MEDINFO 2001
V. Patel et al. (Eds)
Amsterdam: IOS Press

Exploiting the Nursing Minimum Data Set for the Netherlands

William T.F. Goossen

Senior researcher and consultant, Acquest consulting, the Netherlands

Abstract

The purpose of this overview is to examine and illustrate the feasible options for the statistical analysis of the Nursing Minimum Data Set for the Netherlands.
Distinct purposes for information use and for the presentation of information require different approaches for data collection and data analysis. Several explorative studies are discussed, illustrating their goals, data collection methods, data analysis and results.
The overview ends with general recommendations for data collection and analysis of nursing minimum data sets.

Keywords

nursing minimum data set, statistical analysis, ridit analysis, data collection, graphical representations

Introduction

A Nursing Minimum Data Set provides a minimum number of data that are maximally useful to visualize and legitimate nurses' contribution to healthcare and several other purposes [1, 2, 3]. The Nursing Minimum Data Set for the Netherlands (NMDSN) describes for different patient populations the diversity and complexity of patients and their problems, and the variability of patient related nursing care activities [4].

The NMDSN consists of 145 single variables and its content validity is supported by its consistency with the literature, findings from practice documentation, and the judgements of potential users [4, 5]. The categories and data items of the NMDSN include five institution related items, six patient demographic items, seven medical conditions items, and ten items about the use of the nursing process.

In addition, twenty-four nursing diagnoses include information needs, communication, uncertainty, fear, stress, ADL and self care, vitals, pain, risks, elimination, breathing, food intake, bedsores, and sleep, among others. The set contains thirty-two nursing interventions including instructing, supporting, assisting with feeding, breathing, ADL, self care, disorientation, mouth care, prevention of bedsores and risks, checking vital signs, giving medications, infusion, wound care, sampling, preparation and aftercare for tests and operations, among others [4].

Four outcomes of nursing care are included: patient satisfaction with information, pain alleviation, care in general, and the number of patient falls. Finally, three additional NMDSN items measure the complexity of nursing care [4]. To apply the NMDSN, a special form and instruction were developed for manual completion [5].

The NMDSN has been tested in 9 hospitals on 15 wards [5], and its validity and reliability have been explored on an intensive care ward, a nursing home and a residential home [6]. In these tests, the methodology of the Belgian Nursing Minimum Data Set has been applied [3, 7, 8].

Since the major goal of the NMDSN is to allow multiple use of the same data, it is important to identify the purpose of each use, determine appropriate designs and data collection methods, and apply adequate methods to analyze the data accordingly. The question guiding this exploration is as follows:
What are feasible purposes for data use and what are adequate methods for analysis of the nursing data in the NMDSN?

An overview of purposes, methods and results

The first purpose of the NMDSN is to visualize differences in patient populations and nursing activities for management and policy makers. To test the NMDSN for this purpose, data are collected on 15 wards every day for all patients during 5 succeeding days [5]. The sample includes 686 individual patients, while for the data analysis their 2090 patient days in the hospital are used, and analyzed cross-sectional. Frequencies of nursing diagnoses, nursing interventions and results of care per patient day are calculated, and are illustrated in Table 1. This methodology is consistent with the Belgian approach, where such data, after additional statistical analysis, are used to determine part of the national budget for nursing [3, 8].

This kind of reporting frequencies from the NMDSN is useful to get the broad picture of what happens on a ward. Additional ways of data presentation include per age

Table 1 – Examples of frequencies in the test of the NMDSN

NMDSN item / Ward	Ward A	Ward B	Ward C	Ward D
Patient fears	25	33	21	77
Bedsores prevention	44	32	4	25
Level assistance with mobility: a] independent b] partly c] complete	a] 79 b] 58 c] 38	a] 99 b] 62 c] 29	a] 58 b] 27 c] 22	a] 106 b] 76 c] 62

category, per medical diagnosis, or for other relevant groupings of patients. Managerial and policy information can thus be generated from the NMDSN.

A second purpose is to present NMDSN data in a longitudinal design. De Klerk has done so to determine what kind of nursing data are necessary for inclusion in a 'care product' for hernia patients [9]. The care product is the Dutch equivalent of Diagnosis Related Groups. De Klerk daily collected the NMDSN items for 7 patients with hernia from admission to discharge. She reported, with the individual patient as unit of analysis, the total number of nursing diagnoses and nursing interventions and suggested including the nursing variables with the highest frequencies for inclusion in the care product [9]. See table 2 for examples of the frequencies from this study.

Table 2 – Frequencies on individual patient level per episode

Frequencies per Patient	Pt. A	Pt. B	Pt. C	Pt. D	Pt. E	Pt. F	Pt. G
F need for information	2	0	4	4	1	3	4
F total number of nursing problems	46	27	42	65	47	25	38
F days of admission	8	5	8	10	5	6	5
Halloran index	5.75	5.40	5.25	6.50	9.40	4.17	7.60

Reporting the data in a longitudinally way is feasible with the current NMDSN, provided that the data collection is episode based. It generates additional information to the cross-sectional approach. This longitudinal reporting is also useful to generate epidemiological data, such as incidence and prevalence rates on the level of individual patients. This can be done with the NMDSN items, but for adequate reporting it requires that the data collection on the individual level goes alongside with data about the episode and that the unit of analysis is the individual patient, not the patient day.

A third purpose for analyzing the NMDSN is to determine the workload of patient categories. Two methods have been used for the NMDSN until so far. These are the nursing diagnoses index, developed by Halloran [10], and the San Joaquin patient classification as it is used in the Netherlands and Belgium [11, 12].

Medical case-mix alone does not explain all variability in the costs of treating hospitalized patients [10]. Clinical attributes of patients, including demand for nursing care and social services, have been associated with variations in the case cost and length of hospital stay [10]. Together these three discipline-related data sources provide information highly predictive and explanatory of patient resource use, cost and length of stay [10]. Halloran developed a method and algorithm to calculate the adequate nursing staffing levels based on a nursing diagnosis-based patient classification system [10]. The frequencies of nursing diagnoses are counted (scores 1 = condition exists and 0 = condition does not exist), and then indexed by summing all nursing conditions for the entire stay and dividing by the length of stay to determine the relative need for nursing care [10].

Because the nursing diagnoses data from the NMDSN are coded in exactly the same way as in Halloran's method, it is possible to calculate Halloran's nursing diagnoses index from the NMDSN data. In De Klerk's study, this index has been calculated for the 7 patients with hernia disease (See Table 2).

Sermeus developed a systematic algorithm to determine the levels of care dependency for the San Joaquin categories from the Belgian nursing minimum data set [13]. This method has been applied to the NMDSN. This works as follows: standard, every patient falls into San Joaquin category 2 (average care). De patient falls in San Joaquin category 1 (independent) if he is 'completely independent', or if he is independent for mobility, elimination, and feeding, and there is no infusion line. A patient falls in category 3 (heavy) if he needs full assistance with feeding, mobility or elimination, or needs tube feeding, and the self care needs to be taken over completely. Category 4 (extreme) is reached when category 3 applies, and the patient needs to be permanently observed. A patient can go into a higher category when there is disorientation, or assistance necessary with breathing. See Table 3 for the categories for the 2090 patient days in the NMDSN test [5].

Table 3 - Number of patients in the San Joaquin categories.

CAT	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	854	40,9	854	40,9
2	720	34,4	1574	75,3
3	482	23,1	2056	98,4
4	34	1,6	2090	100

To relate these care dependency data to the required time for nursing care, workload sampling is performed simultaneously to the NMDSN data collection [5]. The results of determining the nursing time required for each care dependency category are illustrated in table 4. This kind of information from the NMDSN is useful to support managerial decisions with respect to human resources and budgets.

Table 4 – Required nursing time per patient dependency category of the NMDSN.

CAT	Avg total time	Sd	N
1	56	39	854
2	78	44	720
3	129	70	482
4	153	72	34
Total	82	58	2090

A fourth purpose for the NMDSN is to illustrate the diversity of patient populations and the variations in nursing practice [3, 5, 8]. In the test of the NMDSN, the frequencies of the nursing diagnoses and interventions are further analyzed using ridit analysis [14]. Next, the data are presented graphically in terms of the ridits [7]. Ridit means 'Relative to an Identified Distribution' and is originally developed and applied by Bross for the analysis of ordinal data [14]. The ridit score does not require a normal distribution and preserves differences present in the (ordinal) data. It can be applied to nominal, interval and ratio data as well [7]. Ridit analysis illustrates the relative position of the score of a particular group against an identified distribution of a reference group [7, 14, 15, 16].

Table 5 – Calculation of ridits in the NMDSN

Part 5A calculation for the reference group of all 15 wards for the NMDSN item 'help with mobility'				
Answer Categories	Frequency Reference group	Frekw/2	Cumulative Freq+1	Ridit Score
Independent	1024	512	0	0.255
Partial assistance	589	294.5	1024	0.656
Complete assistance	398	199	1613	0.901
N :	2011			
		Avg. RIDIT =		0.500

Part 5B the ridit calculation for the subgroups for the NMDSN item 'help with mobility'				
Answer Categories	Frequencies Subgroups: Columns 6 to 9 for individual wards			
	Ward K	Ward M	Ward P	Ward Q
Independent	79	58	127	20
Partial assistance	58	27	49	23
Complete assistance	38	22	14	29
N subgroups:	175	107	190	72
Avg. RIDIT =	0.528	0.489	0.406	0.643
Substract 0.500 zero	-0.500	-0.500	-0.500	-0.500
ridit scores per ward	0.028	-0.011	-0.094	0.143

In the test of the NMDSN, the average ridit distribution of all 2090 patient days is used as the reference group against which the 15 individual general hospital wards are plotted. The calculation of the ridit scores in the NMDSN is illustrated in table 5.

In the first column of part 5A, the answer categories are listed from independent to complete assistance. N is the number of patient days, representing the reference group for this item. In column two the frequencies are listed for each answer category. Column three is the frequency of column two divided by 2. In the fourth column, the cumulate of column 2 is displaced one category downward, and the first category starting with a 0 [7, 14, 16]. Column five contains the calculation of the ridit of the reference distribution. This is, for each answer category, the summation of the value of column 3 plus the value of column 4 divided by N. The mean ridit for all patients on these five answer categories is listed on the bottom of column 5, and this is always 0.500 [7, 14, 16]. In column 6 to 9 (Part 5B), per ward and per answer category, the frequency is multiplied with the corresponding ridit score of the reference group. Then all the values are added up and divided by the number of patient days per ward.

A 'fingerprint' is a representation of the ridit score in a graph, illustrating the differences in the occurrence of the variables [5, 7, 13]. According to Bross [14], the normal value of a ridit score is between 0 and 1. To better interpret the 'fingerprint' graphs, Sermeus [7, 13] subtracted 0.5 from the average ridit of the reference group to get the zero value as reference point. This means that the average ridit for the reference group is 0 and the variables of the comparison groups (wards) with higher ridit scores are presented as bars to the right in the graph – illustrating 'more than average'. Similarly, variables with a lower ridit score go to the left, illustrating 'less than average'. Thus, the longer the bars are, the more that particular subgroup differs from the whole population of the reference group. Figure 1 illustrates such a fingerprint for the variables 'information need', 'pain', and 'pressure ulcer'.

To check the ridit scores, it is possible to estimate the confidence interval with the following formula [14]:

$$1 / \sqrt{3N} \qquad (1)$$

For example, the ridit score for risk for patient falls at the neurological ward in the sample is approximately 0.70, and N = 76. Applying the formula gives the following figures:

$$1 / \sqrt{3 * 76} = 1 / \sqrt{228} = 1 / 15.1 = 0{,}07 \qquad (2)$$

Thus the confidence intervals for this ridit score are between 0.7 + 0.07 = 0.77 and 0.7 – 0.07 = 0.63.

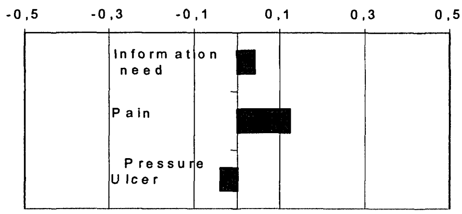

Figure 1: a fingerprint for three nursing diagnoses.

To determine the meaning of the differences between groups, the odds can be calculated by A/B. For example, the probability that a randomly chosen patient day from the surgical oncology ward has more pain than a randomly chosen patient day of the reference group, is 0.624545/0.5, which is about 1.25 to 1 [5]. A more accurate calculation is the odds ratio, which can be calculated based on the following equation [17]:

$$(0.50 + d) / (1- (0.50 + d)) \quad (3)$$

Where d is the numerical difference between average ridits of reference and test distributions. In the example above, this would be (0.5 + 0.124545) / (1 – (0.5 + 0.124545) = 1.66 to 1.

Such calculations are interesting in other group comparisons, e.g. in intervention studies. However, additional data must be available in order to estimate the reliability and significance. The Odds ratio can be used as a confidence limit that an individual will be worse of than an individual in the reference distribution or be worse of than an individual from another response category. As ridit scores are numeric scale values for variable categories, statistics as the arithmetic mean, standard deviation, standard error, analysis of variances, and principal component analysis can be used [7, 14, 15, 16, 17].

A fifth purpose of the NMDSN is to assist in instrument development. The Care Dependency Scale (CDS) is an instrument for assessing functional care demands of psychogeriatric nursing home patients [18]. The criterion-related validity of the CDS was investigated by studying the relationship between the CDS and three other instruments, including the NMDSN [19]. A convenience sample of 53 residential home patients agreed to participate in the study. The Pearson correlation coefficient, *r*, was used to test the relationship between the CDS sum score and relevant items of the NMDSN. Significant relationships (p = .01, 2-tailed) were found between the CDS sum score and selected items of the NMDSN. Two examples include *r* = -.59 for complexity of care, and *r* = .71 for assuming responsibility for self care [19]. This can be interpreted as follows: the higher the sum score on the CDS scale (and thus independence is higher), the lower the complexity of care is, and the higher the rating for assuming responsibility for self care. See Table 6.

Table 6 Pearson correlation between CDS sumscore and three NMSDN-items.

Scale	Number of patients #	Number of items	R
NMDSN-items: - complexity of care	52	1	-.59*
- assume responsibility for self care	53	1	.71*

(*) p= .01, 2-tailed #) Variation in *n* due to lack of response to the item [19].

This example illustrates the feasibility of using the NMDSN in nursing research, as an additional source of data for testing instruments against. Other examples could be to adapt the NMDSN in such a way that scientifically tested instruments become an integral part of it.

A sixth purpose of the NMDSN is to support health policy making thru combining the results for the individual wards into one graph [3, 8, 20]. Griens applied the ALSCAL procedure to combine the NMDSN data in one graph, illustrating the relative positions of wards and variables. See figure 2 [20]. She revealed that beside the differences between independence versus dependence for self care (left to right), a second dimension of fear and information needs versus cognitive problems (bottom to top) is important to illustrate nursing care.

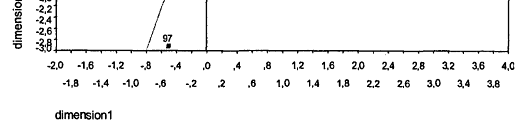

Figure 2: Combined map of NMDSN data for 15 wards.

Discussion and conclusion

This paper discussed feasible purposes the NMDSN. Six purposes of the NMDSN, their adequate methods for data collection, data analysis, and reporting results are described. If the different requirements for data collection can be combined, e.g. in computerized patient records, the NMDSN allows multiple uses. Although this listing of options is not complete, it is obvious that the NMDSN offers a maximum of information out of a minimum of data and effort.

References

[1] Werley HH, Devine EC, Zorn CR, Ryan P, Westra B. Nursing Minimum Data Set: abstraction tool for standardized, comparable, essential data. *Am Jnl of Pub Hlth* 1991: 18 (4): 421-426.

[2] Clark J, Lang NM. Nursings Next Advance: an International Classification for Nursing Practice. *Int Nurs Rev* 1992: 39:109-12.

[3] Sermeus W, Delesie L. The registration of a nursing minimum data set in Belgium: six years of experience. In: Grobe SJ, Pluyter-Wenting, ESP (Eds). *Nursing Informatics: An international overview for Nursing in a Technological Era.* Amsterdam, Elsevier North Holland 1994: pp. 144-149.

[4] Goossen WTF, Epping PJMM, Van den Heuvel WJA, Feuth T, Frederiks CMA, Hasman A. Development of the Nursing Minimum Data Set for the Netherlands (NMDSN): Identification of Categories and Items. *Jnl Adv Nurs*, 2000: 31 (3): 536-547.

[5] Goossen WTF, Epping PJMM, Feuth T, Van den Heuvel WJA, Hasman A, Dassen TWN. Using the Nursing Minimum Data Set for the Netherlands (NMDSN) to illustrate differences in patient populations and variations in nursing activities. In press. *Int Jrnl Nurs Stud* 2000: 38.

[6] Goossen WTF, Dassen TWN, Dijkstra A, Hasman A, Tiesinga L, and Van den Heuvel WJA. Validity and reliability of the NMDSN. In: Goossen WTF. *Towards strategic use of nursing information in the Netherlands.* Groningen, PhD Dissertation Rijksuniversiteit Groningen, 2000: pp. 183 – 201.

[7] Sermeus W, Delesie L. Ridit analysis on ordinal data. *West J Nurs Res* 1996: 18 (3): 351-359.

[8] Sermeus W, Delesie L. Development of a presentation tool for nursing data. In: Mortensen RA (Ed). *ICNP in Europe: Telenurse.* Amsterdam, IOS Press 1997: pp. 167-176.

[9] De Klerk T. *Zorg voor een profiel.* Nijmegen, Diemen, thesis for the BS Nursing Informatics program 1999.

[10] Halloran EJ, Patterson C, Kiley M. Case-Mix: matching patient need with nursing resource. *Nurs Manage* 1987: 18 (3): 27-42.

[11] Frumau DM, Baukema C. Patientenclassificatie en werklastonderzoek: de theorie, *Verpleegkundig Management* 1998: 4 (2): 25-29.

[12] Dierickx H, Sermeus W. *Zorgbehoefte als basis van staffing,* Leuven/Amersfoort, Acco, 1985.

[13] Sermeus W. *Variabiliteit van Verpleegkundige verzorging in Algemene Ziekenhuizen.* Leuven, Dissertatie Fac. der Geneeskunde, school voor de Maatschappelijke Gezondheidszorg 1992.

[14] Bross IDJ. How to use RIDIT analysis. *Biometrics* 1958: 14: 18-38.

[15] Fleiss JL. *Statistical Methods for Rates and Proportions.* New York, John Wiley and Sons 1981.

[16] Jansen ME. RIDIT Analysis, a review. *Statistica Neerlandica* 1984: 38: (3): 141-158.

[17] Kantor S, Winkelstein W, Ibrahim MA . A note on the interpretation of the ridit as a quantile rank. *Am J Epidemiol* 1968: 87 (3): 609-615.

[18] Dijkstra A, Buist G, Dassen ThWN. Nursing-Care Dependency: Development of an assessment scale for demented and mentally handicapped patients. *Scand J Caring Sci* 1996: 10: 137-143.

[19] Dijkstra A, Tiesinga LJ, Goossen WTF, Dassen TWN. Further psychometric testing of the Dutch Care Dependency Scale on two different patient groups. *submitted.*

[20] Griens AMGF. *Exploring the Nursing Minimum Data Set for the Netherlands, using different statistical approaches.* Master thesis, Leiden University, department of Psychology, 2000.

Address for correspondence

Acquest consulting,
Dorpsstraat 50,
2396 HC Koudekerk aan den Rijn,
the Netherlands.
Email: Acquest@Acquest.net