Stafseth, S. A. U.-B. K., Tønnessen, T. I., Diep, L. M. & Fagerström, L. (2018). Testing the reliability and validity of the Nursing Activities Score in critical care nursing. *Journal of Nursing Measurement.26*(1), 142-162.

doi: http://dx.doi.org10.1891/1061-3749.26.1.142

Journal of Nursing Measurement

Test of the reliability and validity of the Nursing Activities Score in critical care nursing

--Manuscript Draft--

Manuscript Number:	JNM-D-17-00005R1
Full Title:	Test of the reliability and validity of the Nursing Activities Score in critical care nursing
Article Type:	Original Study
Keywords:	workload critical care nursing intensity reliability validity structural equation modelling
Corresponding Author:	Siv Karlsson Stafseth, RN, CCN, MNSc Oslo Universitetssykehus Oslo, NORWAY
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Oslo Universitetssykehus
Corresponding Author's Secondary Institution:	
First Author:	Siv Karlsson Stafseth, RN, CCN, MNSc, PhD candidate
First Author Secondary Information:	
Order of Authors:	Siv Karlsson Stafseth, RN, CCN, MNSc, PhD candidate
	Tor Inge Tønnessen, MD, Professor II
	Lien My Diep, Statistician/Biostatistician
	Lisbeth Fagerstrøm, RN, PhD, Professor
Order of Authors Secondary Information:	
Abstract:	Background and Purpose. The purpose of this study was to test the reliability and validity of the conceptual model of critical care nursing using the Nursing Activities Score (NAS) for ICU patients. Methods. An observational study conducted in 2011 using the NAS scores of 219 patients in Norway. The inter-rater reliability was tested by parallel classifications. The validity was explored by an exploratory factor analysis (EFA) and a confirmatory factor analysis (CFA) as the measurement models in the structural equation model. Results. Within the paired ratings the Spearman's correlation coefficient was 0.39. The EFA results explained 77% of the variance with six factors. The reduced CFA model resulted in a 3-factor model: relationship, prevention and treatment. Conclusions. The findings supported the inter-rater reliability and construct validity of the conceptual model of the NAS.

Test of the reliability and validity of the Nursing Activities Score in critical care nursing

Introduction

Patient classification systems are developed for the classification of nursing activities and to document nursing diagnoses, outcomes and interventions (Castellan et al., 2016). Classification is relevant to provide a basis for staff planning and costs connected to the appropriate resourcing of nursing care. The accurate measurement of the nursing workload and intensity with a reliable and valid system is necessary to quantify workload, plan staffing and allocate manpower to match the various needs of patients in ICUs. The nursing intensity can be defined in four dimensions: how sick a patient is, a patient's dependency on nursing, the nursing process and the time required for nursing activities (Morris et al., 2007; Inoue et al., 2014; Fagerstrøm & Vainikainen, 2014). One patient classification system in critical care nursing is the Nursing Activities Score (NAS), which uses a combination of the amount of care and level of activity and comprises 23 items with sub-items (Miranda et al., 2003; Padilha et al., 2015). The NAS tool has been used for more than 10 years in a variety of countries, but its theoretical foundation has not been a subject for tests (Reis Miranda & Jegers, 2012). To the authors' knowledge, there is relatively scarce information concerning the reliability and validity of the NAS. There is a clear lack of studies that have explored and tested the conceptual model of critical care nursing in classification systems.

Background

The first review of the NAS tool in 36 studies explored the contexts, methods of application and purpose for using the NAS in clinical settings showed a widespread use of the instrument (Lachance et al., 2015). The NAS has been tested for several types of validity such as content, construct, criterion (concurrent and predictive) and convergent or discriminant validity in studies from several countries.

Content validity was explored using Italian critical care expert nurse focus groups (*n*=7) and in a translation process on face validity currently used by 10 experts (Palese et al., 2016). The tool NAS was considered not fully adequate to measure current ICU nursing activities and their weighting of average time consumption. The construct validity in an exploratory factor analysis of 164 scores showed that one latent factor explained 58.9% of the variance and that seven latent factors explained 70% of the NAS (Kaiser-Meyer-Olkin= .589). They concluded that the psychometric properties of the Spanish version of the NAS were acceptable (Sánchez-Sánchez et al., 2015).

Criterion-related validity was tested in concurrent validity between two patient classification systems—the TISS-28 and the NAS—in 200 patients during admission to four ICUs in Brazil (Padilha et al., 2008). In the logistic regression model, the results of a high TISS-28 (i.e., above 23 points) increased the possibility of a high NAS by 5.45 times compared to lower values. Concurrent validity was also tested between the recommendation of workload (number of nurses) in the COFEN Resolution No. 293/2004 and the NAS (Inoue & Matsuda, 2010). Based on the results from one ICU (*n*=107), the average of the NAS (697.3 %) showed the following: there was a high nursing workload; the nursing staff of this ICU should have had 40 professionals instead of 28; and the proportion of 35.7% nurses did not agree with the recommended proportion 52.5%. Similar results were evaluated in neonates in 5 ICUs between the COFEN and the NAS (Nunes & Toma, 2013). One study included patients (*n*=149) from a general ward (Panunto & Guirardello, 2009) and found that each NAS % was equivalent to 0.24 hr and that on average, 34.9% of the workload, representing 8.4 hr of nursing service, was delivered during 24-hr care. Compared to the COFEN and the obtained score, this finding was evident for a profile of patients demanding intermediary and semi-intensive care. Criterion-related validity was tested between two patient classification systems on 730 medical patients in one ICU in Spain (Carmona-Monge et al., 2013).

Pearson's correlation coefficient for individual measurements through the NAS and the NEMS corresponded to .67 and .93, respectively, for the daily workload in the unit. The results showed that staffing requirements based on NAS scores were significantly higher than those based on NEMS scores.

A more unusual way to study validity is predictive validity. In Brazil, (Ducci & Padilha, 2008) applied the NAS prospectively and retrospectively to 104 ICU patients. The results were found to show moderate correlation between the prospective and retrospective mean NAS (Pearson .65). They concluded that the prospective NAS performed well in measuring nursing workload for the subsequent 24 hr period at the ICU. This tool can be used in planning for adequate staff numbers.

Convergent validity of the NAS, here in a comprehensive and representative study of workload and APACHE II, was performed in 7 ICUs in Poland on 314 patients (Cudak & Dyk, 2009). There was no association between the proposed measurements with the unrelated construct. A significant difference in high workload (as NAS) was found between survivors and non-survivors. The NAS was empirically tested for validity in different patient samples and countries: trauma patients in Brazil (Nogueira Lde et al., 2014), general and surgical patients in Brazil (Kakushi & Martinez Evora, 2014) and in Norway (Stafseth et al., 2011) and in Italy (Lucchini et al., 2015), and workload per shift in Belgium (Debergh et al., 2012).

A body of research during the past 10 years has focused on how reliable the NAS is. Findings from Brazil indicated a strong association between the NAS score and related factors and outcomes (i.e., the Simplified Acute Physiology Score (SAPS II), length of stay, age and mortality) (Rivera-Fernandez et al., 2007; Padilha et al., 2008; Silva et al., 2011).

The inter-item correlation of the NAS was high (M of .88), and Cronbach's α was 0.99 in a Norwegian study (Stafseth et al., 2011). A Spanish study revealed that the inter-rater concordance and intra-class correlation were high between nurse-raters (Valls-Matarin et al., 2015). Findings from Norway (Stuedahl et al., 2015) on the inter-rater reliability between health professionals (i.e., critical care nurses) reported a mean NAS of 88.4%, whereas physicians had a significantly lower mean of 83.7%. This result can be converted to time and is equivalent to a difference of 1 hr per 24 hr for each patient.

Other empirical studies have focused on nursing workload and intensity. For example, NAS acted as a protective factor against the development of pressure ulcers because of the higher number of nurses who adequately cared for patients in preventing and reducing risk (Cremasco et al., 2013). Excessive nursing workload was the main risk factor for healthcare-associated infections (Daud-Gallotti et al., 2012). The NAS had a positive correlation with total family satisfaction (Gerasimou-Angelidi et al., 2014). No differences in workload were found among elderly patients (Sousa et al., 2008). Theoretical considerations of the NAS in critical care have not been explored and are of interest in the present study.

Conceptual Framework

The concept of nursing can be explained in domains of nursing care. The concepts are the basic building blocks of a theory or a conceptual model (Polit & Beck, 2014). The domains of nursing care are related to a helping role, a teaching-coaching role, diagnostic and patient-monitoring function, effective management of rapidly changing situations, administering and monitoring therapeutic interventions and regimes, monitoring and ensuring the quality of health care practices and organization, and work-role competence (Benner, 2001; Kitson et al., 2010; Kitson et al., 2013; Myny et al., 2013).

In Nordic research (Athlin, 1998; Fagerström, 1999), the core domains of the nursing process have been described as two domains: the interaction between the patient and nurse and the nursing care activities. These two domains are interconnected in surroundings with nursing resources, environment, organization and competence. Nurses’ competence is a major factor of the two domains and is connected to the patients' needs. Technology, severity of illness and a high dependency level of the patient are essential in the critical care nursing process for the optimal balance of a patient's needs.

In the main content of nursing and in critical care of nursing from Benner et al. (1999), the nursing process is divided into functional terms (interventions, instrumental functionality and solving problems) and relational terms (nursing, expressive function, inter-humanity and relationships, including communication). We (the authors) mapped and organized the NAS items with a similar content to identify the domains of critical care nursing, and we identified five latent factors. The result of the conceptual model for critical care nursing and the NAS is presented in Figure 1. The conceptual model of critical care nursing might be explained as prevention, treatment, recovery, respectful attitude and environment with integrity. Prevention is, for example, the vigilant monitoring of vital signs to avoid deteriorations and performing hygienic measures to reduce infections. Treatment includes many of the interventions performed in an ICU. Recovery is closely connected to mobilization and nursing, with increased survival. A respectful attitude is the manner in which nurses work with patients and their families in individual care plans. The environment with integrity is a part of the administration and teamwork in the unit. These five factors are evident in the professional and direct care of ICU patients (Fagerstrøm & Vainikainen, 2014, Gulbrandsen & Stubberud, 2015).

The objective of this study was to test the reliability and validity of the conceptual model of critical care nursing in the NAS for adult ICU patients.

Methods

This study consisted of a prospective observational study that was conducted in September 2011. Daily NAS scores were collected from 219 patients older than 18 years with a length of stay of more than 8 hr. Two hundred fifty-six patients were admitted to two ICUs in specialist hospitals which are run as parts of one university-affiliated organization in Norway. Excluded were 5 kids and 32 adults with LOS < 8 hr. The ICUs admitted patients from all specialties, but most of the patients were from trauma, surgery and infectious disease medicine (Table 1).

Data collection

Caregiving nurses scored the NAS in the internal quality database as part of their daily clinical routine. The nurses had the educational level of a registered nurse (RN) with or without or additional European standard critical care nursing (CCN) specialization of 90 ECTS credits, and they had undergone standardized NAS training sessions. Theoretically, each nurse could have rated one or more patients, and altogether in the two ICUs, 83 and 65 nurses were employed at that time. We did not determine who performed the scoring. Only one NAS rating from each patient on the first day of admission was included in this study. For Item 16, “Hemofiltration techniques, dialysis techniques,” we collected the NAS rating on the very first day of treatment. To test the reliability, two nurses independently rated the NAS for a given patient on the same day. The caregiving nurses on duty provided the first rating, and based on the notes in a patient’s record, a researcher provided the second rating after discharge from the ICU.

Instrument

The NAS is a patient classification system that measures the nursing workload performed at the patient level (Appendix 1). The average time and personal consumption for therapeutic procedures and nursing activities is included. The nursing activities include monitoring, hygiene, mobilization, administrative activities and support to the patient and relatives (Miranda et al., 2003). The NAS tool consists of 23 items with sub-items from the lowest level of care (Item 1a, 4a, 6a, 7a, 8a), and it can range in score from 22.3 to 177. It can be collected retrospectively once per shift or once per day. Each individual item is given a weight of 1.2 to 32, and each represents the percentage of time that is spent by one nurse on a specified activity. A higher score indicates a corresponding increase in nursing time used to care for the patient (Miranda et al., 2003). The translation of the NAS tool to a Norwegian version has been piloted and was explored for the concurrent validity to the NEMS (Stafseth et al., 2011), and this version was applied to the present study.

Ethical considerations

This study was conducted in accordance with the Helsinki Declaration and was approved by the Regional Committees for Medical and Health Research Ethics in Norway (REK 2011/2325). Approval was obtained from the Institutional Review Board at the hospital (#2011/15341). The project did not affect the patients or their care.

Statistical analysis

The data were analysed using SPSS Inc. Released 2009. PASW Statistics for Windows, Version 18.0 Chicago, IL, USA and STATA/SE version 14 for Windows (StataCorp LP, TX, USA). Descriptive statistics were used to describe the clinical characteristics of the sample.

Reliability

To test the reliability, paired ratings data were analyzed by descriptive statistics and Spearman's correlation coefficient. The strength of a relationship is defined as follows: 0.1 to 0.3, weak; 0.3 to 0.5, moderate; and 0.5 to 1.0, strong. Differences between the two raters' data were tested by a one-sample *t*-test. The agreement of the two raters' data was examined by Cohen's weighted or unweighted Kappa (*K* or *Kw*). Cohen's kappa controls for the amount of agreement that might have occurred merely by chance whereas the weighted kappa reflects the relative seriousness in disagreement between raters (Cohen, 1968). The unweighted kappa is a well-established statistical measure of inter-rater agreement for categorical items. It has a maximum of 1.0 if the agreement is perfect. In the guidelines for interpretation of the results as follows: 0.0-0.2, slight; 0.21-0.40, fair; 0.41-0.60, moderate; 0.61-0.80, substantial; and 0.81-1.0, almost perfect (Landis & Koch, 1977). We decided to investigate nursing activities (not therapeutic procedures) because they represent the highest percentage of items in the NAS and have two or three levels. For Items 1, 4, 6 and 8, which represent the nursing activities in the NAS, the proportions and 95% confidence intervals (CIs) of agreement between the two raters were given according to the Guidelines for Reporting Reliability and Agreement Studies (GRRAS) in Kottner et al., (2011). One additional item, Item 7, represents nursing activity and has a high percentage in the NAS; however, it has two ordinal categories.

Validity

A theoretical exploration in a conceptual model was performed to confirm the empirical assessment. To be able to conduct a factor analysis, the number of items (e.g. 23) and domains (e.g. 5) must be proportionate to the number of scorings (Watson & Thompson, 2006). From each patient only one NAS score was collected and we included the patients' NAS scores until we reached > 5 scores in each item. No repeated NAS scores or items were used from a single patient. The sample size was powered to be > 200 (Kaariainen et al., 2011). In the factor analysis, we used the scores from the caregiving nurses as recommended under conditions as close as possible to the daily routine (Kaariainen et al., 2011; Kottner et al., 2011).

In the NAS from the sample from caregiving nurses, four of the 23 items (i.e. q1, q4, q6 and q8) had 3 ordinal categories, which were merged into 2 categories before performing the exploratory factor analysis (EFA) and the confirmatory factor analysis (CFA) embedded in a structural equation model (SEM). The scores of three patients who were assigned to the first category of q1 were excluded because of a large difference in the weight of the categories. The other items displayed very small differences between the last two categories; therefore, the two categories were combined into one category. The scores of two patients for item q4 were missing. To test the validity of the NAS on the very same sample i.e. caregiving nurses, an exploratory factor analysis (EFA) was also conducted on tetrachoric correlations of binary variables (Bock & Lieberman, 1970). The comparison of eigenvalues, which were determined by findings in the matrix such as a vector or size, was performed by parallel analysis with Monte Carlo PCA using 1,000 replications on original data and items.

The CFA was performed as the measurement model in the SEM (Flora & Curran, 2004; Kline, 2016). The tetrachoric correlation matrix was estimated a prior to conducting the EFA and CFA (Flora & Curran, 2004). In a SEM, it is possible to examine the accuracy of the correlation between a theoretical model and an empirical model, resulting in an estimation of the model fit. The model modification indices (by adding more error covariance between the variables) and goodness-of-fit (statistics) were inspected using the post-estimation commands. The goodness-of-fit for the model was assessed by the Akaike information criterion (AIC), the likelihood ratio statistics chi-squared test [χ^2 (*df*)], the root mean square error of approximation (RMSEA), the comparative fit index (CFI) and Tucker-Lewis fit index (TLI) estimates (STATA SEM Manual). We compared the AIC values from the models of interest, and the model with the lowest value was preferred. CFI and TLI levels close to or greater than 0.95 are acceptable (Watson & Thompson, 2006). One and three factor model was fitted and their goodness of fit was compared. On statistical aspect, 1-factor model would have been preferred (because the factors were strong correlated and the sample size may not need to increase). The CFA embedded in the SEM was used to examine and confirm the a priori specified theoretical model consisting of five latent factors related to prevention, treatment, recovery, respectful attitude and environment with integrity. The relationship between the observed and latent factors was first tested as a full factor model with ‘prevention’ added to the ‘recovery’ factor. The second model was modified with item 19 removed from the factor 'treatment' because of a strong correlation between items 16 and 19. The third model was a reduced factor model resulting in three latent factors (Figure 3). The relationship between the dimensions was also tested. The significance level was set to 0.05.

Results

Reliability

A sample of NAS ratings from 219 patients in two ICUs was collected in September 2011. Of the enrolled patients, 141 were male (64.4 %) and 78 (35.6%) were female. The primary diagnosis of admitting patients are presented in Table 1. The average age was 54.7 (*SD* ±20.1). The average length of stay was 2.7 days (*SD*=6.2) and the median was 1.0 day. The length of stay varied between 8 hours and 58 days. There was no significant gender difference regarding the length of stay (*U* test=5228, *z*=-.62, *p*=.54). A one-sample *t*-test showed a significant difference in the NAS score between the caregiving nurses (NAS values *M*= 99.1, *SD* = 33.7, with min-max of 36.9-168.6) and the research nurses (NAS values *M*= 114.3, *SD* = 23.6, with min-max of 54.3-160.7) (*p* = .03). The Spearman correlation between the raters was 0.39 (95% CI: 0.25-0.51) and was significant (*p* = .01). The average inter-rater reliability was substantial using Cohen's weighted kappa (*Kw*) with a value of 0.85 in items 1, 4, 6, and 8 of the NAS. As illustrated, the inter-rater reliability (IRR) for nursing activities (Table 2) ranged from 0.48 to 0.91 in the unweighted Kappa (*K*), and the proportions of agreement varied from 0.37-0.52.

Validity

The findings of the EFA and the correlation matrix revealed the presence of many coefficients of 0.4 and above (Table 3). An initial analysis was performed to obtain eigenvalues for each component in the data. The scree plot (Figure 2) warranted six components, which were retained for the final analysis. Six components explained 77% of the variance. All factors are interrelated to some degree. The items that cluster on the same factors suggest that Factor 1 represents “treatment” and Factor 2 “prevention” the results cannot explain all factors in our conceptual model. The decision of six components was supported by the results of the parallel analysis, which showed components with eigenvalues exceeding the corresponding criterion values for a randomly generated data matrix of the same size (23 variables x 219 scores). Table 3 shows the six-component solution with the pattern matrix (regression coefficients for each variable on each factor).

The results from the CFA embedded in the SEM showed a conceptual model for critical care nursing, which a reduced factor model was resulting in three latent factors (Figure 3 and Table 4). The standardized factor loading between the latent variables was significant in the model. The model displays an acceptable fit to the data because the results showed the following: the likelihood ratio for the model vs. saturated in non-significant chi-square $\chi 2$ $(170) = 188.2$, $n = 214$ and $p = 0.161$ with CFI = 0.98, TLI = 0.98, AIC = 11,939.3 and RMSEA = 0.02. The goodness of fit for 1-factor model is also shown in Table 5. We investigated the information and found cross loadings between the error measurements. The theoretical connection between cross loading occurred if the items had something in common, for example, item 16 “Hemofiltration techniques, dialysis techniques” and item 19 “Treatment of complicated metabolic acidosis or alkalosis” (tetrachoric correlation coefficient = .99).

Discussion

This study demonstrated that a conceptual model for critical care nursing of observed NAS could be confirmed in a reduced CFA embedded in an SEM. The three latent factors “relationship”, “prevention” and “treatment” could easily be recognized in our conceptual model. For validity, we compared the results from the CFA and EFA and discovered similarities, and for reliability, a moderate correlation was found between raters in the ICUs.

A well-established fact when studying the construct of a tool is that both empirical and theoretical considerations must be made. The latent construct is formed and determined through a combination of items in the NAS, i.e., the workload and time spent on various nursing activities. The NAS consists of several themes or areas, which are not interchangeable. Adding or dropping an item might increase or decrease only the total score but not the domain of the construct. We had to theoretically identify the NAS items by the terms “functional” and “relational”, with five latent factors to describe the conceptual model for critical care nursing (Figure 1).

Interestingly in our research we initially treated the binary items as continuous in the CFA, and the conceptual model for critical care nursing was not confirmed by the observed data. When the CFA was conducted on tetrachoric correlations, assuming that our items reflected underlying and unobserved items, the model did confirm. The statistical results indicated that the sample size should have been larger, or strong correlated items should be removed to reduce the numbers of items in the conceptual model and avoid collinearity which may disturb the estimation of parameters (Watson & Thompson, 2006). After performing several analyses described earlier, the CFA was finally reduced to three latent factors which were strong correlated with each other (see the covariance in Table 4). On clinical aspect the 3-factor model would be preferred. We made the choice of model: the 1-factor model could not describe the “reality” or the empirical world of critical care nursing. These results can be interpreted in light of Nordic research in two domains: the interaction between the patient and nurse and the nursing care activities (Athlin, 1998; Fagerström, 1999). As mentioned previously, the surroundings/environment and nursing competence are essential in critical care nursing. We found a perfect negative correlation between the latent factor “prevention” and “treatment”. “Prevention” is a nursing activity that can prevent or even avoid deterioration of patients and the necessity of starting “treatment”. A strong positive correlation was found between “prevention” and “relationship”, which corresponded well to the importance of continuity of nurses being able to get to know the patient. Finally, a strong negative correlation between “treatment” and “relationship” was determined: if patients are under heavy “treatment”, most are on medication (in medically induced or natural comas). The findings in the present study of strong/ perfect correlations indicate a NAS tool which describes interrelationships’ between all the three factors. However it must be understood that the NAS mostly assesses direct care and not all types of indirect care (Kakushi & Martinez Evora, 2014). To our knowledge, this study is the first to explore the NAS both empirically and theoretically from the critical care nursing point of view.

Our results showed that six latent factors in the EFA explained 77% of the variance; these results were similar to a study performing an EFA on the NAS (the only study identified in the literature), which studied 164 patients and found that 70% of the variance could be described by seven latent factors (Sánchez-Sánchez et al., 2015). We must briefly discuss the differences and similarities of the statistical results from the EFA in three ways (internal validity): first, items on the original scale without dichotomization; second, items on a dichotomized/binary level; and third, items on a binary dichotomized/binary level using a tetrachoric correlation matrix. For the original scale, we performed the analysis without dichotomization; for the items on a dichotomized/binary level, we performed an EFA on the binary items with rotation (results not shown in paper). For the items on a binary dichotomized/binary level, we performed the analysis on a tetrachoric level for the complex interpretation of CFA (on same level); some of the results are reported in this paper. Surprisingly in the latent factor “treatment”, 11 out of 12 items in NAS had negative correlations (significant) between the observed variables. That indicates us to remove the factor from the model because the model is statistically overfitted and too strong correlated with “prevention”. However, in the tool NAS “treatment” represents 41.6 % (of 177 %). The "treatment" factor correlated positively with its corresponding items when the "prevention" factor was not included. The statistical results and goodness of fit without the "prevention" factor in the model were not confirmed by SEM because the fitted model was not full rank, but the model without the "treatment" factor in the model fitted very well. It seems likely that the "prevention" factor has a priority from the items when both the "prevention" and "treatment" factors were in the model. However it might be in comparison of "treatment" to the nursing activities (i.e. five items in NAS) that have high scores in the instrument. In the clinical setting and in critical care nursing, the treatment of ICU patients counts towards highly specialized health care and its complexity.

We have followed the tradition and development of patient classification system from large systems to index systems, and to date, the NAS is beneficial because it is a system that is derived from nursing time and is able to explain 81% of the nursing time (Miranda et al., 2003; Watson & Thompson, 2006; Reis Miranda & Jegers, 2012). It is important to recall the background and history of the development of systems, particularly the tool NAS, from the assessment of medical interventions to adding five areas of nursing activities (monitoring, hygiene, mobilization, administrative activities and support to patients and relatives). As mentioned previously, we could only identify one earlier factor analysis on the NAS, and it would be of interest to see whether similar results are reported in other countries in larger samples.

In the literature on reliability tests of tools, several methods have been reported (Polit & Beck, 2012). We tested reliability in paired NAS ratings; in our study, the caregiving nurses (daily bedside) and the researcher derived NAS ratings based on notes in the patient's records. The method was chosen because of the short time period to obtain the data. In our study, the raters (nurses) were responsible for daily ratings, and because of organizational and financial constraints, we could not reduce the numbers of raters. Presenting scores from several raters might have strengthened the results. One reliable method, in which a health professional scored the same patient as the other raters, was used in a study of inter-rater reliability (Stuedahl et al., 2015). One earlier study conducted a pre-test (reliability) of understanding of each item in the NAS (Ducci & Padilha, 2008). In their study, the nurse and the researcher scored the NAS from records for both retro- and prospective applications. They found no significant differences in the average NAS between raters, which do not surprise us because the information that can be assessed from records is limited. In the same study, the degree of agreement was described to range from 23-71.8%, and the Kappa (*Kw*) of nursing activities with sub-items was found to be low, ranging from 0.03 to 0.46 (Ducci & Padilha, 2008). In scientific debate, there are arguments that reliability and agreement should be investigated under conditions as close as possible to the daily routine (Kottner et al., 2011). Therefore, we used the daily scores from care-giving nurses for many of the analysis in the present study.

The tool NAS is constructed with a variation in item measures, for example, item 1 a, b or c. There are variations in time-consumption and the level of treatment. Therefore, it is evident that in the present study, the NAS from the researcher might have been the result of nurse documentations in the records. Our results indicated that the researcher could have both under or overestimated the NAS from the patients' records. This improper estimation could be a result of difficulties in determining the level and weighting of each nurse's time consumption and workload from the records, or a result of an inaccuracy in the interpretation of items. Several other correlation studies (Padilha et al., 2008; Altafin et al., 2014; Nogueira Lde et al., 2014) were based only on data from researchers and not from caregiving nurses. A better method might be to test the reliability independently from two caregiving nurses (raters) on the same day rather than from records.

As a patient classification system, the NAS has limitations in the assessment of workload. The NAS can explain approximately 81% of direct care; hence, it is not possible to assess all nursing activities (Miranda et al., 2003). One major aspect to consider is the actual number of nurses on duty, which might result in a higher score. For example, for item 6 “Mobilization and positioning”, if three or more nurses perform the procedure, a high weight is given to item 6c, but if fewer nurses are available, it is only possible to score a or b with lower weights.

In summary, the inter-rater reliability of the NAS was significant and provided empirical support for the stability of the system. The test of validity performed well in the EFA and CFA. We recommend the use of whole ratings in the NAS, as suggested in previous studies (Miranda et al., 2003; Reis Miranda & Jegers, 2012). Critical care nursing is complex and has a huge and obvious human relations aspect; nurse competency requires more than technical skill alone. The factors related to respectful attitude and providing an environment with integrity (Ääri et al., 2008; Lakanmaa et al., 2012) are not easy to assess or quantify, and communication is one of the core values in critical care nursing. Nevertheless, different instruments are only as good as each assessment and performance in delivering care to intensive care patients. Caring is more than the sum of time consuming activities (Fagerström, 1999), and more than the sum of NAS percentages. We thus encourage future nursing research to explore the tool NAS by recalibrating the weights attributed to the items in the NAS.

Limitations and strenghts

This study provides new insight, both theoretical and empirical, into critical care nursing. However, some limitations have been mentioned and should be considered when interpreting the results. A strength in the data collection was the equivalence of a system by inter-rater reliability, in which two persons scored the NAS independently using standardized scoring criteria. It is uncertain whether the methodological treatment of the data to recode 3 ordinal categories into 2 might have influenced the findings. The number of scores of respectively a, b or c; whereas for example in Item 8. However, we cannot find any evidence that the recoding influenced our results.

We chose to treat the data as binary data, with items representing underlying continuous variables. The technique for estimating binary data in estimating correlation is applicable for tetrachoric correlations. This technique is perfect for rating scales (such as the NAS) with a small number of observations. We treated our variables as binary in the preferred model (Figure 3). For the analysis by EFA, a more comprehensive sample size would have strengthened the study.

Implications for Practice

The Nursing Activities score will allow nurses and nurse managers to assess critical care nursing. Use of the tool NAS could be valid for nurse allocation in critical care, and in nursing management. This knowledge will increase the novel understanding of assessment of nurses' workload in intensive care units.

Conclusion

The tests of reliability and validity of the conceptual model in the NAS on ICU patients provided empirical support for its usefulness in the assessment of critical care nursing. The findings supported the inter-rater reliability and construct validity of the NAS scores in two ICUs in Norway. Continued amendment and validation of the NAS on a larger sample of patients from different ICUs and countries would be required to estimate the workload.

References

Altafin, J.A., Grion, C.M., Tanita, M.T., Festti, J., Cardoso, L.T., Veiga, C.F., Kamiji, D., Barbosa, A.R., Matsubara, C.C., Lara, A.B., Lopes, C.C., Blum, D. & Matsuo, T. (2014) Nursing Activities Score and workload in the intensive care unit of a university hospital. *Rev Bras Ter Intensiva, 26(3),* 292-8.

Athlin, E. (1998) Nursing based on an interaction model applied to patients with eating problems and suffering from Parkinsons's disease and dementia. In *University of Umeå, Sweden*, Vol. Dissertation University of Umeå, Umeå.

Benner, P., Hooper-Kyriakidis, P. & Stannard, D. (1999) *Clinical wisdom and interventions in critical care: a thinking-in-action approach,* Saunders, Philadelphia.

Benner, P. (2001) *From Novice to Expert. Excellence and Power in Clinical Nursing Practice,* Prentice-Hall, Upper Saddle River NJ USA.

Benner, P., Hooper-Kyriakidis, P. & Stannard, D. (1999) *Clinical wisdom and interventions in critical care : a thinking-in-action approach,* Saunders, Philadelphia.

Bock, R.D. & Lieberman, M. (1970) Fitting a response curve model for dichotomously scored items. *Psychometrika, 35*, 179-198.

Carmona-Monge, F., Rollán Rodríguezb, G., Quirós Herranz, C., Gómez, S. & Marín-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. *Intensive and Critical Care Nursing, 29*, 228-233.

Castellan, C., Sluga, S., Spina, E. & Sanson, G. (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.

Cohen, J. (1968) Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychologic Bulletin, 70*, 213-220.

Cremasco, M.F., Wenzel, F., Zanei, S.S. & Whitaker, I.Y. (2013) Pressure ulcers in the intensive care unit: the relationship between nursing workload, illness severity and pressure ulcer risk. *J Clin Nurs, 22(15-16)*, 2183-91.

Cudak, E. & Dyk, D. (2009) [Nursing demand in intensive therapy units]. *Anestezjologia intensywna terapia, 42(2)*, 70-74.

Daud-Gallotti, R.M., Costa, S.F., Guimaraes, T., Padilha, K.G., Inoue, E.N., Vasconcelos, T.N., da Silva Cunha Rodrigues, F., Barbosa, E.V., Figueiredo, W.B. & Levin, A.S. (2012) Nursing workload as a risk factor for healthcare associated infections in ICU: a prospective study. *PLoS One, 7(12)*, e52342.

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. *Intensive Care Med, 38(9)*, 1438-44.

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(4)*, 581-7.

Fagerström, L. (1999) The Patient's Caring Needs To understand and measure the unmeasurable. In *Åbo Akademi University*, Vol. Doctoral Åbo Akademi University, Åbo.

Fagerstrøm, L. & Vainikainen, P. (2014) Nurses' Experiences of Nonpatient Factors That Affect Nursing Workload: A Study of the PAONCIL Instrument's Nonpatient Factors *Nursing Research and Practice, 9*, ISSN 2090-1429.

Flora, D.B. & Curran P.J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods, 9*, 466-491.

Gerasimou-Angelidi, S., Myrianthefs, P., Chovas, A., Baltopoulos, G. & Komnos, A. (2014) Nursing Activities Score as a predictor of family satisfaction in an adult Intensive Care Unit in Greece. *J Nurs Manag, 22(2)*, 151-8.

Gulbrandsen, T. & Stubberud, D.-G. (2015) *Intensivsykepleie,* Cappelen Damm akademisk, Oslo.

Inoue, K.C. & Matsuda, L.M. (2010) Sizing the nursing staff in an intensive care unit for adults. *Acta Paulista de Enfermagem, 23(3)*, 379-384.

Inoue, K.C., Versa, G.L. & Matsuda, L.M. (2014) Stress level among intensive care nurses in the municipality of Parana (Brazil). *Invest Educ Enferm, 32(1)*, 69-77.

Kaariainen, M., Kanste, O. & Elo, S. (2011) Testing and verifying nursing theory by confirmatory factor analysis. *Journal of Advanced Nursing, 67(5)*, 1163-1172.

Kakushi, L. & Martinez Evora, Y. (2014) Direct and indirect nursing care time in an intensive care unit. *Rev Lationo-Am Enfermagem, 22(1)*, 150-157.

Kitson, A., Conroy, T., Wengstrom, Y., Profetto-mcgrath, J. & Robertson-malt, S. (2010) SCHOLARLY PAPER: Defining the fundamentals of care. *International Journal of Nursing Practice, 16(4)*, 423-434.

Kitson, A., Marshall, A., Bassett, K. & Zeitz, K. (2013) What are the core elements of patient centred care? A narrative review and synthesis of the literature from health policy, medicine and nursing. *Journal of Advanced Nursing , 69*, 4-15.

Kline, R.B. (2016) *Principles and Practice of Structural Equation Modeling,* The Guilford Press, New York NY USA.

Kottner, J., Audige, L., Brorson, S., Donner, A., Gajewski, B.J., Hróbjartsson, A., Roberts, C., Shoukri, M. & Streiner, D.L. (2011) Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *International Journal of Nursing Studies, 48(6)*, 661-671.

Lachance, J., Douville, F., Dallaire, C., Padilha, K.G. & Gallani, M.C. (2015) The use of the Nursing Activities Score in clinical settings: an integrative review. *Revista da Escola de Enfermagem da USP, 49(SPE)*, 147-156.

Lakanmaa, R.-L., Suominen, T., Perttilä, J., Puukka, P. & Leino-Kilpi, H. (2012) Competence requirements in intensive and critical care nursing – Still in need of definition? A Delphi study. *Intensive and Critical Care Nursing, 28(6)*, 329-336.

Landis, J.R. & Koch, G.G. (1977) The Measurement of Observer Agreement for Categorical Data. *Biometrics, 33*, 159-174.

Lucchini, A., Elli, S., Bambi, S., Becattini, G., Vanini, S., Piantanida, C., Iacca, C. & Sartori, D. (2015) Nursing activities score: differences in nursing workload in three intensive care units. *Assist Inferm Ric, 34(1)*, 6-14.

Miranda, D.R., Nap, R., de Rijk, A., Schaufeli, W. & Iapichino, G. (2003) Nursing activities score. *Crit Care Med, 31(2)*, 374-82.

Morris, R., MacNeela, P., Scott, A., Treacy, P. & Hyde, A. (2007) Reconsidering the conceptualization of nursing workload: literature review. *J Adv Nurs, 57(5)*, 463-71.

Myny, D., De Bacquer, D., Van Hecke, A., Beeckman, D., Verhaeghe, S. & Van Goubergen, D. (2013) Validation of standard times and influencing factors during the development of the Workload Indicator for Nursing. *J Adv Nurs, 70(3)*,674-686.

Nogueira Lde, S., Domingues Cde, A., Poggetti, R.S. & de Sousa, R.M. (2014) Nursing workload in intensive care unit trauma patients: analysis of associated factors. *PLoS One, 9(11)*, e112125.

Nunes, B.K. & Toma, E. (2013) Assessment of a neonatal unit nursing staff: application of the Nursing Activities Score. *Rev Lat Am Enfermagem, 21(1)*, 348-55.

Padilha, K.G., de Sousa, R.M., Queijo, A.F., Mendes, A.M. & Reis Miranda, D. (2008) Nursing Activities Score in the intensive care unit: analysis of the related factors. *Intensive Crit Care Nurs, 24(3)*, 197-204.

Padilha, K.G., Stafseth, S.K., Solms, D., Hoogendoom, M., Monge, F.J.C., Gomaa, O.H., Giakoumidakis, K., Giannakopoulou, M., Gallani, M.C. & Cudak, E. (2015) Nursing Activities Score: an updated guideline for its application in the Intensive Care Unit. *Revista da Escola de Enfermagem da USP, 49(SPE)*, 131-137.

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*,548-59.

Panunto, M.R. & Guirardello, E.d.B. (2009) Nursing workload at a gastroenterology unit. *Revista Latino-Americana de Enfermagem, 17*, 1009-1014.

Polit, D.F. & Beck, C.T. (2012) *Nursing Research: Principles and Methods,* Lippincott Williams &Wilkins, Philadelphia Pa USA.

Polit, D.F. & Beck, C.T. (2014) *Essentials of nursing research: appraising evidence for nursing practice,* Wolters Kluwer/Lippincott Williams & Wilkins, Philadelphia.

Reis Miranda, D. & Jegers, M. (2012) Monitoring costs in the ICU: a search for a pertinent methodology. *Acta Anaesthesiol Scand, 56(9)*, 1104-13.

Rivera-Fernandez, R., Nap, R., Vazquez-Mata, G. & Reis Miranda, D. (2007) Analysis of physiologic alterations in intensive care unit patients and their relationship with mortality. *J Crit Care, 22(2)*, 120-8.

Sánchez-Sánchez, M., Arias-Rivera, S., Fraile-Gamo, M., Thuissard-Vasallo, I. & Frutos-Vivar, F. (2015) Validación de la versión en castellano del Nursing Activities Score. *Enfermería Intensiva, 26(2)*, 63-71.

Silva, M.C., Sousa, R.M. & Padilha, K.G. (2011) Factors associated with death and readmission into the Intensive Care Unit. *Revista Latino-Americana de Enfermagem, 19*, 911-919.

Sousa, R.M., Gonçalves, L.A., Toffoleto, M.C., Leão, K. & Padilha, K.G. (2008) Predictors of nursing workload in elderly patients admitted to intensive care units. *Revista latino-americana de enfermagem, 16(2)*, 218-223.

Stafseth, S.K., Solms, D. & Bredal, I.S. (2011) The characterisiation of workloads and nursing staff allocation in intensive care units: a descriptive study using the Nursing Activities Score for the first time in Norway. *Intensive and Critical Care Nursing, 27(5)*, 290-294.

STATA STRUCTURAL EQUATION MODELING REFERENCE MANUAL RELEASE 14, Factor analysis and SEM with tetrachoric and polychoric correlations: Example 15 titled "Higher-order CFA" (p.229), and MV (Multivariate statistics) (p.316 – 357), Uebersax JS,Texas.

Stuedahl, M., Vold, S., Klepstad, P. & Stafseth, S.K. (2015) Interrater reliability of Nursing Activities Score among Intensive Care Unit health professionals. *Revista da Escola de Enfermagem da USP, 49(SPE)*, 117-122.

Valls-Matarin, J., Salamero-Amoros, M. & Roldan-Gil, C. (2015) Analysis of the workload and the use of the nursing resources in an intensive care unit. *Enfermeria Intensiva, 26 (2)*,72-81.

Watson, R. & Thompson, D.R. (2006) Use of factor analysis in Journal of Advanced Nursing: literature review. *Journal of Advanced Nursing, 55(3)*, 330-341.

Ääri, R.-L., Tarja, S. & Helena, L.-K. (2008) Competence in intensive and critical care nursing: A literature review. *Intensive and Critical Care Nursing, 24(2)*, 78-89.

Figure 1

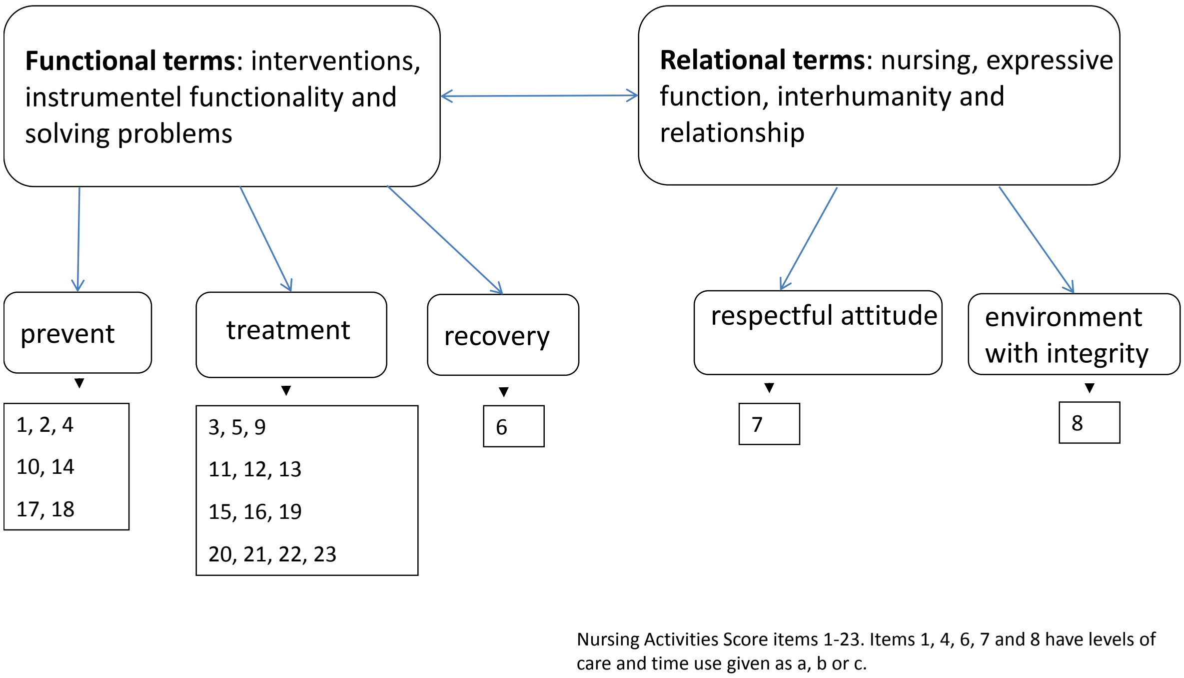

Figure 1 Conceptual model of critical care nursing using NAS items 1-23.

Figure 2

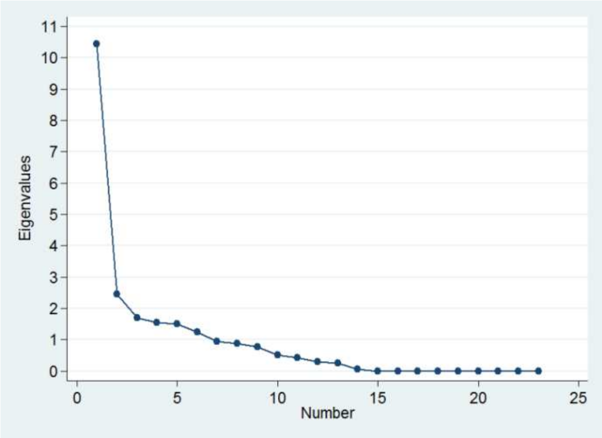

Figure 2 A scree plot of eigenvalues for components within the NAS of 23 items by EFA. Data are based on tetrachoric correlation.

Figure 3

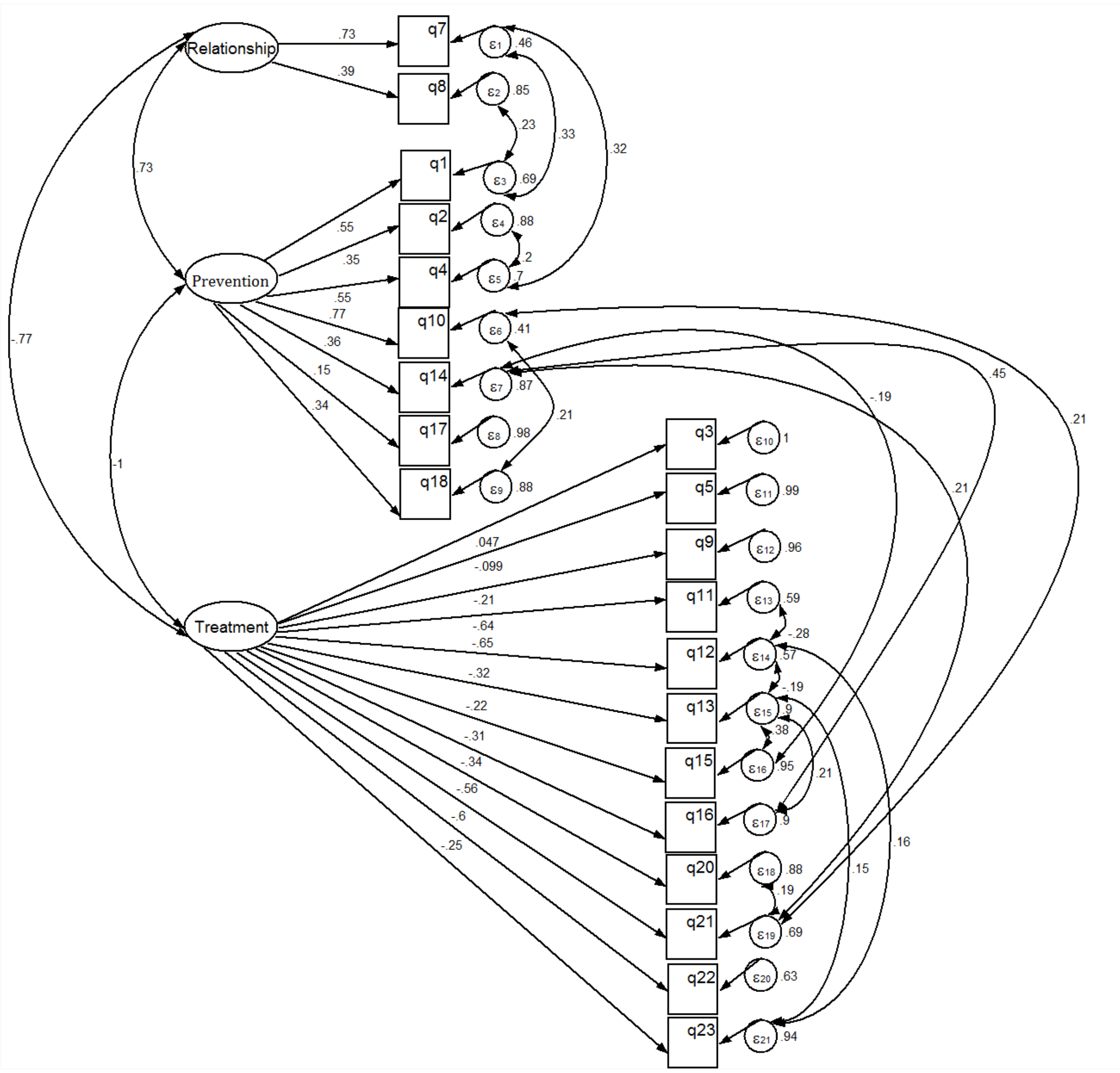

Figure 3 Reduced factor model CFA embedded in SEM with three latent factors in the NAS. The relationship between the observed variables and underlying latent factors in the NAS instrument are shown. The ovals contain latent factors, the boxes contain items in the NAS (q), and all significant correlations are presented and are estimated factor loadings (standardized).

→ Unidirectional effect between variables

↔ Correlation or covariance between variables

Appendix 1. Nursing Activities Score in basic activities and %, (Reproduced from Miranda et al,, 2003; p 378, with permission from the author).

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 thoseexamples 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; movingfrom 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/m2/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.

Table 1 Primary diagnosis of patients from an internal quality database.

Diagnosis	
1. Acute surgery	*n*=108 (49.3%)
2. Trauma	*n*=69 (31.5%)
3. Medical	*n*=37 (16.9%)
4. Cancer	*n*=5 (2.3%)
Total	*N=219 (100%)*

Table 2 Inter-rater reliability of nursing activities from paired ratings of four NAS items (N = 219).

Item in the NAS (short labels)	Unweighted Kappa (K)	Proportions of agreement (95% CI of observed)
1. Monitoring	0.48	0.52 (0.45-0.58)
4. Hygiene	0.91	0.50 (0.43-0.57)
6. Mobilization	0.55	0.37 (0.31-0.44)
8. Administrative	0.67	0.47 (0.40-0.54)

Values are given between caregiving nurses' scores and researchers' scores using unweighted Kappa (K); proportions of agreement and confidence interval (CI) are shown when there is agreement on items.

Table 3 EFA with oblimin rotation factor loading in a pattern matrix of six NAS item factors.

Variable	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
q1	**-0.75**	-0.23	0.03	-0.01	-0.18	-0.07
q2	**-0.71**	-0.29	-0.07	0.05	0.11	**-0.59**
q3	0.13	0.21	**0.96**	-0.02	0.01	-0.01
q4	**-0.88**	0.07	0.04	0.07	-0.12	-0.01
q5	0.04	0.00	-0.09	-0.03	-0.02	0.09
q6	**-0.82**	-0.01	0.20	0.00	0.10	0.29
q7	**-0.77**	-0.07	0.04	0.06	-0.12	-0.05
q8	**-0.61**	-0.07	-0.04	-0.01	-0.12	0.07
q9	**0.47**	-0.34	0.13	**0.77**	0.05	0.19
q10	**-0.71**	0.01	0.16	0.24	-0.13	-0.28
q11	**0.83**	-0.07	-0.01	0.08	0.05	0.06
q12	**0.61**	0.00	-0.19	0.02	0.27	0.06
q13	**0.75**	0.15	0.33	-0.01	0.29	0.11
q14	**-0.81**	0.38	0.28	0.00	0.04	-0.02
q15	0.25	**0.87**	0.30	-0.02	0.04	0.07
q16	**0.88**	-0.04	**0.41**	0.03	-0.05	-0.04
q17	**-0.51**	0.31	-0.16	**0.77**	0.03	0.00
q18	-0.37	**0.72**	0.01	0.02	-0.27	-0.04
q19	**0.87**	0.03	**0.43**	0.03	-0.09	-0.09
q20	**0.51**	-0.04	-0.07	**0.40**	-0.01	**0.71**
q21	**0.82**	0.01	-0.18	0.00	-0.14	0.21

q22	**0.75**	**0.52**	0.00	0.14	0.02	-0.38
q23	0.23	-0.07	0.00	0.04	**0.96**	-0.03
Eigenvalue	10.05	1.88	1.76	1.45	1.31	1.29
Proportion	0.44	0.08	0.08	0.06	0.06	0.06

Note: Boldfaced values are above 0.40, and the pattern matrix is based on tetrachoric correlations.

Table 4

Table 4 Influence of factors on the responses to items in the reduced factor model.

Latent factor	Item	Regression/ Correlation coefficient	Standard error (*SE*)	95% CI	*p*-value
RELATIONSHIP	q7	0.73	0.09	0.56-0.91	< 0.001
	q8	0.39	0.07	0.25-0.54	< 0.001
PREVENTION	q1	0.55	0.05	0.45-0.66	< 0.001
	q2	0.35	0.07	0.22-0.47	< 0.001
	q4	0.55	0.05	0.45-0.66	< 0.001
	q10	0.77	0.04	0.69-0.84	< 0.001
	q14	0.36	0.06	0.23-0.48	< 0.001
	q17	0.15	0.07	0.01-0.29	0.034
	q18	0.34	0.07	0.22-0.47	< 0.001
TREATMENT	q3	0.05	0.07	-0.10-0.19	0.525
	q5	-0.10	0.07	-0.24-0.04	0.176
	q9	-0.21	0.07	-0.35-0.07	0.003
	q11	-0.64	0.05	-0.74-0.54	< 0.001
	q12	-0.65	0.05	-0.75-0.56	< 0.001
	q13	-0.32	0.07	-0.45-0.18	< 0.001
	q15	-0.22	0.07	-0.36-0.08	0.002
	q16	-0.31	0.07	-0.44-0.18	< 0.001
	q20	-0.34	0.07	-0.47-0.22	< 0.001
	q21	-0.56	0.05	-0.66-0.45	< 0.001

	q22	-0.60	0.05	-0.70-0.51	< 0.001
	q23	-0.25	0.7	-0.39 -0.11	< 0.001
Covariance	PR and TR	-1	0.03	-1-0.98	< 0.001
	PR and RE	0.73	0.10	0.53-0.93	< 0.001
	TR and RE	-0.77	0.10	-0.96-0.58	< 0.001

PR = prevention, TR = treatment, RE = relationship

Correlation coefficients (i.e., standardized coefficients) and corresponding standard errors (*SE*), 95% confidence intervals and *p*-values, and correlation between the factors, N = 214 are shown.

Table 5 The fit statistics for the degree of similarity between the observed and expected models and confirmation of the factor structure for the 3-factor models and 1-factor model.

CFA model embedded in SEM	Latent factors	Items in NAS or error covariance added /removed	x^2 (*df*) = value for the present model vs. the saturated model	CFI	TLI	AIC	RMSEA
1. Full factor model	5	Item 6 was added to treatment factor, and no error covariance was added	x^2 (*227*) = 898.5 $p < 0.001$	0.53	0.47	13,275.7	0.12
2. Modified factor model	5	Item 6 was removed, and error covariance was added	x^2 (*186*) = 209.6 $p = 0.113$	0.98	0.98	12,118.3	0.02
3. Reduced factor model	3	Item 6 and item 19 were removed, and error covariance for items 13 and 23 was added	x^2 (*170*) = 188.2 $p = 0.161$	0.98	0.98	11,939.3	0.02
4. 1- factor model	1	Item 6 and item 19 were	x^2 (*173*) = 196.0	0.97	0.97	11,941.1	0.03

		removed and error covariance for item 13 and 23 was added	$p = 0.111$				

CFI = comparative fit index, TLI = Tucker–Lewis fit index, AIC = Akaike information criterion, RMSEA = root mean square error of approximation, which indicates the amount of unexplained variance or residuals, *df* = degrees of freedom.

Comment: the CFA embedded in the SEM confirmed the reduced factor model (even though the model was overfitted), but there was no confirmation from the CFA for the full and modified models.

Test of the reliability and validity of the Nursing Activities Score in critical care nursing

Siv K. Stafseth* RN, CCN, MNSc, PhD Candidate

Dept. of Research and Development, Division of Emergencies and Critical Care, Oslo University Hospital and Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Norway

Tor Inge TØNNESSEN, MD, PhD, Professor

Division of Emergencies and Critical Care, Oslo University Hospital and Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Norway

E-mail: t.i.tonnessen@medisin.uio.no

Lien My DIEP, Statistician/Biostatistician, Oslo Centre for Biostatistics and Epidemiology Research Support Services, Oslo University Hospital, Oslo, Norway

E-mail: l.m.diep@medisin.uio.no

Lisbeth FAGERSTRÖM, RN, PhD, Professor, Faculty of Health Sciences, University College of Southeast Norway, Drammen, Norway; Professor, Faculty of Education and Welfare, Åbo Akademi University, Finland

E-mail: lisbeth.fagerstrom@usn.no

*Correspondence E-mail: sistaf@ous-hf.no , Adress: Box 4950 Nydalen, 0424 Oslo, Norway

Tel.: +47-22117950 or +47-99266142 Fax: +4723073714

Conflict of interest

No conflict of interest has been declared by the authors.

Acknowledgements

This study was financially supported by a grant (13/0028) from the Norwegian Nurses Organization.

Author contributions

S.K.S and L.F were responsible for the study conception and design. S.K.S performed the data collection. L.M.D performed the analyses and provided statistical expertise. S.K.S, L.M.D, L.F and T.I.T were responsible for drafting the manuscript.

Summary Statement

Why is this research needed?

- The tool Nursing Activities Score (NAS) is a promising approach for the allocation of nurses and resources in critical care.
- Testing the conceptual model for critical care nursing is intended to provide more valid information about NAS assessments and its usefulness in nursing practice.

What are the key findings?

- The inter-rater reliability was moderate between raters (on-duty caregiving and researchers using patient records).
- Three latent factors—relationship, prevention and treatment—were identified and confirmed the a priori conceptual model.
- We have achieved a novel understanding of assessment in critical care nursing.

How should the findings be used to influence policy/practice/research/education?

- Routine daily assessment of the NAS could be used for nurse allocation in critical care and in nursing management.
- We recommend additional tests of inter-rater reliability and validity in intensive care units using a larger sample size.