

Development and validation of nursing resource weights for the Belgian Nursing Minimum Dataset in general hospitals: A Delphi questionnaire survey approach

Walter Sermeus^a, Pierre Gillet^b, Daniel Gillain^c, Jan Grietens^d,
Nancy Laport^e, Dominik Michiels^d, Olivier Thonon^{e,f}, Guy Vanden Boer^d,
Pieter Van Herck^{a,*}, Koen Van den Heede^a

^a Centre for Health Services and Nursing Research, Catholic University Leuven, Belgium

^b University Hospital of Liège, Belgium

^c Department of Biostatistics, University Hospital of Liège, Belgium

^d Business Intelligence, University Hospitals Leuven, Belgium

^e Medical and Economical Information Department, University Hospital of Liège, Belgium

^f Department of Nursing, University Hospital of Liège, Belgium

Received 16 June 2008; received in revised form 28 August 2008; accepted 6 September 2008

Abstract

Background: Internationally, nursing is not well represented in hospital financing systems. In Belgium a nursing weight system exists to adjust budget allocation for differences in nurse staffing requirements, but there is a need for revision. Arguments include the availability of a nursing minimum dataset and the adverse consequences of the current historically based nursing weight system.

Objectives: The development and validation of nursing resource weights for the revised Belgium nursing minimum dataset (NMDS).

Design: Two independent cross sectional Delphi—surveys.

Setting and participants: A convenience sample of 222 head nurses from 69 Belgian hospitals participated in the cross sectional survey methods. To assess validity 112 patient case records from 61 nursing wards of 35 Belgian general hospitals representing general, surgical, pediatric, geriatric and intensive care were selected.

Methods: Nursing resource weights were constructed based on Delphi survey results by NMDSII intervention. The patient case Delphi survey results were used as the primary source for validation. A series of additional validation measures were calculated, based on the different patient classification systems. Finally, three validated nursing resource weighting systems were compared to the constructed NMDSII weighting system: the use of ‘Closon’, ‘Ghent’ and WIN weights.

Results: A coherent set of nursing resource weights was developed. The comparison of nurse resource weights, based on the survey per NMDS intervention versus the survey on patient cases, yielded high correlations: $r = 0.74$ to $r = 0.97$ ($p < 0.01$) between three case rating questions, as an indication of reliability in terms of internal consistency, and $r = 0.90$ ($p < 0.01$) between summed intervention weights and patient case weights, as an indication of criterion validity in terms of concurrent validity. Other concurrent validity measures based on summed intervention weights versus patient classification dependency weights showed a correlation ranging from $r = 0.14$ to $r = 0.74$. The correlation of summed intervention weights with the

* Corresponding author at: Centre for Health Services and Nursing Research, Catholic University Leuven, Kapucijnenvoer 35, 4th floor, 3000 Leuven, Belgium. Tel.: +32 16336982; fax: +32 16336970.

E-mail address: Pieter.vanherck@med.kuleuven.be (P. Van Herck).

Closon, Ghent and WIN weights ranged from $r = 0.93$ to $r = 0.96$ ($p < 0.01$), as a third indication of concurrent validity.

Conclusions: A system of valid nursing resource weights has been developed. The system should be further validated within an international context.

© 2008 Elsevier Ltd. All rights reserved.

Keywords: Budgets; Diagnosis-related groups; Economics; Hospitals; Nursing; Personnel staffing and scheduling

What is already known about the topic?

- In a limited number of countries hospital reimbursement is adjusted for nursing care, often lacking transparency and not taking into account appropriate nursing care needs at the patient level.
- General hospitals make local use of a diverse array of patient classification systems and nursing care workload measures to reallocate nursing staff. This requires a high amount of additional data collection efforts and investments.
- The lack of standardization obscures comparison and impedes national applications.

What this paper adds?

- This study provides a set of validated nursing resource weights as an input for nursing care adjustment of hospital reimbursement.
- The system is standardized, allowing for comparison between hospitals and nursing wards.

1. Introduction

Nursing is not well represented in hospital reimbursement, despite of being the biggest healthcare profession involved and despite of driving the most substantial cost component of the hospital budget. What is more, recently the impact of nurse staffing on patient safety has been highlighted in several studies (Aiken et al., 2002; Kane et al., 2007; Needleman et al., 2002). General staffing levels and reallocation of nursing personnel are likely to be influenced by a hospital's financial resources and how nursing is addressed in the reimbursement system.

In most countries there is no adjustment for a nursing care case mix in the hospital financing system, although there is a significant variability in nursing intensity and direct nursing costs between and within similar adult medical/surgical units (Welton et al., 2006a). Nursing intensity is the amount of direct and indirect patient care activity required to carry out the nursing function and the factors that have an impact on the level of work required to perform that activity (Morris et al., 2007). Under the umbrella of intensity of nursing care fall the concepts of nursing workload, patient acuity and time taken to administer patient care.

A recent literature review by Laport et al. (2008) found that countries like Denmark, Italy, the Netherlands, Portugal,

Spain, and the US reimburse nursing as a part of a general 'room and board' fee, without going further into specific nursing care characteristics. This means that differences in nursing hours delivered are not addressed. Welton et al. (2006b,c) found that the use of a fixed daily room rate in the US led to an underestimation of nursing costs by 32.2%.

Most countries that do adjust for nursing care, do this by applying an average nursing resource weight per Diagnosis Related Group (DRG). The system clusters patient stays into groups that are homogeneous in terms of clinical characteristics and resource use. A relative weight represents differences in hours and minutes of care on a pre specified level such as a patient level, DRG level, or nursing ward level. It summarizes resource consumption as a function of nursing time needs. For example, on a DRG level, a DRG with a relative weight of 4.0 is four times more nursing resource intensive than a DRG with a relative weight of 1.0. Examples of an average nursing resource weight per DRG can be found in the systems used in Australia, Canada, New Zealand and Switzerland (Laport et al., 2008). Nursing cost allocation studies provide the necessary information to develop such average nursing resource weights in most of these countries. The averaging method however does not take the variability of nursing intensity within DRG's into account. Including additional nursing care data, combined with DRG use, can improve the overall explanation of variance in length-of-stay, use of intensive care, hospital charges, hospital death and discharge to nursing home by respectively 29.3%, 28.3%, 27.5%, 146.4% and 92.4% (Welton and Halloran, 2005).

In Belgium a mixed system is used (Sermeus, 2006). All Patient Refined Diagnosis Related Groups (APR-DRG's v. 15.0) are only used in the volume component, by specifying length of stay, and are not related with nursing care characteristics. Nursing care adjustment takes place as a part of the hospital budget price calculation. The main price calculation system allocates relative parts of the national hospital operating budget to hospitals, independent of nursing costs. An average price per patient day is calculated, comparable to the 'room and board' approach, based on minimal nurse staffing ratios per type of nursing unit.

However, about 6.5% of the national budget is reserved for supplementing the average price with an additional budget allocation, partly based on nursing intervention data. The Nursing Minimum Data Set (NMDS) is used to classify inpatient days into 28 zones. Each zone is weighted based on actual staffing level (number and qualification level). The

additional budget allocation is based on the number of inpatient days per zone and their weight (Sermeus, 2006).

The need for change of this system is clear, for multiple reasons. Firstly, in 2006 NMDS was thoroughly updated towards a system of 79 nursing intervention items (Sermeus et al., 2005). NMDSII is the result of broad qualitative sector participation and a statistical quantitative reconfiguration of the system. It is up to date with current nursing practice. It is based on the Nursing Interventions Classification (NIC) as an international nursing intervention ‘language’. It is a much more accurate representation of what nursing care incorporates in all its different dimensions when compared to the previous version of NMDS, although it’s still a selection when compared to the 433 NIC interventions (Sermeus et al., 2006). Secondly, the current system lacks in one very important aspect: the NMDS – weighting as financial driver is based on a historically determined staffing levels per zone. There is no transparent relationship between these weights and the real need of nursing care.

Thirdly, in the current system there is no nursing care adjustment for other than medical, surgical or paediatric specialties, like geriatrics.

And finally, the nursing care adjustment is done independently of the APR-DRG medical case mix grouping, and is therefore also not related to length of stay in the Belgian system. The presently common evolution of length of stay reduction leads to a linear reduction of allocated resources, without accommodating for the likely probability that the remaining hospitalization days will lead to a higher level of nursing care intensity (Graf et al., 2003; Welton et al., 2006d).

As length of stay decreases, by postponing admission and accelerating discharge, patients are generally sicker during their stay. A combined higher patient acuity and care complexity leads to a higher need for nursing staff resources, both in number, knowledge and skill level. However, reversely, the greater demand for nurses and the upward pressure on wages has often led to a decrease of nursing to patient staffing ratios. In addition, next to increasing nursing intensity, the compression of days of stay creates greater variability in nursing care needs of hospitalized patients. Current cost accounting systems in most countries are lacking in identifying direct nursing costs on the level of an individual patient. Length of stay is often used as a surrogate cost marker, but doesn’t account for increased nursing resource needs and increased variability in nursing care needs. A potential solution, as an alternative to highly sophisticated cost accounting systems, is the introduction of nursing intensity billing instead of patient location billing, combined with length of stay data. The biggest barrier to overcome in most countries is the lack of data to identify actual and needed nursing resources. Often, when hospitals do collect relevant data, the second barrier concerns issues of validity and reliability of these measures. Belgium, disposing of NMDSII data, can be a frontrunner in improving nursing care billing and reimbursement policy. This paper presents

the first steps of creating such a nursing intensity reimbursement system by addressing nursing care needs, expressed as time and number of nurses, and the validity of the system. Knowledge and skill level are not addressed.

The study was aimed at redirecting the supplementary part of the Belgian financing system from actual towards justified staffing needs as a key criterion for resource allocation. This was part of the initial development and validation of a Belgian nursing resource weighting system, based on NMDSII. The system had to be applicable for all nursing units (e.g. including geriatrics).

2. Materials and methods

Multiple methods can be used to construct nursing resource weights. Nursing resource needs are commonly estimated in terms of nursing time requirements by time-and-motion studies, work sampling, and subjective evaluation. (Colombet et al., 2005) Patient dependency data can also be used in combination with one of these methods as an implicit algorithm that indicates the accuracy and consistency of nurses’ and observers’ data collection or subjective evaluation (Hurst, 2005; Hurst, 2008). One of the main challenges is capturing the whole patient care context and its inherent complexity of nursing interventions and inter-relations between interventions within the determination of nursing resource weights.

Time studies measure the allocation of nursing staff resources in the most objective and accurate way. However, it is difficult to fully comprehend nursing care by means of time studies as, by its nature, it is complex and multidimensional. Nurses often ‘multitask’ by doing more than one activity at a time. (Gran-moravec and Hughes, 2005) Traditional time studies are also expensive and time consuming. Both time-and-motion and work-sampling methods are subject to the Hawthorne effect, which consists of workers changing their habitual work pattern when observed (Colombet et al., 2005).

Self reporting is a low-cost means of quantifying time allocation by nursing care staff (Burke et al., 2000). Subjective evaluation usually takes the form of interviews or questionnaires (Colombet et al., 2005). These are subject to personal biases such as participants’ problems with memory, selective recall and correct question interpretation (Gran-moravec and Hughes, 2005).

Whatever the method, several authors stress the importance of linking allocation of nursing resources to a quality of care criterion, based on an intended result of nursing care norm setting. A transition from a historical and actual to a justified and required level of nursing care resources should be central to the methods applied.

Based on these considerations, with all methods having advantages and disadvantages, a multi-methods approach was followed. As part of subjective evaluation, a Delphi survey was used to construct a nursing resource weight per

NMDS intervention. To validate the nursing resource weights, a second subjective evaluation approach, using patient cases, was applied as part of a second Delphi survey. To counterbalance the disadvantages of these methods we also validated the results based on patient dependency data using patient classification systems, as observed by nurses while providing nursing care. And finally, alternative nursing resource weight systems as applied on the aggregated data set were used as source of validation. Some of these latter systems are based on time studies, using work sampling (Schoupe et al., 2007). All of these comparisons address concurrent validity, as a specific form of criterion related validity. Reliability in terms of internal consistency was assessed as part of the second Delphi survey. In attempting to establish criterion related validity the emphasis is on establishing the relationship between the instrument, i.e. NMDS weights, and other criteria. The instrument is valid if its scores correlate highly with the criteria (Polit and Hungler, 1999). Subjective patient case evaluation weights, patient dependency data weights, and time study weights are used as available criteria, with a differing degree of validity and reliability.

2.1. Delphi survey per NMDS intervention

Activity methods involve an assessment of the patient for the activities involved in providing the nursing care required. Each activity is been allocated a time to carry it out. The sum of a patient's required nursing activity times should supply the total time required to care for that patient in that shift or day (Hughes, 1999). Some reported difficulties are: perceptual differences of what constitutes an activity (Burke et al., 2000) and difficulties to assess time requirements for non-physical needs of patients (Hughes, 1999).

Since NMDSII is considered as a good descriptor of nursing interventions based on international grounding, nurse practice participation and statistical analysis, it is expected to be a useful tool to apply an activity method in nursing resource weighting.

Therefore, all NMDSII interventions were rated on nursing care needs separately, independently of any patient context, by 20 head nurses. They were selected randomly out of 222 candidates, working in 69 Belgian hospitals. There was a zero non-response. The rating was based on a survey distributed by e-mail. During a second round the raters had the opportunity to modify their estimates, based on the feedback of the first round's results. Seven of 20 raters chose to do so. The survey was aimed at the following question: How much time does a nursing team spend on average on . . . [specific NMDSII nursing intervention] . . . in caring for a typical patient to ensure quality of care? This question was posed for each of the 79 NMDSII items and was further subdivided by NMDSII sub item categories. To account for the potential problem as mentioned above concerning what constitutes a nursing intervention, the main question per nursing intervention is embedded into the

official NMDSII registration manual as available in August 2006 (Belgian Federal Public Service Health, Food Chain Safety and Environment, 2006). This manual provided a clear definition incorporating all relevant modalities of execution. Most interventions could be assessed directly on a 24 h basis. For example: how much nursing time is needed for adequate hygienic care with partial assistance can be estimated for 24 h. But e.g. blood administration depends on the frequency of administrations and the number of units per administration. Therefore this item was assessed on a blood unit level. Other examples are 'bladder catheterization', 'extra ward patient transport', 'administration of subcutaneous, intradermal or intramuscular medication', 'venous blood sampling', etc. As such, 21 of the 79 NMDSII interventions were frequency or quantity based and assessed accordingly (see Appendix A). The other NMDSII items are conceived as a nominal or ordinal scale. Adding the relative points of the 24 h based items and multiplying with frequency where necessary gives the total nursing resource weight for one patient day.

2.2. Validation

2.2.1. Development of patient cases

One hundred and twelve real patient cases were written, based on patient records from 61 nursing wards in 35 hospitals. This represents a convenience sample of about one fourth of all Belgian hospitals. Criteria of representativeness such as hospital size, teaching status, and region were applied. Patient records in combination with additional information from the head nurse, involved in the specific care delivered, were the basis for case construction. The head nurses selected had a minimum of five years' experience as nurses and a one year's experience as head nurses. All information was obtained by way of nursing ward visits. All data were gathered and all cases were written using a standard format. A case describes the whole of nursing care delivered for a specific patient, during one day of stay (24 h). Three forms of information were gathered: (1) information about the patient day of stay and care given, (2) information about the context of care and (3) registered NMDSI and NMDSII data.

The 24 h of nursing care were described in detail separately for the morning, evening and night shift. A chronological description of all nursing interventions during the patient day of stay was constructed. Lab and other results of clinical investigations were included since these can imply a substantive amount of additional nursing care.

Patient characteristics such as age, gender, etc., which are of influence on nursing care, were given. The same holds true for example for the medication taken by the patient at home, which is continued during the hospital stay. The medical history of the patient gave additional information about comorbidities that necessitate additional care. Other important patient characteristics were also included. Elements of the whole hospital stay included the reason of admission, the

preliminary length of stay, the care rendered previously and the treatment stage which is now considered (e.g. first day post operative).

Forty medical-surgical cases, 25 paediatric cases, 22 intensive care unit (ICU) cases and 25 geriatric cases were developed. Cases differed in length from three up to nine pages. After case construction, an additional feedback of the involved nursing staff was sought to confirm that the case was a genuine description of care as it was rendered in practice.

2.2.2. Delphi survey on patient cases

These cases were rated on nursing care needs by 202 head nurses from 69 hospitals. This sampling was done independently of the sampling for case construction. However, the same criteria were applied. Eighty-seven head nurses rated medical-surgical cases, 40 rated paediatric cases, 36 rated ICU cases and 39 rated geriatric cases. Each respondent assessed eight to twelve cases. This enabled an average of 10 ratings per case. The cases were randomly assigned to the respondents, stratified by specialty.

Cases were distributed to the respondents on an individual basis by e-mail. The rating itself took place using a web based survey. All cases were rated during two consecutive rounds, as part of an adapted Delphi approach. More exactly, all respondents received feedback per case. The feedback enabled the respondent to change his or her original rating or not. It is an important method for achieving consensus on issues where none previously existed (Keeney et al., 2006). A Delphi approach is a validated technique to obtain the most reliable consensus of opinion (Campbell et al., 2003). Recommendations as presented by Hasson et al. (2000) were followed.

For round I a time frame of one month and for round II a time frame of two weeks was provided. Reminder e-mails were sent to minimize non-response. Additional individual respondent phone calls were a further means of stimulation. The response during round I is equal to 92%. During the second round 20% of the participants chose to revise their estimates, based on the feedback received. None of the participants informed the research team to choose to withdraw from the study. The variation in response between specialties was minimal.

Three main questions were formulated concerning nursing care needs: 'How much is the required nurse time needed for care delivery as described in the specific patient case to ensure quality of care?', 'Taking into account current level of ward staffing, how many patients with this nursing care profile can one nurse care for?', 'Suppose there would be no limitations on ward staffing, how many patients with this nursing care profile can one nurse care for?'. The different questions allowed evaluating internal consistency of the rating procedure.

The combination of NMDSII items in each patient case makes it possible to sum and combine the constructed nursing resource weights. This makes the findings of both methods directly comparable as a measure of criterion validity. Items that are frequency based in NMDSII were

taken into account by multiplying their weight with their frequency before summation.

2.2.3. Alternative patient classification systems and nursing resource weight systems

A number of alternative patient classification systems were used on a patient case level, to evaluate concurrent validity: the Therapeutic Intervention Scoring System (TISS) for ICU cases (Reis et al., 1996), NARVEL for paediatric cases (Dierickx and Sermeus, 1985), San Joaquin for medical-surgical cases (Hadley et al., 1978) and Autonomie Gérontologique Groupe Iso-Ressources (AGGIR) for geriatric cases (Benaim et al., 2005).

Two validated nursing resource weight systems, based on NMDSI were compared to the constructed NMDSII weighting system: the use of the 'Closon' and 'Ghent' NMDS weighting system. (Closon, 1991; Ghent, submitted for publication)

Finally, the Workload Indicator for Nursing (WIN) weights developed as part of a NMDSII based staffing allocation tool, were compared to the constructed nursing resource weights. These weights were constructed specifically for general, surgical and geriatric care by a multi moment work sampling time measurement approach. Schoupe et al. (2007) used external observation during two morning and two evening shifts in ten general Belgian hospitals, four nursing wards each. These staff allocation weights were, as part of their study, also assessed on face validity by 226 registered nurses.

The 'Closon', 'Ghent' and 'WIN' weights were compared to the developed NMDS II weighting system based on an existing dataset. NMDS-II information was collected during the pilot study of the actualisation of NMDS during 2003–2004. The dataset consists of 66.827 inpatient days from 59 Belgian hospitals concerning 18.148 patient stays. A balanced sample was obtained for the following medical specialties: geriatrics, paediatrics, intensive care, chronic illness, maternal services, general internal medicine and general surgical procedures.

2.2.4. Statistics

Sample sizes were assessed using power analysis based on Altman's nomogram (1982) and Lehr's formula (1992). A power level of 0.80 to 0.90 is maintained.

The random allocation algorithm was written in R, a statistical computing platform, developed as part of the R project (see <http://www.r-project.org>). Ratings of the Delphi study per NMDS intervention were statistically analysed in the following way: Descriptive measures, including the mean, trimmed mean, median, standard deviation, and range, were calculated for each of the NMDS items.

For calculating average time estimates per nursing intervention and patient case, robust measures were used. They are less sensitive to outliers, but don't exclude them from analysis. Normality checks were applied, including the Shapiro Wilk test. If the Shapiro-Wilk normality test showed

a significant deviation from normality ($p < 0.05$), the Huber robust mean was selected. Else the mean was considered a fair measure. Time estimates per NMDSII intervention were transformed into relative weights by dividing them by five. This approach is identical to the one applied in other internationally existing nursing resource weight systems, such as the Programme Recherche Nursing (P.R.N.) system, used in Luxembourg (Tilquin et al., 1989; O'Brien-Pallas et al., 1992).

Statistical analysis of the ratings of the Delphi study on patient cases, in terms of descriptives calculation, normality checks and the selection of robust measures, corresponded with the method described above.

To analyse concurrent validity and internal consistency Pearson correlations were calculated. Non-parametrical alternatives such as Spearman's rho and Kendall Tau showed negligible differences.

3. Results

3.1. Delphi survey per NMDS intervention

The rating of nursing care needs based on separate nursing interventions varied from five minutes for nursing intervention 'capillary blood sampling', to be multiplied with the sampling frequency, up to 240 min for 'relaxation care in preparation of child birth'. The mean time estimate across not frequency based items was 39.2 min per item, during 24 h of care. For frequency based items this was 22.3 min per intervention. A complete overview of nursing resource weights per NMDSII item is presented in Appendix A. All NMDS II interventions are listed. Time estimates are expressed as relative points. Specific care modalities are specified where necessary.

Low weighted items include 'urinary elimination follow-up', 'fecal elimination follow-up' and 'contact with other institutions'. Highly weighted items comprise 'dialysis regulation' items, 'artificial ventilation' items, 'cardio circulation support by mechanical aids' items, and most items within the 'birth care' class in domain V.

3.2. Validation

3.2.1. Delphi survey on patient cases

The rating of nursing care needs based on patient cases varied from a minimum of 15.08 min up to 356.87 min, with an average of 91.74 min. Fig. 1 presents an overview of case ratings per specialty. Intensive care and paediatric care require higher amounts of nursing care, compared to other specialties.

Internal consistency correlations between the three case rating questions ranged from $r = 0.76$ to $r = 0.97$ ($p < 0.01$).

There is a high correlation ($r = 0.90$, $p < 0.01$) between the sum of nursing care weights per intervention and the

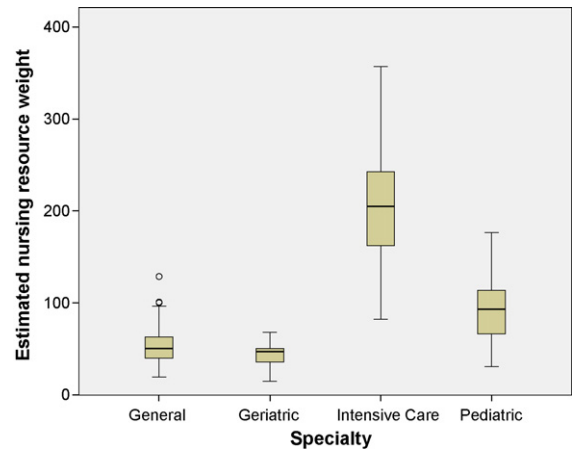


Fig. 1. Overview of case ratings per specialty. Legend: X-axis = patient cases ordered by specialty, Y-axis = estimated nursing resource weight, expressed in minutes.

patient case rating of nursing care needs, as a criterion related validity measure, based on subjective evaluation (Fig. 2).

3.2.2. Other patient classification systems and nursing resource weight systems

There was a high correlation with TISS ($r = 0.74$, $p < 0.01$) for ICU cases, a medium correlation with AGGIR ($r = 0.41$, $p < 0.01$) for geriatric cases, a low correlation with San Joaquin ($r = 0.27$, $p < 0.05$) for general and surgical cases and a very low correlation with NARVEL ($r = 0.14$, ns) for pediatric cases.

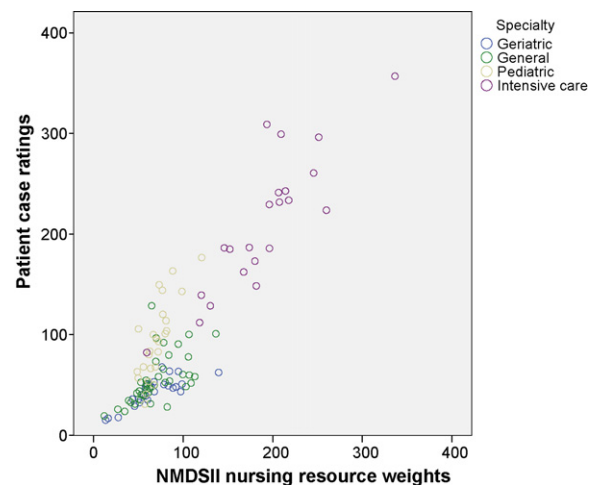


Fig. 2. Visual comparison between patient case ratings and NMDSII nursing resource weights. Legend: X-axis = NMDSII rating based on the summation of robust selected intervention ratings, Y-axis = patient case rating, based on the robust selected mean for 24 h of nursing care, expressed in minutes.

The correlation between the newly developed nursing care weights for NMDS-II and the Closon, Ghent and WIN weights for NMDS is respectively equal to $r=0.93$ ($p < 0.01$), $r=0.95$ ($p < 0.01$) and $r=0.96$ ($p < 0.01$).

4. Discussion

The study results in validated nursing resource weights per item of the Belgian NMDSII. This means that the system is currently up to date with nursing practice, based on NIC as an international nursing language, and that it incorporates a broader representation of professional nursing care. More importantly, it justifies nursing resource needs in a relative way, as opposed to using historical staffing levels. Compared to the current complex weighting system of the 28 nursing care zones, the new tool is likely more easily for nurses to understand and to relate to their own practice. The notion of nursing resource weights, explicitly based on required time estimates, and the role of nurse experts in its development support this hypothesis. Additionally, the perceptions of nurses with regard to what they consider appropriate resource levels is built into the instrument's development by relating time allocation to a perceived level of sufficient quality of care as part of the Delphi survey questions. Finally, the developed nursing resource weights serve a broader set of medical specialities.

The innovative method of using both patient cases and nursing intervention estimates enables a strong validation of the system's accuracy. It was a remarkable finding that although interrater variability in both approaches is high in determining nursing resource needs in absolute terms, the criterion validity in relative terms was very high. Most of the other criterion validity measures support the high level of validity of the weighting system. Here also the correlation between relative estimation and other nursing intervention weighting systems, such as Ghent, Closon, and WIN weights was remarkably high.

Furthermore, a very strong relationship existed between the estimated time and the TISS patient classification system regarding intensive care. This confirms previous research. The geriatric AGGIR – estimated time relationship was also strong. However the relationship with San Joaquin for general care was weak and with NARVEL for paediatric care is the relationship weak and not-significant. NARVEL is an older, internally developed patient classification system. The relationship of estimated time in general patient cases with San Joaquin is hampered by the limited patient case coverage of nursing workload in categories I and IV. This is partly due to the very high criteria in San Joaquin which define its category IV, e.g. constant monitoring and observation of patient's medical condition. At the other end of the continuum patients within category I are almost totally independent for activities of daily living within the San Joaquin classification. Both situations are rare on a surgical or general care ward. Hence the lack of sufficient variability

of the patient case nursing care needs within the San Joaquin classification. This impeded accurate relationship estimation by using correlation measures.

In general the validity of the nursing resource weighting system has been confirmed by a number of surprisingly strong measures.

The use of the nursing resource intensity weights is not restricted to hospital financing. Our findings have also high potential to guide nurse staffing decisions at both the level of the Ministry of public Health as the hospital level. Next to the registration of 79 nursing interventions (and the according nursing resource weights) also the number and educational level of the employed nursing staff are registered. This offers the possibility to develop a nationwide benchmarking tool in which actual staffing patterns are compared with required staffing (based on the scored B-NMDS nursing interventions and the according nursing intensity weights). The Ministry of public Health can use this information to monitor nurse staffing levels over time. When drastic changes in the nursing intensity are observed, actions to adjust nurse staffing levels can be undertaken. The comparison of actual and required staffing levels can also guide hospital managers allocate nurse staffing budgets within their institutions.

The nursing resource weights can also contribute to the research domain of nurse staffing and patient safety. After all, it is widely recognized that nurse staffing levels should be adjusted for differences in nursing intensity (Needleman et al., 2007). However, this is seldom done due to the absence of uniformly collected data about nursing intensity. The linkage of the B-NMDSII to the Belgian Hospital Discharge Dataset (from which patient outcomes can be derived) offers the possibility to fill this gap in the nurse staffing and patient safety research.

This study has resulted in a set of validated nursing resource weights on the level of individual patients' days. Currently a study is being prepared to make the transfer towards the complete patients' stay level. This will provide in the identification of individual patient nursing intensity profiles, taking into account the patient's evolution through his length of stay. Also, a model will be developed which counters the possibility to game the system.

In long term the results of the current study contribute to two evolutions:

Firstly, the evolution towards a combined medical and nursing case mix data system, in which medical DRG's and Nursing Related Groups (NRG's) are positioned independently next to each other, forming a matrix to identify an individual patient's hospital health care needs and resource needs status. In such a scenario the sources of variability are more comprehensively addressed. It represents nursing as a resource for meeting performance targets rather than exclusively as a cost center.

Secondly, the evolution towards linking payment to performance. An adequate nursing resource allocation system cannot only contribute to the prevention of

adverse events, and therefore saving costs and reducing length of stay. Moreover, it is a valuable tool within a Pay For Performance (P4P) context. In fact, what the system actually does is linking structural resources such as nursing personnel to processes of care on an individual patient's level, as identified by NMDSII. Nursing interventions represent processes of care. These can be translated into evidence based process indicators, based on their relationship with nurse sensitive patient outcomes. The direct relationship between nursing resource levels and patient outcomes can also be modelled as part of P4P, since both types of variables are present in the dataset on the patient level. As such a P4P system based on nursing performance measures can be developed. Reversely, the impact of P4P on nursing labour can be monitored (Bodrock and Mion, 2008; Clarke et al., 2008).

Further research is needed to study the criterion related validity of NMDS II weights, using alternative general care and paediatric care patient classification systems. In addition, future research could replicate the patient case estimation approach within an international context. This would clarify international and cultural differences in nursing practice and the associated resource utilization.

5. Conclusions

A system of valid nursing resource weights has been developed to assist in fair and justified hospital budget allocation in Belgium. The system is based on the revised

Nursing Minimum Dataset, which is internationally grounded, up to date, and supported by both quantitative and qualitative research. To integrate the system into APR DRG further modeling is required, based on a national representative NMDSII sample. National NMDSII data collection has started in March 2008.

Funding

This study was funded by the Belgian Health Care Knowledge Centre.

Ethical Approval

Ethical approval was given by all participating general hospitals. There was no direct patient intervention in this study.

Conflict of Interest

No conflicts of interest to declare.

Acknowledgments

The study was funded by the Belgian Healthcare Knowledge Centre, Brussels, Belgium under grant HSR 2005-19, D/2006/10.273/06. The authors also thank Sigrid Aerts, Natasha Deprez, Nicolette Van Braekel and Ann Janssens for their valuable contribution.

Appendix A

Relative nursing resource weights per NMDSII item

Domain I: Care for elementary physiological functions

Class A Support of activities and physical movement

item	Care description	Specific care modality	Nursing care weight
A100	Structured physical exercises		12
Class B Care for elimination			
B100_1	Elimination child care	Toilet trained child day and night time	10
B100_2		Toilet trained child night time	8
B100_3		Non-toilet trained child day and night time	6
B210	Urinary elimination follow-up		2
B220	Support of urinary continent patient		6
B230	Care for the urinary incontinent patient		10
B240	Care for urinary stoma		7
B250	Care for urinary catheter		5
B300	Bladder catheterization	× frequency	4
B410	Fecal elimination follow-up		2
B420	Support of fecal continent patient		5
B430	Care for the fecal incontinent patient		7
B440	Care for fecal stoma or pouch		5
B500	Constipation prevention or treatment		4
B600	Elimination care education		6

Appendix A (Continued)

Domain I: Care for elementary physiological functions

Class A Support of activities and physical movement

item	Care description	Specific care modality	Nursing care weight
Class C Care for patient mobility			
C110	24 h bedridden patient installation		7
C120_1	Non-24 h bedridden patient installation	Supervision	6
C120_2		Partial assistance	6
C120_3		Complete assistance	8
C200_1	Support of intraward patient mobility	Supervision	5
C200_2		Complete assistance	5
C300	Extra ward patient transport	× frequency	3
C400	Care for traction		6

Class D Care for feeding

D110_1	Support of feeding (patient room)	Supervision	3
D110_2		Partial assistance	6
D110_3		Complete assistance	13
D110_4		Specific feeding needs	13
D120_1	Support of feeding (dining room)	Supervision	3
D120_2		Partial assistance	5
D120_3		Complete assistance	9
D120_4		Specific feeding needs	15
D130	24 h sober patient care		2
D200	Care for child bottle and breast feeding		15
D300_1	Administration of gastro enteral tube feeding	Gastric tube	8
D300_2		Stoma	8
D400	Administration of Total Parenteral Nutrition		4

Class E Comfort support

E100	Symptoms management pain		5
E200	Symptoms management nausea and emesis		4
E300	Symptoms management tiredness		3
E400	Symptoms management sedation		5

Class F Personal care support

F110_1	Hygienic care at lavatory, bed or incubator	Supervision	3
F110_2		Partial assistance	4
F110_3		Complete assistance	6
F110_4		Permanent presence and guidance	5
F120_1	Hygienic care in bath or shower	Supervision	3
F120_2		Partial assistance	4
F120_3		Complete assistance	6
F120_4		Permanent presence and guidance	5
F200	Hygienic care education and training		5
F300	Support of day clothing		3
F400	Support of self image		4
F500	Special mouth care		6

Domain II: Care for complex physiological functions

Class G Care for base acid and electrolyte balance

G100_1	Regulation of hydration and nutritional balance	Hydration and nutrition 1/day	4
G100_2		In out 1/day	4
G100_3		In out 2–6/day	5
G100_4		In out 7–12/day	8
G100_5		In out >12/day	13
G100_6		In out electronic	4
G200	Care for evacuating gastric tube		5
G300_1	Regulation of glycemic balance	Without education	6
G300_2		With education	9
G400	Regulation of blood balance		6

Appendix A (Continued)

Domain I: Care for elementary physiological functions

Class A Support of activities and physical movement

item	Care description	Specific care modality	Nursing care weight
G500_1	Dialysis regulation	Peritoneal dialysis	21
G500_2		Discontinual hemodialysis	19
G500_3		Continual hemodialysis	15
Class H Care for drug use			
H100	Administration of SC, ID or IM medication	× frequency	2
H200	Number of different administered IV drugs	× frequency	2
H300	Most frequently administered IV drug	Not applicable	/
H400	Administration of aerosol, puff or oxygen tent medication	× frequency	2
H500	Administration of vaginal medication	× frequency	3
Class I Neurological care			
I100	Neurological function follow-up using GCS	× frequency	2
I200_1	Pressure monitoring of intracranial fluid	Without drainage	5
I200_2		With drainage	6
Class K Care for breathing			
K100	Aspiration of airways		7
K200_1	Supportive means of breathing function	Mask, goggles, nasal tube, oxygen tent	3
K200_2		Endo tracheal tube, larynx mask	7
K200_3		Trachea canulae	12
K300_1	Artificial ventilation	Regular	23
K300_2		Special type	21
Class L Skin and wound care			
L100	Supervision of wound dressing, materials and near skin		6
L200	Care for sutures and inward materials points	× frequency	4
L300	Simple care for open wound	× frequency	4
L400	Complex care for open wound	× frequency	7
L500	Care for dermatological lesions	× frequency	4
Class M Regulation of temperature			
M100	Thermal regulation follow-up		11
Class N Care for tissue circulation			
N100	Administration of blood and blood components	× frequency	4
N200	Artificial entry point supervision and/or care	× frequency	2
N300	Venous blood sampling	× frequency	2
N400	Arterial blood sampling	× frequency	2
N500	Capillary blood sampling	× frequency	1
N600	Cardio circulation support by electrical aids		9
N700_1	Cardio circulation support by mechanical aids	Internal assistance device	18
N700_2		External assistance device	19
N700_3		ECMO	19
Domain III: Behavioral care			
Class O Behavioral therapy			
O100_1	Activity support	Group	12
O100_2		Individual	5
O200	Behavioral dysfunction care		13
Class P Cognitive therapy			
P100_1	Care for patients with reduced cognitive abilities	Occasional	11
P100_2		Standard plan	11
Class Q Communication support			
Q100	Support of communication problems		7
Class R Problem handling support			
R110	Basic emotional support		6
R120	Specific emotional support		11

Appendix A (Continued)

Domain I: Care for elementary physiological functions

Class A Support of activities and physical movement

item	Care description	Specific care modality	Nursing care weight
R130	Emotional crisis support		10
Class S Patient education			
S100_1	Specific education	Occasional	5
S100_2		Standard plan	6
S200_1	Pre investigation or surgical procedure education	Occasional	5
S200_2		Standard plan	6
Domain IV: Safety care			
Class V Risk management			
V100	Pressure ulcer prevention by means of dynamic materials		5
V200	Pressure ulcer prevention by repositioning	× frequency	2
V300	Continual monitoring of vital parameters		11
V400	Discontinual monitoring of vital parameters		4
V500	Tissue or excremental sampling	× frequency	2
V600_1	Isolation care	Minimal 2 elements of {apron, gloves, mask, garbage handling}	11
V600_2		Minimal 3 elements and separate patient room	9
V700	Protective measures with des orientation		7
Domain V: Family care			
Class W Birth care			
W100	Relaxation care in preparation of child birth		48
W200	Ante partum care: monitoring uterine activity	× frequency	13
W300	Child birth delivery		17
W400	Post partum follow-up	× frequency	3
W500	Kangaroo care		8
Class X Family care			
X100	Rooming in of family or significant others		4
Domain VI: Healthcare management			
Class Y Care counseling			
Y100	Cultural brokerage		4
Y200	Anamnesis at intake		4
Class Z Management of care provisions and information			
Z100	Functional, mental, psychosocial assessment		6
Z200	Physician support in direct medical care	× frequency	7
Z300	Multidisciplinary conference		6
Z400	Contact with other institutions		2

References

- Aiken, L.H., Clarke, S.P., Sloane, D.M., 2002. Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction. *Journal of the American Medical Association* 288 (16), 1987–1993.
- Altman, D.G., 1982. How large a sample? In: Gore, S.M., Altman, D.G. (Eds.), *Statistics in Practice*. Blackwell Publishing Ltd, BMA, London.
- Belgian Federal Public Service Health, Food Chain Safety and Environment. Dienst Datamanagement. Codeerhandleiding Minimale Ziekenhuisgegevens. Officiële versie 1.2. Mei 2006. Retrieved June, 2006 from: https://portal.health.fgov.be/portal/page?_pageid=56,6450388&_dad=portal&_schema=PORTAL.
- Benaim, C., Froger, J., Compan, B., Pelissier, J., 2005. The assessment of autonomy in elderly people. *Annales de Readaptation et de Medecine Physique* 48 (6), 336–340.
- Bodrock, J.A., Mion, L.C., 2008. Pay for performance in hospitals: Implications for nurses and nursing care. *Quality Management in Health Care* 17 (2), 102–111.
- Burke, T.A., McKee, J.R., Wilson, H.C., Donahue, R., Batenhorst, A.S., Pathak, D.S., 2000. A comparison of timeand-motion and self-reporting methods of work measurement. *Journal of Nursing Administration* 30, 118–125.
- Campbell, S.M., Braspenning, J., Hutchinson, A., Marshall, M.N., 2003. Research methods used in developing and applying quality indicators in primary care. *British Medical Journal* 326, 816–819.

- Clarke, S.P., Raphael, C., Disch, J., 2008. Challenges and directions for nursing in the pay-for-performance movement. *Policy, Politics, and Nursing Practice* 9 (2), 127–134.
- Closon, M.C., 1991. Le financement des hôpitaux en fonction de la structure des pathologies. Thèse de doctorat. Ecole de Santé Publique UCL.
- Colombet, I., Aguirre-Junco, A., Zunino, S., Jaulent, M.C., Leneveu, L., Chatellier, G., 2005. Electronic implementation of guidelines in the EsPeR system: A knowledge specification method. *International Journal of Medical Informatics* 74, 597–604.
- Dierickx, H., Sermeus, W., 1985. Patientenclassificatie. Zorgenbehoefte als basis van staffing. ACCO, Leuven/Amersfoort.
- Ghent, internal document Centre for Health Services and Nursing Research, Catholic University Leuven, Belgium, submitted for publication.
- Graf, C.M., Millar, S., Feilteau, C., Coakley, P.J., Erickson, J.I., 2003. Patients' needs for nursing care: beyond staffing ratios. *Journal of Nursing Administration* 33 (2), 76–81.
- Gran-moravec, M., Hughes, C., 2005. Nursing time allocation and other considerations for staffing. *Nursing and Health Sciences* 7, 126–133.
- Hadley, J., Murphy, L., Dunlap, L., Williams, M., Mcathy, M., 1978. *Methods for studying. Nurse staffing in a patient unit.* U.S. Department of Health, Education and Welfare. Publication nr. HRA 78-3. 226pp.
- Hasson, F., Keeney, S., McKenna, H., 2000. Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing* 32 (4), 1008–1015.
- Hughes, M., 1999. Nursing workload: an unquantifiable entity. *Journal of Nursing Management* 7, 317–323.
- Hurst, K., 2005. Relationships between patient dependency, nursing workload and quality. *International Journal of Nursing Studies* 42, 75–84.
- Hurst, K., 2008. UK ward design, patient dependency, nursing workload, staffing and quality- an observational study. *International Journal of Nursing Studies* 45, 370–381.
- Kane, R.L., Shamliyan, T., Mueller, C., Duval, S., Wilt, T.J., 2007. Nurse staffing and quality of patient care. Agency for Healthcare Research and Quality. Publication Nr. 07-E005. Evidence Report Nr. 151, 539 pp.
- Keeney, S., Hasson, F., McKenna, H., 2006. Consulting the oracle: ten lessons from using the Delphi technique in nursing research. *Journal of Advanced Nursing* 53 (2), 205–212.
- Laport, N., Sermeus, W., Vanden Boer, G., Van Herck, P., 2008. Adjusting for nursing care casemix in hospital reimbursement. A review of international practice. *Policy, Politics & Nursing Practice* 9 (2), 94–102.
- Lehr, R., 1992. Sixteen s squared over d squared: a relation for crude sample size estimates. *Statistics in Medicine* 11, 1099–1102.
- Morris, R., MacNeela, P., Scott, A., Treacy, P., Hyde, A., 2007. Reconsidering the conceptualization of nursing workload: literature review. *Journal of Advanced Nursing* 57 (5), 463–471.
- Needleman, J., Buerhaus, P., Mattke, S., Stewart, M., Zelevinsky, K., 2002. Nurse-staffing levels and the quality of care in hospitals. *New England Journal of Medicine* 346, 1715–1722.
- Needleman, J., Kurtzman, E.T., Kizer, K.W., 2007. Performance measurement of nursing care: state of the science and the current consensus. *Medical Care Research and Review* 64 (Suppl. 2), 10S–43S.
- O'Brien-Pallas, L.L., Cockerill, R., Leatt, P., 1992. Different systems, different costs? An examination of the comparability of workload measurement systems. *Journal of Nursing Administration* 22, 17–22.
- Polit, D.F., Hungler, B.P., 1999. *Nursing Research. Principles and Methods*, sixth ed. Lippincott, Philadelphia.
- Reis, M., 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.
- Schoupe, L., Defloor, T., Gobert, M., Van Goubbergen, D., 2007. *Workload Indicator for Nursing.* Federaal Wetenschapsbeleid, Eindrapport, 322 pp.
- Sermeus, W., Delesie, L., Michiels, D., Van den Heede, K., Van Herck, P., Van Landuyt, J., 2006. Actualisatie van de minimale verpleegkundige gegevens. Ministerie van Sociale Zaken, Volksgezondheid en Leefmilieu, Eindrapport, Brussel.
- Sermeus, W., 2006. *De Belgische ziekenhuisfinanciering ontcijferd.* ACCO, Leuven.
- Sermeus, W., Van den Heede, K., Michiels, D., Delesie, L., Thonon, O., Van Boven, C., Codognotto, J., Gillet, P., 2005. Revising the Belgian Nursing Minimum Dataset: From concept to implementation. *International Journal of Medical Informatics* 74, 946–951.
- Tilquin, C., et al., 1989. PRN 87—Measuring the Level of Nursing Care Required. E.R.O.S., Montréal, 175 pp.
- Welton, J., Halloran, E., 2005. Nursing diagnoses, diagnosis related group, and hospital outcomes. *Journal of Nursing Administration* 35, 541–549.
- Welton, J., Fischer, M., DeGrace, S., Zone-Smith, L., 2006a. Hospital nursing costs, billings and reimbursement. *Nursing Economics* 24, 239–245.
- Welton, J., Fischer, M., DeGrace, S., Zone-Smith, L., 2006d. Nursing intensity billing. *Journal of Nursing Administration* 36 (4), 1–9.
- Welton, J., Zone-Smith, L., Fischer, M., 2006b. Adjustment of inpatient care reimbursement for nursing intensity. *Policy Politics & Nursing Practice* 7, 270–280.
- Welton, J., Fischer, M., DeGrace, S., Zone-Smith, L., 2006c. Nursing intensity billing. *Journal of Nursing Administration* 36, 181–188.