.307)
	54.7 - < 70.3	1.251	(1.090 – 1.435)	1.203	(1.036 - 1.397)
	≥ 70.3	1.262	(1.092 – 1.458)	1.256	(1.074 - 1.469)
Mean NNR	< 39.4	-	(reference)	-	(reference)
	39.4 - < 54.8	1.115	(0.975 – 1.276)	1.011	(0.872 – 1.172)
	54.8 - < 69.9	1.312	(1.142 – 1.507)	1.191	(1.024– 1.385)
	≥ 69.9	1.318	(1.138 – 1.525)	1.230	(1.049 – 1.442)

CI Confidence Interval; NNR NAS per Nurse Ratio; OR Odds Ratio
Association models:
[a] Hospital mortality ~ (1|hospital of ICU admission) + NNR
[b] Hospital mortality ~ (1|hospital of ICU admission) + NNR + comorbidities + age + admission type
NNR calculated including all nurses and all type of patients.

Appendix G. Association PNRs and in-hospital mortality (*n* = 29,445)

			[a]Model: PNR		[b]Model: PNR + adjustment
Covariate	**Range**	**OR**	**(95% CI)**	**OR**	**(95% CI)**
PNR day 1	< 0.90	-	(reference)	-	(reference)
	0.90 - < 1.19	1.133	(0.975 – 1.315)	1.164	(0.988 - 1.370)
	1.19 - < 1.74	1.265	(0.982 – 1.491)	1.214	(0.980 - 1.451)
	≥ 1.74	1.219	(0.998 – 1.466)	1.292	(0.999 - 1.579)
Mean PNR	< 0.90	-	(reference)	-	(reference)
	0.90 - < 1.19	1.165	(0.998– 1.360)	1.171	(0.988 - 1.388)
	1.19 - < 1.72	1.172	(0.987 – 1.625)	1.205	(0.989 - 1.511)
	≥ 1.72	1.199	(0.988 – 1.455)	1.228	(0.993 - 1.519)

CI Confidence Interval; PNR Patients per Nurse Ratio; OR Odds Ratio
Association models:
[a] Hospital mortality ~ (1|hospital of ICU admission) + PNR
[b] Hospital mortality ~ (1|hospital of ICU admission) + PNR + comorbidities + age + admission type
PNR calculated including all nurses and all type of patients.

Appendix H. Data dictionary of variables in the model

Model item	Explanation
Central nervous system	
Neurological checks[*]	The frequency of neurological checks (GCS, delirium, sedation scores).
Cardiovascular	
Administer inotropes	Inotropes include (nor-) adrenalin, dopamin, dobutamin, vasopressin, terlipressin, levosimendan and phosphodiesterase inhibitors.
Cardioversion[*]	>=1x cardioversion during the shift.
Administer antiarrhythmics[*]	Including all 4 Vaughan Williams classes of antiarrhythmics.
Administer anticoagulation	Including administering of thrombolysis, therapeutic- and prophylactic anticoagulation.
Medication via arterial line[*]	Including arterial thrombolysis and arterial administering of calcium solution.
Arterial pressure measurement	Invasive arterial pressure measurement.
Central line	Proceedings at the central line.
Central venous pressure measurement	Central venous pressure measurement.
Respiratory	
Oxygen[*]	Additional administering of oxygen.
Spontaneous breathing via ETT or trach	Including Hudson Klep, One Way Valve, HME etc.
Pressure support[*]	Proceedings of breathing equipment by which the patient breaths (partially) on his own.

Controlled ventilation*	Proceedings around controlled form of breathing.
ECMO*	Extra Corporal Mebran Oxygenation.
Tracheal suction	Proceedings to remove sputum out of the lungs.
Tractus digestive	
Parenteral feeding*	Feeding via continuous infusion.
Enteral feeding	Enteral feeding via stomach, duodenum, jejunum or gastrostomy.
Gastro- or colposcopy*	Gastro- or colposcopy.
Renal	
Building up renal replacement therapy*	Number of times per shift.
RRT by an ICU or dialysis nurse	All forms of renal replacement therapy.
Urinary tract catheter	Proceedings at the urinary tract catheter.
Blood	
Administer blood products	Including erythrocytes, plasma, thrombocytes and other blood products.
Blood sampling*	The number of unique blood samplings per shift.
Monitoring, titration, mobilization and positioning	
Bedside	Time of nurse being present around and active at bedside due to safety of the patient, severity of illness or specific nursing proceedings.
Patient in sitting position	Bringing patient in sitting position.
Positioning	Mobilization or positioning

Infectology	
Administer antibiotics*	The number of unique times of antibiotic administering.
Isolation	Including barrier isolation, droplet isolation, cytostatic barrier and reversed isolation.
Hygiene	
Hygiene activities*	Hygienic proceedings like caring of wounds, changing bedsheets, washing a patient etc.
Administration and support	
Care for patient or relatives*	Support or care for patient and family with full commitment.
Administration*	Administration or management tasks which require full dedication of the nurse.
Other	
Other medication (other vasoactive medication and epileptic medication via IV) *	Other vasoactive medication and epileptic medication via IV.
Thermoregulation	Including acquiring/maintaining normothermia and induced hypothermia.
Monitoring (ECG monitoring, heart minute volume measurement and intracranial pressure measurement)	ECG monitoring, heart minute volume measurement and intracranial pressure measurement.
Drain	Proceedings at drain in situ.
Accompanied transport	Accompanied transport outside the ICU by a nurse.

* Present in the final reduced model, after Lasso regression.

Appendix I. Performance measures

Performance measure	Interpretation	Formula	Range
Pearson's R	The covariance of the two variables divided by the product of their standard deviations.	R $= \frac{\sum_i \left[y_i - \frac{1}{n}\sum_i y_i \right]^2 \left[\hat{y}_i - \frac{1}{n}\sum_i \hat{y}_i \right]^2}{\sum_i \left[y_i - \frac{1}{n}\sum_i y_i \right]^2 \sum_i \left[\hat{y}_i - \frac{1}{n}\sum_i \hat{y}_i \right]^2} = \left[\frac{\mathrm{Cov}(Y,\hat{Y})}{\sigma(Y)\cdot\sigma(\hat{Y})} \right]$	0-1
R²	The percentage of explained variance which is explained by a linear model. Higher values signify better prediction.	$\mathrm{R}^2 = \frac{\sum_i \left[y_i - \frac{1}{n}\sum_i y_i \right]^2 \left[\hat{y}_i - \frac{1}{n}\sum_i \hat{y}_i \right]^2}{\sum_i \left[y_i - \frac{1}{n}\sum_i y_i \right]^2 \sum_i \left[\hat{y}_i - \frac{1}{n}\sum_i \hat{y}_i \right]^2} = \left[\frac{\mathrm{Cov}(Y,\hat{Y})}{\sigma(Y)\cdot\sigma(\hat{Y})} \right]^2,$	0-1
Root Mean Squared Prediction Error (RMSPE)	The mean residual, or unexplained, standard error of predictions obtained using a model. Lower values signify better prediction.	$\mathrm{RMSPE} = \sqrt{\frac{1}{n}\sum_i \left(\hat{y}_i - y_i \right)^2}$	Depends on dependent variable.
Mean Absolute Percentage Error (MAPE)	Accuracy measure based on errors. Lower values signify better prediction.	$\mathrm{MAPE} = \frac{1}{n}\sum_i \left| \hat{y}_i - y_i \right|$	Depends on dependent variable.
Prediction bias	Measures the magnitude and direction of the average model, when compared to observed data. Values closer to zero signify better prediction.	$\mathrm{Bias} = \frac{1}{n}\sum_i \hat{y}_i - \frac{1}{n}\sum_i y_i \,.$	Depends on dependent variable.

n = the number of patients in the dataset.
y_i = the observed time for patient *i.*
$\hat{y}_i$ = the predicted time for patient *i.*
Cov((Y,Ŷ), *σ(Y)* and *σ(Ŷ)* are the covariance and standard deviations of the observed and predicted time.

Appendix J. Log transformed coefficients per nursing activity in the final model[1]

Model item	**Coefficients, in minutes**
Intercept	3.84741108
Central nervous system	
Neurological checks, 1 time /shift	0.06571384
Neurological checks, 2-4 times /shift	0.04057786
Neurological checks, >=4 times /shift	0.15510167
Cardiovascular	
Administer inotropes, >1 times/shift	0.15208661
Cardioversion*	1.47712125
Administer antiarrhythmics	0.08660284
Administer anticoagulation	0.02848972
Medication via arterial line	0.05734708
Arterial pressure measurement	0.03556968
Respiratory	
Oxygen	0.13041269
Pressure support	0.11085906
Controlled ventilation	0.18456842
ECMO*	1.77815125
Tractus digestive	
Enteral feeding	0.11654908
Gastro- or coloscopy*	1.77815125
Renal	
Building up renal replacement therapy	0.35300802
RRT by an ICU or dialysis nurse	0.03077794
Urinary tract catheter	0.01884063
Blood	
Administer blood products, 1 or more times/shift	0.13557211
Administer blood products, 3 or more times/shift	0.31034380
Administer blood products, 5 or more times/shift	0.22663366
Blood sampling, 3 times/shift	0.06237777
Blood sampling, 4 times/shift	0.01651692
Blood sampling, 5 or more times/shift	0.06397588

Monitoring, titration, mobilization and positioning	
Bedside, <1 hour	0.22269618
Bedside, >1 hour	0.34226152
Positioning by one nurse	0.07485767
Positioning by two nurses	0.24228694
Positioning by three or more nurses	0.36714461
Infectology	
Administer antibiotics, 1 or 2 times/ shift	0.03642946
Administer antibiotics, 3 or more times/ shift	0.04037633
Isolation (barrier isolation, droplet isolation, cytostatic barrier and reversed isolation)	0.02498903
Hygiene	
Hygiene activities, <1 hour/shift	0.17098585
Hygiene activities, >1 hour/shift	0.56896408
Administration and support	
Care for patient or relatives, <1 hour/shift	0.26892228
Care for patient or relatives, >=1 hour/shift	0.77750255
Administration, <1 hour/shift	-0.42323767
Administration, >2 hours/shift	0.24435897
Administration, >3 hours/shift	0.46908082
Other	
Other medication (other vasoactive medication and epileptic medication via IV)	0.03733068
Thermoregulation	0.09087546
Care for drains	0.01474637

[*] *For these activities an expert opinion is used, since these activities occurred less than 10 times during the measurements, but do sometimes occur in daily practice.*

[1] See Appendix H for the definitions of the items used in the model. These items form together the final model (after Lasso regression). Activities can be add up if they occurred and the sum should be back transformed to obtain the total time per patient.

Appendix K. Patient characteristics of patients in the time-and-motion techniques. N = 345 nursing shifts, 287 unique patients.

Variable	***Categories (median observed nursing time/points per patient, IQR)***			
Total observed time (h)	1.88 [1.37 – 2.81]			
Admission type	**Medical (N=147)**	**Emergency surgical (N=69)**	**Elective surgical (N=110)**	
Observed time (h)[a]	2.18 [1.57 – 3.03]	1.72 [1.25 – 2.79]	1.55 [1.25 – 2.20]	
APACHE IV mortality risk	**Low (<0.30) (N=153)**	**Medium (0.30 – 0.69) (N=96)**	**High (>=0.70) (N=47)**	
Observed time (h)[b]	1.72 [1.36 – 2.65]	2.02 [1.45 – 2.92]	2.43 [1.57 – 3.13]	
BMI group	**Underweight (BMI < 18.5) (N=2)**	**Normal weight (BMI 18.5 – 24.9) (N=125)**	**Overweight (BMI >=25 – 29.9) (N=245)**	**Obesity (BMI >=30) (N=69)**
Observed time (h)	3.20 [3.040 – 3.36]	1.82 [1.35 – 2.21]	1.88 [1.40 – 2.81]	1.99 [1.52 – 2.73]
Age	**<50 years (N=52)**	**50-79 years (N=227)**	**>=80 years (N=49)**	
Observed time (h)	2.12 [1.50 – 2.90]	1.78 [1.40 – 2.68]	1.97 [1.22 – 2.90]	
Patient subgroups	**Sepsis (N=18)**	**CAP (N=10)**	**OHCA (N=30)**	
Observed time (h)	2.88 [2.26 – 3.51]	2.99 [1.91 – 3.54]	2.26 [1.86 – 2.87]	

OHCA: Out of Hospital Cardiac Arrest; CAP: Community Acquired Pneumonia
[*] Significantly different between all categories; [a] Significantly different between medical and emergency surgical patients; medical and elective surgical; [b] Statistically different between low and high APACHE IV mortality risk.

Appendix L. Contextual factors of the patients in the time-and-motion techniques. N = 345 nursing shifts, 287 unique patients.

Variable	***Categories (median total observed nursing time/points per patient, IQR)***		
Shift type	**Day shift (N=128)**	**Evening shift (N=122)**	**Night shift (N=95)**
Observed time (h) *	2.27 [1.53 – 3.20]	1.90 [1.41 – 2.74]	1.55 [1.06 – 2.18]
Number of beds per ICU	**ICUs with less than 12 beds (N=37)**	**ICUs with 12 beds or more (N=308)**	
Observed time (h) *	2.71 [1.65 – 3.34]	1.78 [1.36 – 2.64]	

* Significantly different between all categories

Curriculum Vitae

Charlotte Claire Margadant was born on September 16th 1992 in Apeldoorn, the Netherlands, where she grew up as youngest of two brothers and one sister. In 2009 she finished her higher general secondary education with a special interest in nature and health. In the same year, she started a Bachelor in Nursing at the Saxion Hogeschool Deventer, of which she obtained her degree in 2013. After a premaster and master at the Free University (VU) Charlotte graduated in 2015 from the Master Health Sciences with specialization 'Nutrition and Health'. She was happy to work on her master thesis on 'Lower muscle density is associated with major postoperative complications in older patients after surgery for colorectal cancer' which got published in 2016. After obtaining her masters' degree Charlotte traveled for a longer period and started in 2016 working as a data manager and PhD student at the Dutch National Intensive Care Evaluation (NICE) registry at the Department of Medical Informatics, Amsterdam University Medical Center (AUMC), University of Amsterdam (UVA), the Netherlands. Under supervision of Prof. Dr. Nicolette F. de Keizer and Dr. Sylvia Brinkman she worked on the research described in this thesis.

Charlotte lives in the center of Amsterdam, and enjoys traveling as much as her free time allows, spending time with family, cycling, skating, swimming, surfing, playing the piano, taking care of plants, baking pies, and cooking food for her family and friends.

Portfolio

PhD candidate:	Charlotte C. Margadant
Period:	May 2016 to June 2020
Promotor:	Prof. Dr. Nicolette F. de Keizer
Co-promotores:	Dr. Sylvia Brinkman
	Dr. Jan Jaap Spijkstra

PhD training	Year	ECTS
General courses		
AMC World of Science	2016	0.7
Project Management	2017	0.6
Specific courses		
Practical Biostatistics	2016	1.1
Endnote	2016	0.1
Clinical Data Management	2017	0.9
Computing in R	2017	0.4
Systematic Reviews	2017	
Searching for a systematic review	2017	0.1
Scientific Writing for publication	2017	1.5
Citation Analysis and Impact Factors	2018	0.1

Seminars		
Amsterdam Public Health (APH) research institute annual meeting (Amsterdam, The Netherlands)	2019	0.2
PhD days (Breukelen, The Netherlands)	2016-2020	1

NICE discussion meeting (Nieuwegein, The Netherlands)	2016-2020	1
Oral presentations, (inter)national		
Measuring nursing workload in Dutch ICUs, nursing management team, AUMC, (Amsterdam, The Netherlands)	2018	0.5
Nursing workload module and new study results, NICE Discusson meeting, (Nieuwegein, The Netherlands)	2016-2020	1
Nursing workload and mortality, ISQUA International conference (Cape Town, South Africa)	2019	0.5
PhD days, Department of Medical Informatics AUMC (Breukelen, The Netherlands)	2016-2020	1
Conferences (inter)national		
Annual Audit & Feedback Symposium 2019 (Amsterdam, The Netherlands)	2019	1
International Society for Quality in Health Care, ISQUA (Cape Town, South Africa)	2019	1
Other PhD training		
Attending research meetings, Department KIK AUMC (Amsterdam, The Netherlands)	2016-2020	3
Attending book discussions, Department KIK AUMC (Amsterdam, The Netherlands)	2016-2020	2
Supervising		
9-month master Medical Informatics internship: Validating the Therapeutic Intervention Scoring System (TISS) and Nursing Activity Score (NAS)	2017-2018	2

9-month master Medical Informatics internship: Nursing Workload and Mortality in Dutch ICUs	2018-2019	2
Grants		
Amsterdam Public Health (APH-QoC) travel grant	2018	
Other activities		
Collecting data and performing analysis and interviews at EPIMED and Brazilian ICUs (Rio de Janeiro, Brazil)	2018	4

1 European Credit Transfer System (ECTS) = 28 hours

List of publications

Publications in this thesis

- Hoogendoorn, M. E., **Margadant, C. C.**, Brinkman, S., Haringman, J. J., Spijkstra, J. J., & de Keizer, N. F. (2019). Workload scoring systems in the Intensive Care and their ability to quantify the need for nursing time: a systematic literature review. International journal of nursing studies, 103408.
- **Margadant, C. C.**, Wortel, S. A., Hoogendoorn, M.E., Bosman, R. J., Spijkstra, J. J., Brinkman, S., & de Keizer, N. F. (2020). The Nursing Activities Score Per Nurse Ratio Is Associated With In-Hospital Mortality, Whereas the Patients Per Nurse Ratio Is Not. Critical Care Medicine, 48(1), 3-9.
- **Margadant, C. C.**, Hoogendoorn, M.E., Bosman, R.J., Spijksta, J.J., Brinkman, S., de Keizer, N.F. Validation of the Nursing Activities Score (NAS) using time-and-motion measurements in Dutch Intensive Care Units: an observational study [in press].
- **Margadant, C. C.**, de Keizer, N.F., Hoogendoorn, M.E., Bosman, R.J., Spijkstra, J.J., Brinkman, S. Nurse Operation Workload (NOW): a new nursing workload model for Intensive Care Units based on time measurements [in press].
- **Margadant, C. C.**, Brinkman, S., Hoogendoorn, M.E., Bosman, R.J., Spijkstra, J.J., de Keizer, N.F. Effect of patient characteristics and contextual factors on needed nursing time in Intensive Care Units [in press].

Other publications

- **Margadant, C. C.**, Bruns, E. R. J., Sloothaak, D. A. M., van Duijvendijk, P., van Raamt, A. F., van der Zaag, H. J., ... & van der Zaag, E. S. (2016). Lower muscle density is associated with major postoperative complications in older patients after surgery for colorectal cancer. *European Journal of Surgical Oncology (EJSO)*, *42*(11), 1654-1659.

Author contributions

Chapter 2 - Workload scoring systems in the Intensive Care and their ability to quantify the need for nursing time: A systematic literature review.

Study conception and design: Marga Hoogendoorn, Charlotte Margadant, and Nicolette de Keizer. Acquisition of data: Marga Hoogendoorn and Charlotte Margadant. Drafting of manuscript: Marga Hoogendoorn. Analysis and interpretation: Marga Hoogendoorn and Charlotte Margadant. Critical revision: Charlotte Margadant, Sylvia Brinkman, Jan Jaap Spijkstra, and Nicolette de Keizer.

Chapter 3 - Validation of the Nursing Activities Score (NAS) using time-and-motion measurements in Dutch Intensive Care Units: an observational study.

Study conception and design: Charlotte Margadant, Sylvia Brinkman, and Nicolette de Keizer. Acquisition of data: Charlotte Margadant. Drafting of manuscript: Charlotte Margadant. Analysis and interpretation: Charlotte Margadant and Sylvia Brinkman. Critical revision: Sylvia Brinkman, Marga Hoogendoorn, Rob Bosman, Jan Jaap Spijkstra, and Nicolette de Keizer.

Chapter 4 The Nursing Activities Score per Nurse Ratio is associated with in-hospital mortality, whereas the Patients per Nurse Ratio is not

Study conception and design: Charlotte Margadant, Sylvia Brinkman, Nicolette de Keizer, Jan Jaap Spijkstra, and Rob Bosman. Acquisition of data: Charlotte Margadant. Drafting of manuscript: Charlotte Margadant. Analysis and interpretation: Charlotte Margadant and Sylvia Brinkman. Critical revision: Sylvia Brinkman, Marga Hoogendoorn, Rob Bosman, Jan Jaap Spijkstra, and Nicolette de Keizer.

Chapter 5 Nurse Operation Workload (NOW): a new nursing workload model for Intensive Care Units based on time measurements

Study conception and design: Charlotte Margadant, Sylvia Brinkman, and Nicolette de Keizer. Acquisition of data: Charlotte Margadant. Drafting of manuscript: Charlotte Margadant. Analysis and interpretation: Charlotte Margadant and Sylvia Brinkman. Critical revision: Safira Wortel, Sylvia Brinkman, Marga Hoogendoorn, Rob Bosman, Jan Jaap Spijkstra, and Nicolette de Keizer.

Chapter 6 Effect of patient characteristics and contextual factors on needed nursing time in Intensive Care Units

Study conception and design: Charlotte Margadant and Sylvia Brinkman. Acquisition of data: Charlotte Margadant. Drafting of manuscript: Charlotte Margadant. Analysis and interpretation: Charlotte Margadant. Critical revision: Sylvia Brinkman, Marga Hoogendoorn, Rob Bosman, Jan Jaap Spijkstra, and Nicolette de Keizer.

DANKWOORD

Dit is hem dan! Mijn proefschrift is klaar. Vier jaar aan werk is nu gebundeld in dit boek. Er hebben veel mensen aan de inhoud van dit boek meegewerkt of mij geweldig gesteund en die wil ik graag allemaal bedanken.

Allereerst wil ik mijn promotor en copromotor bedanken. Nicolette, dank voor de vele waardevolle inzichten die mij zo ver hebben doen komen. Sylvia, je grote statistische kennis en waardevolle feedback hebben me enorm geholpen! Ik heb altijd heel fijn met je samengewerkt. Jan Jaap, Marga en Rob, bedankt voor jullie kritische blik op mijn stukken en input als coauteurs. Mijn lieve kamergenootjes, Safira, Marie-José en Juliët, dank voor de gezelligheid. Hugo, met jou was het altijd dolle pret! Ik ben blij dat we nu nog altijd lekker sportief samen wielrennen of zwemmen.

I am really thankful for the lovely people I worked with at Epimed, Rio de Janeiro. Jorge, thank you for showing me Brazilian ICUs and provide me with all the information I needed (even if it was about great Brazilian fruit juices). Victor and Caio, I loved to work with you guys. We made so much fun! And with special thanks to Lunna. You have learned me a lot, from eating brigadeiro's, Casa de Papel insights, and nail colors to the most complex R codes.

Marthe en Tinka, wat hebben we veel meegemaakt samen en verdrietige maar vooral heel veel hele blije momenten beleefd. Wat moet ik gaan afkicken van jullie dagelijks zien!! Ik ga onze dansjes, grapjes en gesprekken zo missen! Ik ben heel (lees: héél) erg blij dat ik jullie heb ontmoet door deze PhD! Lieve Lau, mijn beste vriendin, dankje voor al je liefde. Ik weet dat je er altijd voor alles voor me bent en dat zal ik ook altijd voor jou zijn. Je betekent zoveel voor me! Ebru, dankjewel voor me altijd de slappe lach geven met je leuke grappen en voor alle gezellige dagen waarop we samen zijn en de uren zo snel voorbij vliegen! Paulette, dank voor de gezellige avondjes salsa dansen waarbij ik mijn werk even helemaal vergat! ¡Bailemos para celebrar esto! Marije, Marleen, Ahmed en Lara, dank voor het zijn van zulke lieve en gezellige vrienden!

Lieve schoonfamilie, wat ben ik blij met jullie! Ik voel me thuis wanneer we samen zijn en kom na een avond samen altijd uitgeput en propvol thuis van het vele lachen en het heerlijke Surinaamse eten. Jullie zijn schatten en samen hebben we van de af en toe minder leuke momenten in de afgelopen jaren er gewoon een grap van gemaakt waardoor alles weer luchtig en behapbaar werd. Dank voor al jullie liefde! Ik bof zo erg met jullie! Dankjewel lieve Albert en Sita, voor al jullie goede zorgen. Dankjewel lieve Kavish en Rowena, voor alle gezelligheid altijd. Dankjewel lieve Shalini voor alle grappen, je enthousiasme en lieve steun! Dankjewel lieve Rishaan, Eshan en Ishani voor alle uren samen lekker spelen, lachen en knuffelen.

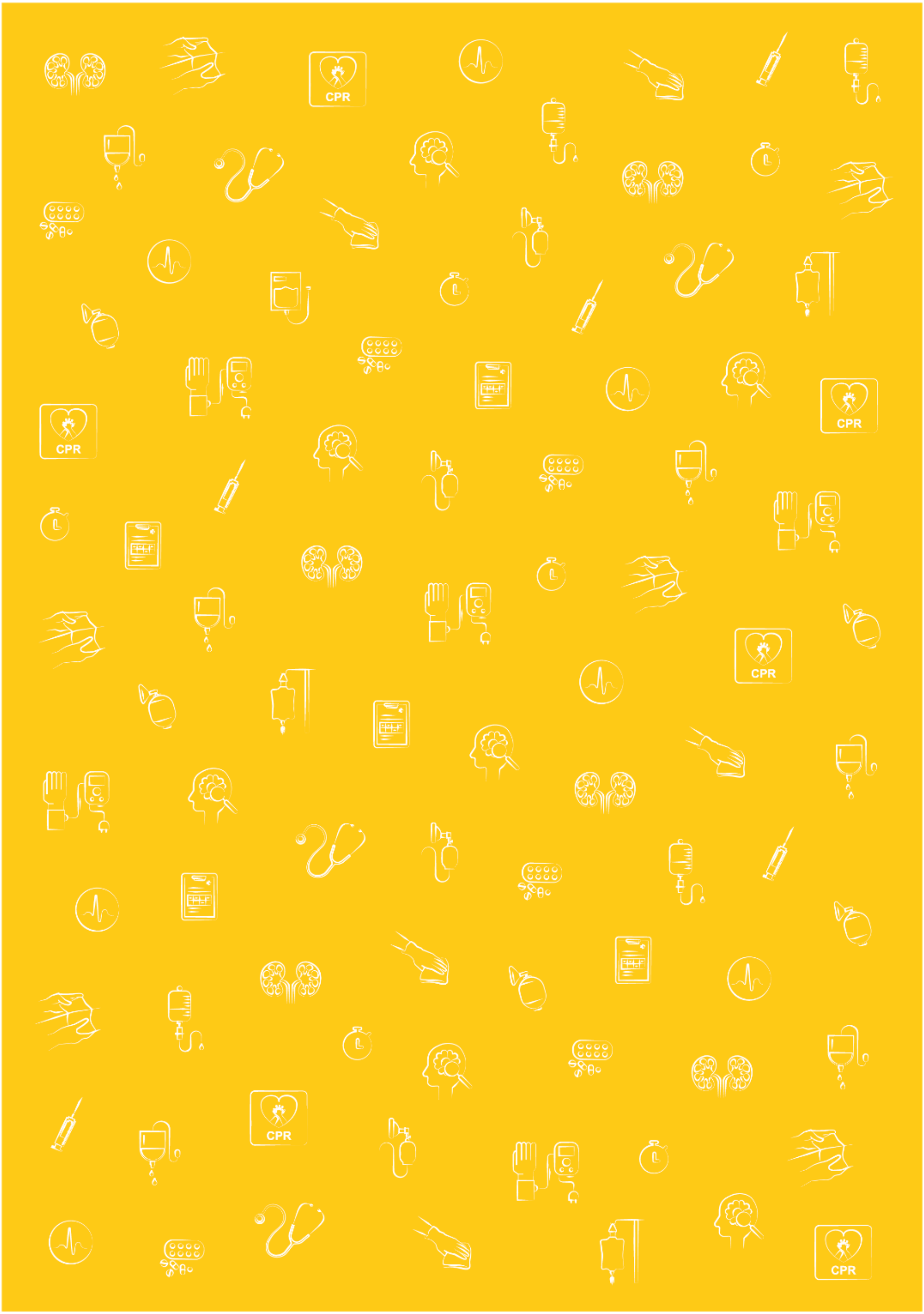
Lieve broers, zussen en inmiddels vele nichtjes en neefjes, Martijn, Monique, Eva, Fleur, Leon, Machteld, Jacobien, Pieterbas, Valentijn, Roderik, Judith en Luna, dank voor het zijn van een hele warme familie. Ik ben altijd jullie 'kleine zusje' geweest maar nu ben ik echt groot hoor! Martijn, je bent een schat van een broer, de meest fanatieke sporter die ik ken en een enorm lieve papa. Monique, je bent mijn zus geworden in al die jaren dat je al bij ons bent en daar ben ik heel blij mee! Ik ben een héle trotse peetmama van jullie lieve meisjes. Eva en Fleur, jullie zijn de grappigste en allerliefste kindjes die ik ken. Ik geniet altijd zo erg van alle tijd die we samen doorbrengen. Juud, dank voor je enthousiasme altijd, het zijn van zo'n geweldige vriendin voor Ro en het zijn van zo'n lieve mama voor Lun! En natuurlijk voor het ontwerpen van de voorkant van dit boek!!! Ro, lieve broer, je bent er altijd voor me en ik lijk zo op je. Je hebt altijd energie voor honderd dingen tegelijk, geniet volop van het leven, hebt zoveel kennis en je bent nu ook nog een geweldige papa. Ik ben zo trots op je. En natuurlijk Luna. Je bent nog zo klein maar je laat iedereen al stralen met je lieve lach!

En nu jullie pap en mam! Per dag zat ik ongeveer twee uur op de fiets om op werk te komen en naar huis te gaan. Elke rit belde ik gezellig met één van jullie en zo hielden we elkaar van alles op de hoogte. Ik stond nooit ergens alleen voor en dat gaf me zoveel kracht! Lieve mama, je hebt altijd benadrukt dat ik alles aankan en in mezelf moet geloven. En hoppa, kijk waar ik nu sta! Je dochter is een doctor! Dank voor alle moed die je me mijn hele leven hebt ingesproken!! Die ongelooflijk positieve en moedgevende instelling ga ik nu doorgeven aan mijn studenten zodat zij ook het allerbeste uit zichzelf kunnen halen.

Liefste papa, je weet hoeveel je voor me betekent. Je kent me als geen ander omdat we zo hetzelfde zijn. We kunnen over alles praten, vrolijken elkaar op als het tegenzit en maken elkaar nog vrolijker als we al vrolijk zijn. Veel mensen kunnen leren van hoe optimistisch en lief jij in het leven staat. Je hebt mij zoveel geleerd en ik weet helemaal zeker dat ik zonder jouw hulp nooit zover was gekomen. Je bent mijn beste vriend en ik zal altijd van je blijven leren. Dankje voor alles lieve schat!!

Mijn lieve Mano, je hebt me enorm geholpen met het tegenlezen van mijn stukken en je vele kennis en inzichten. Slimmerd! Je laat me elke dag voelen dat we samen de hele wereld aankunnen en dat is zo fijn. Elke dag na werk maakte je me zo blij met je leuke humor en je stralende grote lach! We hebben de afgelopen jaren samen de wereld verkend door elke vakantiedag die we kregen optimaal te benutten en ik kijk enorm uit naar alle leuke avonturen die we in de toekomst samen gaan meemaken!

Liefs Lot


CPR
CPR
CPR
CPR
CPR
CPR