

BMJ Open Nursing workload, patient safety incidents and mortality: an observational study from Finland

Lisbeth Fagerström,^{1,2} Marina Kinnunen,³ Jan Saarela¹

To cite: Fagerström L, Kinnunen M, Saarela J. Nursing workload, patient safety incidents and mortality: an observational study from Finland. *BMJ Open* 2018;**8**:e016367. doi:10.1136/bmjopen-2017-016367

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2017-016367>).

Received 12 March 2017
Revised 31 January 2018
Accepted 26 March 2018



¹Faculty of Education and Welfare Studies, Åbo Akademi University, Vaasa, Finland

²Faculty of Health and Social Sciences, University College of Southeast Norway, Drammen, Norway

³Vaasa Central Hospital, Vaasa, Finland

Correspondence to

Professor Lisbeth Fagerström; lisbeth.fagerstrom@abo.fi

ABSTRACT

Objective To investigate whether the daily workload per nurse (Oulu Patient Classification (OPCq)/nurse) as measured by the RAFAELA system correlates with different types of patient safety incidents and with patient mortality, and to compare the results with regressions based on the standard patients/nurse measure.

Setting We obtained data from 36 units from four Finnish hospitals. One was a tertiary acute care hospital, and the three others were secondary acute care hospitals.

Participants Patients' nursing intensity (249 123 classifications), nursing resources, patient safety incidents and patient mortality were collected on a daily basis during 1 year, corresponding to 12 475 data points. Associations between OPC/nurse and patient safety incidents or mortality were estimated using unadjusted logistic regression models, and models that adjusted for ward-specific effects, and effects of day of the week, holiday and season.

Primary and secondary outcome measures Main outcome measures were patient safety incidents and death of a patient.

Results When OPC/nurse was above the assumed optimal level, the adjusted odds for a patient safety incident were 1.24 (95% CI 1.08 to 1.42) that of the assumed optimal level, and 0.79 (95% CI 0.67 to 0.93) if it was below the assumed optimal level. Corresponding estimates for patient mortality were 1.43 (95% CI 1.18 to 1.73) and 0.78 (95% CI 0.60 to 1.00), respectively. As compared with the patients/nurse classification, models estimated on basis of the RAFAELA classification system generally provided larger effect sizes, greater statistical power and better model fit, although the difference was not very large. Net benefits as calculated on the basis of decision analysis did not provide any clear evidence on which measure to prefer.

Conclusions We have demonstrated an association between daily workload per nurse and patient safety incidents and mortality. Current findings need to be replicated by future studies.

INTRODUCTION

Many studies have shown that insufficient nurse staffing in hospital-based care negatively affects outcomes such as mortality, infections and failure to rescue.^{1–6} However, the results are inconsistent and indicate a complex and non-linear relationship between nursing workload (NWL), mortality and other patient

Strengths and limitations of this study

- The study is the first to assess the relationship between nursing workload and patient outcomes based on data obtained on a daily basis.
- The instrument used here takes patient characteristics, such as age, sex and diagnoses, into account.
- The study provides some evidence to suggest that the traditional nurse staffing measure, the patients-to-nurse ratio, may partly fail to control for patient severity and casemix.
- The study does not address the potential influence of skill mix, competence level, work experience or the professionals' patient-related direct time.

outcomes.^{7–12} The strength of the evidence underpinning the association between nurse staffing and outcomes in previous studies can be challenged. Poor research designs, measurement problems and/or imprecise data that do not take into account daily variations in patients' care needs may contribute to the mixed findings.⁸ Higher nurse staffing and richer skill mix are associated with improved patient outcomes.^{4 8 10} Therefore, higher ratios have been recommended for improving patient safety and outcomes.^{1 9} However, it is difficult to set fixed, standard patient-to-nurse ratios for units in acute care hospitals, as evidenced in systematic reviews and other studies.^{7 10 13–15} Staffing levels must instead match patients' nursing care needs.^{8 16 17}

In an attempt to accommodate some of these issues, the RAFAELA patient classification system was developed in the 1990s in Finland.^{16 18 19} As compared with most other patient classification systems that use fixed patient-to-nurse ratios, the RAFAELA system use daily data on patients' care needs and the workload per nurse. The main purpose of the RAFAELA system is to ensure an appropriate allocation of nurse staff resources and, thus, a preferable NWL, which has been labelled as an optimal NWL. The latter term refers to

a situation when patients' care needs are assumed to be in balance with the nursing resources, and that working conditions can be assumed as being favourable, most desirable or satisfactory for the realisation of good nursing care.^{16 18–22} While certain realities such as economic restraints cannot be disregarded, the intention with the RAFAELA system is to provide a NWL measure dedicated to the reduction or elimination of adverse events.

In the RAFAELA system, NWL is based on daily assessments of patients' care needs and the registration of the nursing staff resources.¹⁶ The PAONCIL method ('Professional Assessment of Optimal Nursing Care Intensity Level') is used to establish an assumed optimal NWL for a specific ward. Daily measurements of NWL (Oulu Patient Classification (OPC)/nurse) are subsequently compared with this level, and resources are considered to be appropriately allocated when the actual NWL is at this level.^{19 21} This would mean that a satisfactory number of nurses, neither too many nor too few, are being allocated to provide care for the actual patient group.

We have found only two studies^{8 18} on the relationship between NWL based on assessed requirements for care (as opposed to nurse patient ratios or equivalent measures) and patient outcomes. Needleman *et al*⁸ found a significant association between patient mortality and increased exposure to unit shifts when nurse staffing was below the target level. In a recent study by Junttila *et al*¹⁸ based on monthly means, the incidence rate of death when average daily NWL was above the assumed optimal level was 13-fold that when the average daily NWL was below this level. However, to our knowledge, no studies exist on this relationship using daily-level data.

The aim of this observational study was therefore to investigate whether the daily workload per nurse (OPC/nurse), as a measure based on the RAFAELA system, correlates with patient safety incidents and patient mortality, using data collected on a daily basis. In addition, we wanted to compare the estimates with those based on the standard patients-to-nurse ratio (patients/nurse).

METHODS

Study setting

We obtained data from 36 units from four Finnish hospitals. One (A, 9 units) is a tertiary acute care hospital, whereas the three others (B, 14 units; C, 4 units; D, 9 units) are secondary acute care hospitals. The following specialties were included in the data material: internal medicine (eight units), surgical (eight units), paediatrics (five units), gynaecology (four units), maternity (two units), neurology (two units), orthopaedics (two units), oncology (one unit), rehabilitation (one unit), lung (one unit) and otology (one unit). Inclusion criteria were daily use of the RAFAELA system according to standards, reliable nursing intensity data as expressed in terms of a yearly reliability test done by parallel classifications (requirement is that unanimity is over 70%), and applicable nursing intensity

level measured with the PAONCIL method.^{16 19–21} Units that had undergone major organisational changes over the previous year were excluded. The A and B data represent the period 1 January to 31 December 2012, and the C and D data represent the period 1 January to 31 December 2013.

We did not include any sensitive health-related data of patients in the study, or any information regarding characteristics of the nurses. The RAFAELA is owned by the Association of Finnish Local and Regional Association Authorities and governed by non-commercial Finnish Consulting Group.

Measurement of NWL in the RAFAELA nursing intensity and staffing system

The RAFAELA is a standardised, person-centred, evidence-based system for nurse staffing that was developed in the 1990s.^{16 19} The feasibility, validity and reliability of the RAFAELA have been tested with good results.^{16 18 21 22} It is now used in about 90% of the hospitals in Finland, and has lately been implemented in Iceland, the Netherlands, Sweden and Norway.²² A requirement for users of the RAFAELA system is that the inter-rater reliability for nursing intensity measurements should be tested yearly.

The daily nursing intensity of each unit is assessed by all the responsible registered nurses on each day. One registered nurse usually classify one to six patients per day. The assessment is done every day by classifying each patient's care needs by the OPC instrument. This instrument consists of six subareas of nursing care. The nursing intensity level varies from 6 to 24 points for an individual patient per calendar day.^{16 19} The nurses' workload is calculated by dividing the total amount of nursing intensity points on the unit, for example, 350, with the number of nurses who take care of patients, for example, 12, during the same 24 hours. In this example, the patient-related NWL will then be 29.2 OPC points per nurse (hereafter referred to as OPC/nurse).

The underlying assumption of the RAFAELA system is that the nature and characteristics of nursing care differ between wards. The recommended NWL of each ward therefore has to be determined by the PAONCIL method. The development, testing and description of this method has been reported in several publications.^{16 19–22} Thus, the method is used to assess each ward's recommended optimal NWL including various contextual and organisational factors.²¹ The recommendation is that this level has to be reassessed by conducting the PAONCIL study every second year. The ones used in this study were not older than 2 or 3 years. The basic idea of the RAFAELA system is that the observed NWL (eg, 29.2 points/nurse) is compared with the established preferred for the same unit (eg, 22–30 points/nurse). If the observed NWL lies within the established limits, the nursing intensity is considered to be at the assumed optimal level.

The data we use in this paper consist of daily measurements based on the RAFAELA system.¹⁹ They correspond to every admitted patient's nursing intensity during 1 year

and were based on 249 123 classifications of patients' nursing intensity (OPC classifications). Each day, the patient-related nurse resources were also recorded, using a standardised model where non-patient-related time was excluded. Apart from each day's staff data (OPC/patient, OPC/nurse, etc), there was daily information also on patient incidents and patient mortality.

All data were collected during a period of 1 year, meaning that there were 12 475 data points (not approximately 13 140, since some wards were closed for shorter periods, foremost because of holidays). **Table 1** provides the central variables of the data in terms of each unit's PAONCIL level, daily mean number of classified patients, daily mean number of OPC classifications, total OPC points, nursing staff resources, number of patients per nurse, OPC points per nurse, incidents and deaths.

Outcomes

Data on incidents were collected daily from The Reporting System for Safety Incidents in Health Care Organizations (HaiPro), which is a comprehensive and standardised patient safety system in Finland.^{23 24} As defined by HaiPro, an incident is a safety hazard that may harm or harm the patient. Incidents are classified into 14 categories,²⁴ but there are two main categories: near miss, which may have caused harm to the patient, but was prevented by chance or by timely preventive actions and adverse events, which are negative events that caused harm to the patient. To roughly capture the severity of an event, we categorised incidents in four ways¹: whether at least one incident, of any type, occurred (incident),² whether a patient was affected to any degree (patient affected),³ whether the incident caused harm to the patient (harm to patient), and⁴ whether there was more than one incident, of any type, on the same day (>1 incident), within the available follow-up of 365 days. In addition, we used patient mortality (death) as a fifth type of adverse event. Some wards had no deaths during the study period, but excluding them from the analyses would not affect the results to any noteworthy degree. The mortality data were retrieved from the local mortality register of each hospital.

Statistical analyses

Using logistic regression analyses, associations were estimated on the daily level between nursing intensity per nurse (OPC/nurse) in relation to the assumed optimal level and each type of outcome, that is, each of the four types of incidents and mortality. For each of type of outcome, the event was coded as 0 or 1, meaning that either there was no event during that specific day, or there was an event. The use of logistic regression models accommodate any issues related to non-normal distributions. We estimated associations both in unadjusted models and in models that adjusted for ward-specific effects and effects of day of the week, holiday and season, using dummy variables. Thus, we allowed for heterogeneity in the intercept term, which was motivated by the

fact that across-ward variability was fairly modest. The categories of the variables are described in the footnotes of **table 2**. Parallel analyses were performed with the standard measure of NWL, patients/nurse. Supplementary electronic files provide full details of the models estimated (see online supplementary file 1).

We report results in which evaluations based on the RAFAELA system (OPC/nurse) were assessed using the assumed optimal level with a $\pm 15\%$ deviation around this point,^{16 19 21} and in which the patients/nurse measure was assessed using a categorisation with three equally large groups. The results reported in **table 2** were consequently based on 20 different regressions. Model fit indices (-2 log likelihood, Akaike information criterion and Nagelkerke's R^2) are provided to facilitate comparisons between regressions based on the OPC/nurse measure and the patients/nurse measure. The analyses were performed using SPSS V.21. All estimates are expressed in terms of ORs with 95% CIs.

Apart from comparing the predictive accuracy of the models that use the OPC/nurse measure and the patients/nurse measure, respectively, we have also used decision-analytic methods.²⁵ These ascertain the value of prediction models by incorporating information on consequences and they require explicit valuation of outcomes. The technique may thus help in deciding on which measure to prefer, that is, the one with a higher net benefit.

RESULTS

When OPC/nurse was above the assumed optimal level, the unadjusted odds for a patient safety incident were 1.28 (95% CI 1.13 to 1.45) that of the assumed optimum level (see **table 2**). Corresponding ORs for the other types of incidents, patient affected, harm to patient and >1 incident were 1.13 (95% CI 0.96 to 1.32), 1.16 (95% CI 0.93 to 1.45) and 1.25 (95% CI 0.95 to 1.66), respectively. ORs for patient mortality were even higher or 1.42 (95% CI 1.19 to 1.69). If OPC/nurse was below the recommended optimal level, the ORs for incidents and patient mortality were conversely lower or around 0.67 for the different types of incidents, and 0.55 for patient mortality.

When ward-specific effects and effects of day of the week, holiday and season were adjusted for, the ORs diminished somewhat (see **table 2**). NWL above the assumed optimal level was associated with 8%–34% higher odds of an incident, depending on the type of incident, and 43% higher odds of patient mortality, as compared with if it was at the assumed optimal level. If OPC/nurse was below this level, the OR for an incident and for patient mortality was approximately 25% lower. Adding the ward-specific effects improved model fit considerably. Also the variables for weekday, holiday and season improved the model fit, except for the outcomes >1 incident and death. The odds for incidents were in general least likely to occur on Saturdays and on holidays, whereas there were no obvious seasonal effects (not shown here). Complete descriptions

Table 1 Assumed optimal workload, and daily mean values of workload, staff resources and adverse events per ward in the data

Ward_id	Optimal load, lower bound	Optimal load, upper bound	Optimal OPC per nurse	Patients per nurse	OPC per patient	Number of nurses	Incident affected	Patient affected	Harm to patient	>1 incident	Death	n	Number of OPC classifications
D Infectious*	18.90	25.56	21.06	1.48	251.22	12.03	0.09	0.06	0.02	0.01	0.04	365	17.70
D Internal†	21.85	29.55	24.45	1.60	308.81	12.74	0.13	0.07	0.04	0.02	0.07	362	20.18
D Cardiology	21.88	29.60	25.98	1.89	355.99	13.98	0.26	0.11	0.08	0.03	0.08	365	25.90
D Surgical	16.85	22.79	23.88	1.41	425.36	17.80	0.05	0.02	0.01	0.00	0.08	365	25.19
D Orthopaedics	15.99	21.63	21.92	1.41	350.84	16.15	0.12	0.08	0.04	0.02	0.05	365	22.62
D Surgical	14.54	19.66	17.58	1.24	243.46	13.75	0.07	0.05	0.04	0.01	0.02	330	17.15
D Maternal	25.47	34.47	33.83	2.48	343.10	10.12	0.04	0.02	0.01	0.00	0.00	364	25.17
D Oncology	14.36	20.12	22.84	1.38	249.46	10.88	0.26	0.11	0.05	0.03	0.18	362	15.10
D Neurology	20.05	27.13	25.12	1.34	392.29	15.76	0.13	0.05	0.03	0.02	0.08	365	20.96
B Internal†	16.14	21.83	20.40	1.49	182.62	13.75	0.12	0.09	0.07	0.02	0.06	341	13.29
B Internal†	20.37	27.56	25.13	2.00	290.05	11.59	0.23	0.16	0.10	0.06	0.08	365	23.04
B Internal†	17.16	23.21	19.94	1.47	171.98	13.54	0.20	0.11	0.07	0.06	0.05	365	12.69
B Internal†	19.65	26.59	24.16	1.82	292.34	13.24	0.11	0.07	0.05	0.01	0.13	365	22.09
B Surgical	21.47	29.04	29.16	2.00	544.95	14.59	0.20	0.16	0.10	0.01	0.16	365	37.34
B Surgical	18.50	25.10	23.29	1.77	355.67	13.18	0.08	0.05	0.03	0.02	0.08	365	26.95
B Surgical	23.27	31.48	28.04	1.88	474.66	14.90	0.12	0.08	0.01	0.02	0.05	365	31.85
B Gynaecology	18.87	25.54	23.17	1.68	131.87	14.04	0.13	0.10	0.02	0.02	0.02	177	9.45
B Maternity	19.05	25.78	20.69	1.37	338.39	15.04	0.03	0.02	0.02	0.00	0.00	365	22.48
B Paediatrics	11.51	15.58	13.08	0.75	134.87	17.39	0.04	0.03	0.01	0.00	0.00	362	7.72
B Paediatrics	12.16	16.45	8.65	0.49	108.10	17.52	0.02	0.01	0.00	0.00	0.00	365	6.17
B Otology	17.45	23.60	22.45	1.81	224.67	12.36	0.11	0.07	0.04	0.01	0.01	298	18.12
B Neurology	16.25	21.98	19.37	1.37	245.29	14.24	0.08	0.06	0.01	0.01	0.07	365	17.30
B Lung	20.02	27.09	19.43	1.50	273.30	12.97	0.10	0.08	0.01	0.02	0.11	365	21.05
C Surgical	22.60	30.60	24.33	1.41	321.42	17.29	0.17	0.08	0.03	0.07	0.06	365	18.58
C Internal†	18.70	25.30	25.15	1.65	397.05	15.20	0.17	0.09	0.01	0.05	0.15	366	26.12
C Internal†	20.40	27.60	21.77	1.29	208.62	16.87	0.10	0.09	0.01	0.02	0.03	291	12.33
C Surgical	19.40	26.30	23.70	1.51	228.64	15.64	0.04	0.03	0.00	0.01	0.00	291	14.58
A Gynaecology	19.50	26.40	24.92	1.83	290.61	13.60	0.03	0.02	0.01	0.00	0.00	366	21.14
A Gynaecology	25.60	42.10	39.84	2.81	510.25	14.21	0.04	0.03	0.02	0.00	0.00	225	35.94
A Gynaecology	33.50	45.30	40.54	2.85	607.06	14.22	0.04	0.02	0.02	0.00	0.00	366	42.64
A Orthopaedics	16.80	22.70	20.29	1.33	430.29	15.18	0.10	0.05	0.02	0.02	0.04	366	28.30
A Paediatrics	12.30	14.90	13.33	0.97	165.18	13.85	0.15	0.10	0.08	0.04	0.00	366	11.95

Continued

Table 1 Continued

	Optimal load, lower bound	Optimal load, upper bound	Optimal OPC per nurse	Patients per nurse	OPC	OPC per patient	Number of nurses	Incident affected	Patient affected	Harm to patient	>1 incident	Death	n	Number of OPC classifications
A Paediatrics	11.00	14.20	11.16	0.70	117.12	16.05	10.57	0.09	0.06	0.02	0.01	0.01	360	7.31
A Paediatrics	8.90	12.00	9.29	0.54	91.18	17.36	9.67	0.09	0.06	0.02	0.01	0.00	363	5.24
A Surgical	21.10	28.50	24.45	1.48	273.20	16.50	11.24	0.02	0.02	0.01	0.00	0.06	364	16.54
A Rehabilitation	15.10	20.50	20.57	1.49	198.98	13.89	9.51	0.14	0.12	0.03	0.04	0.00	315	14.52
Total	18.52	25.22	22.41	1.53	294.59	14.97	12.93	0.11	0.07	0.03	0.02	0.05	12475	19.97

Assumed optimal load refers to the established interval of optimal workload according to the PAONCIL measurement (OPC per nurse).

The variables are described in more detail in the main text.

*Infectious diseases.

†Internal medicine.

OPC, Oulu Patient Classification; PAONCIL, Professional Assessment of Optimal Nursing Care Intensity Level.

of all estimates and the models estimated, with predictive indices, can be found in the supplementary electronic file (see online supplementary file 1).

The two lower panels in table 2 provide results of parallel analyses when NWL was measured according to the standard patients-to-nurse ratio (patients/nurse). As compared with results based on the RAFAELA system, there are three main issues to be pointed out. First, effects sizes in terms of ORs were consistently smaller with the patients/nurse approach than with the OPC/nurse approach, irrespective if unadjusted or adjusted models are compared. For instance, in the fully adjusted model, the OR of an incident was 1.13 (95% CI 0.96 to 1.33) if workload was in the highest one-third, and 0.89 (95% CI 0.75 to 1.05) if it was in the lowest one-third, as compared with if it was in the middle one-third. These effects were notably smaller than the estimated relative effect sizes for being above and below the recommended optimum according to the RAFAELA system, which were 1.24 (95% CI 1.08 to 1.42) and 0.79 (95% CI 0.67 to 0.93), respectively. Second, in almost all instances, the estimates of the patients/nurse approach had smaller statistical power in terms of wider CIs (ie, larger SEs). However, far from all estimates for the OPC/nurse measure, or for the patients/nurse measure, were statistically significant at the 5% level. Third, when comparing results for the patients/nurse measure to the OPC/nurse measure for otherwise similar models and outcomes, the model fit of the former was consistently poorer (values of the log likelihood and AIC were higher and R square lower). It nevertheless needs to be stressed that the difference was not very large.

We experimented also with other ways to categorise NWL. For OPC/nurse, we used an alternative with a halved deviation from the recommended optimal point, that is, $\pm 7.5\%$ instead of $\pm 15\%$, and with a doubled deviation, that is, $\pm 30\%$ from the optimal point. The patient-to-nurse measure was also assessed using alternative categorisations, such as five and seven equally large groups, respectively. Results of these additional regressions supported the overall conclusions as reported above. In models using the patients/nurse measure, associations were mostly weaker, came with lower statistical power, and they were less systematic, as compared with models based on the OPC/nurse measure (see online supplementary file 1).

Hence, our analyses suggest that, in terms of predictive accuracy, models estimated on basis of NWL according to the RAFAELA system are slightly to be preferred above otherwise similar models that use the standard patients/nurse measure. It is not evident, however, which measure is to be preferred when it comes to decision-making. Figures 1–5 summarise net benefit values calculated based on the models estimated for each type of patient safety incident and patient mortality, respectively; see Vickers and Elkin²⁵ for technical details. The values have been calculated over a reasonable range for the probability of an event (type of incident or mortality). Models

Table 2 OR for an adverse event (with 95% CI) for four types of patient safety incidents and for patient mortality, according to NWL measurement by the RAFAELA system (OPC/nurse) and the standard NWL measurement system (patients/nurse), unadjusted and adjusted estimates

	Incident	Patient affected	Harm to patient	>1 incident	Death
OPC/nurse, unadjusted model					
Below optimum	0.67 (0.58–0.78)	0.68 (0.56–0.82)	0.66 (0.50–0.88)	0.67 (0.47–0.95)	0.55 (0.43–0.70)
At optimum	1	1	1	1	1
Above optimum	1.28 (1.13–1.45)	1.13 (0.96–1.32)	1.16 (0.93–1.45)	1.25 (0.95–1.66)	1.42 (1.19–1.69)
–2 log likelihood	8577.5	6169.3	3523.0	2406.4	4958.6
Akaike information criterion	8561.5	6173.3	3527.0	2410.4	4962.6
Nagelkerke's R ²	0.0106	0.0056	0.0052	0.0056	0.0160
OPC/nurse, adjusted model					
Below optimum	0.79 (0.67–0.93)	0.78 (0.64–0.96)	0.85 (0.63–1.14)	0.73 (0.50–1.07)	0.78 (0.60–1.00)
At optimum	1	1	1	1	1
Above optimum	1.24 (1.08–1.42)	1.08 (0.91–1.28)	1.11 (0.88–1.41)	1.32 (0.98–1.79)	1.43 (1.18–1.73)
–2 log likelihood	8010.8	5856.3	3211.1	2187.9	4286.5
Akaike information criterion	8106.8	5952.3	3307.1	2283.9	4382.5
Nagelkerke's R ²	0.0960	0.0688	0.1050	0.1041	0.1733
Patients/nurse, unadjusted model					
First group	0.74 (0.64–0.86)	0.85 (0.71–1.02)	0.79 (0.61–1.04)	0.80 (0.58–1.10)	0.47 (0.38–0.58)
Second group	1	1	1	1	1
Third group	1.09 (0.95–1.25)	1.18 (0.99–1.41)	1.24 (0.96–1.58)	0.95 (0.70–1.30)	0.97 (0.81–1.17)
–2 log likelihood	8589.1	6180.9	3525.1	2416.5	4958.8
Akaike information criterion	8593.1	6184.9	3529.1	2420.5	4962.8
Nagelkerke's R ²	0.0055	0.0033	0.0045	0.0010	0.0159
Patients/nurse, adjusted model					
First group	0.89 (0.75–1.05)	0.98 (0.80–1.21)	0.90 (0.66–1.23)	1.01 (0.71–1.44)	0.86 (0.68–1.08)
Second group	1	1	1	1	1
Third group	1.13 (0.96–1.33)	1.15 (0.94–1.41)	1.03 (0.77–1.39)	1.15 (0.81–1.64)	1.20 (0.97–1.49)
–2 log likelihood	8029.8	5863.1	3213.4	2196.1	4301.8
Akaike information criterion	8.125.8	5959.1	3309.4	2292.1	4397.8
Nagelkerke's R ²	0.0931	0.0674	0.1043	0.1004	0.1698
Number of events	1367	848	400	246	636

The table summarises results from 20 different models estimated on 12 475 calendar days, representing 36 wards at 4 hospital units.

Adjusted model refers to models adjusted for ward-specific effects and effects of the of the week, holiday and season.

Estimates for ward-specific effects and effects of day of the week, holiday and season are found in the online supplementary electronic files.

At optimum refers to the assumed optimal nursing intensity point with $\pm 15\%$ deviation, as defined by the RAFAELA system.

Patients/nurse refer to a categorisation into three equally large groups.

Categories used for day of the week are Monday, Tuesday, Wednesday, Thursday, Friday, Saturday and Sunday.

Categories used for holiday are No or Yes, where Yes refers to Easter, Midsummer, Christmas and New Year.

Categories used for season are January–March, April–May, June–August, September–October and November–December.

NWL, nursing workload; OPC, Oulu Patient Classification.

based on the OPC/nurse measure and the patients/nurse measure are to be compared by looking at the net benefit values (see figures 1–5). The one with higher net benefit values is to be preferred above the other. As shown by the figures, there is no clear discrepancy. For some threshold probabilities, the OPC/nurse measure lies above the patients/nurse measure, while for others, the situation is the opposite. In addition, for each event (type of incident and mortality), the two curves are essentially overlapping,

and in most instances the difference in net benefit values is rather modest. In terms of the magnitude of the benefit for patients, it is consequently not evident which measure of NWL is to be preferred.

DISCUSSION

We find that the odds for a patient safety incident were 10%–30% higher, and for patient mortality about 40%

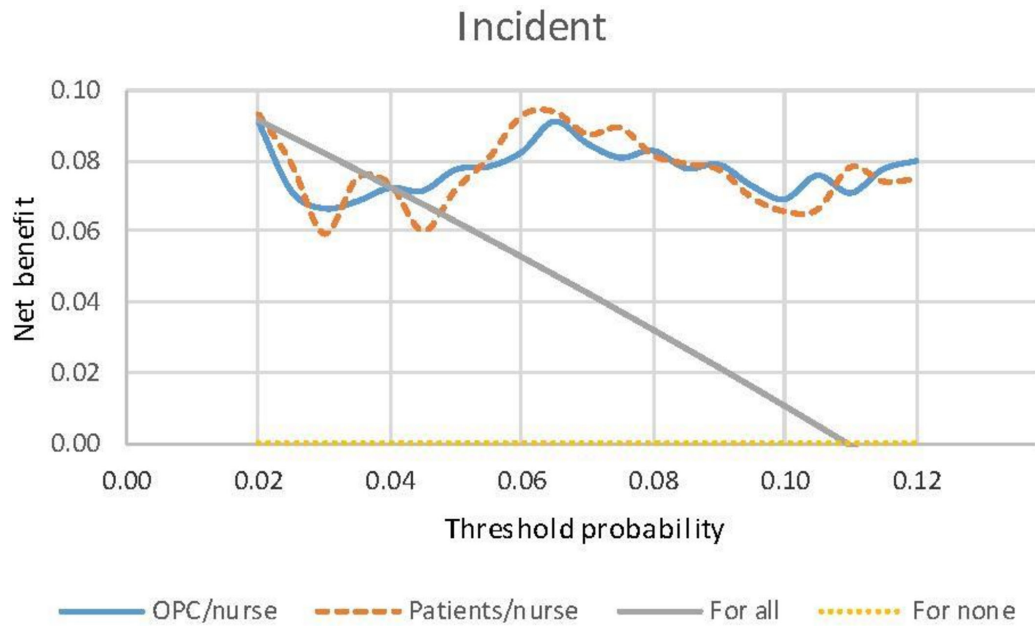


Figure 1 Decision curves for incident according to the OPC/nurse measure and patients/nurse measure, respectively. OPC, Oulu Patient Classification.

higher, if the NWL as measured by the RAFAELA system (OPC/nurse) was above the assumed optimal level, as compared with if it was at this level. If OPC/nurse was below the level, the odds for a patient safety incident and for mortality were approximately 25% lower. The latter situation would mean that nurses have more time for caring and observing each patient, which may reduce the risk for adverse events and accordingly prevent the patient's health condition from deteriorating.

Previous research⁹ did not find significant changes in patient safety associated with decreased NWL and could not confirm compliance with ratios per shift.

Other studies used hospital-level administrative data that imprecisely allocated staffing to patients' care needs.^{8 11} We think that such associations between nurse staffing, patient outcomes and mortality may be challenged.^{12 18} Needleman *et al*⁸ found similar results between mortality and day-to-day, shift-to-shift variation in staffing, and Junttila *et al*¹⁸ between mortality and days with NWL over optimal level on a monthly level. The OPC/nurse measure is more detailed than the traditional patients/nurse measure. While comparable to the 'hours per patient day model',²⁶ its accuracy of nursing resources is higher. For example, if a nurse becomes sick during a

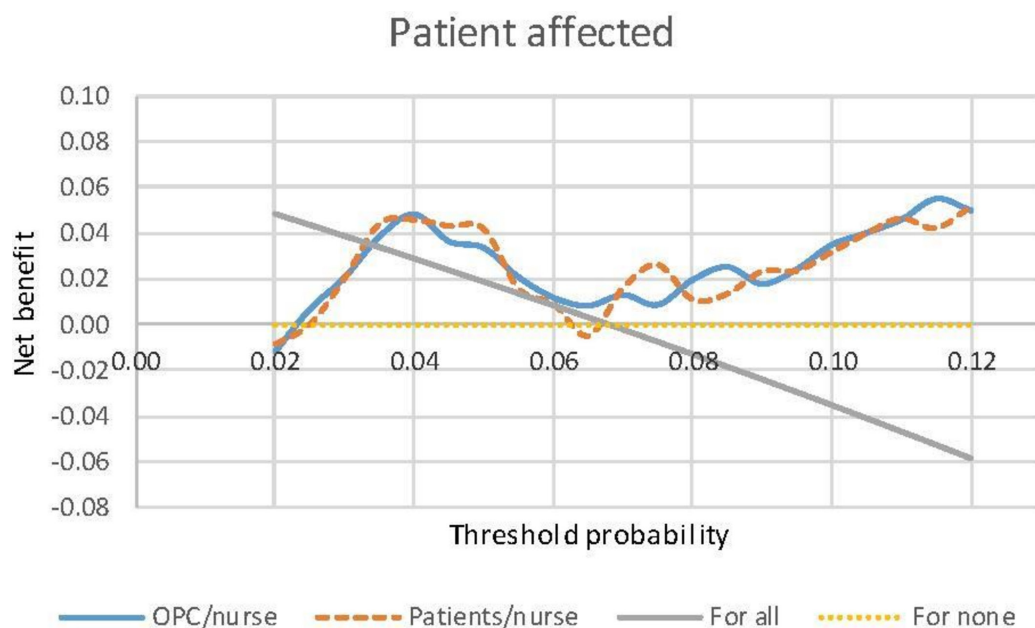


Figure 2 Decision curves for patient affected according to the OPC/nurse measure and patients/nurse measure, respectively. OPC, Oulu Patient Classification.

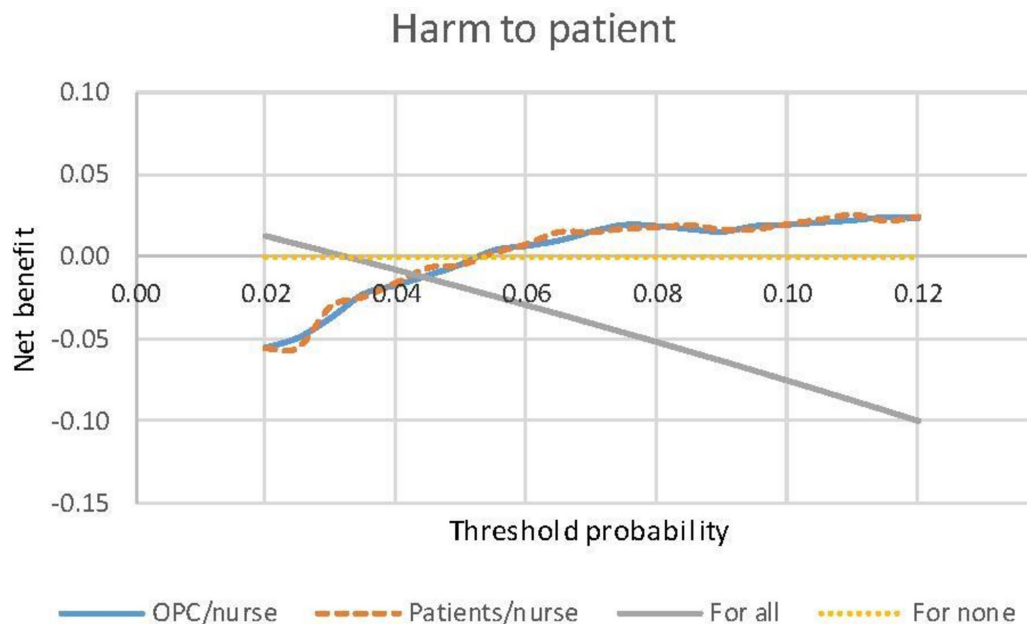


Figure 3 Decision curves for harm to patient according to the OPC/nurse measure and patients/nurse measure, respectively. OPC, Oulu Patient Classification.

shift and leaves the unit, the nurse in charge will deduct these hours from the unit's resources.

Several factors affect the reporting of incidents, for example, staff's lack of motivation or knowledge, nurse staff shortage, stressful situations or burn-out. A reasonable argument is therefore that a very high NWL indicates a working situation where the nurse staff resources are too low. Still, too few resources can result in the de-prioritisation of the registration of adverse events and thus the under-reporting of incidents connected to high NWL, which may affect the results of our study and the conclusions that we draw.

Our study provided results based on daily measures of all-in-hospital patients' actual nursing intensity, including detailed registration of used staff resources and the association with incidents and mortality on daily levels. The HaiPro database, on which our analyses were based, meets WHO criteria for a good reporting system.^{23 24} However, we know that despite a good reporting system, incidents reports are missing due to several reasons, such as lack of time, personnel's involvement etc. The Global Trigger Tool is another method to analyse patient safety, which has been recommended.²⁷ However, it collects triggers and patient safety incidents from treatment periods, not on a daily basis, whereas

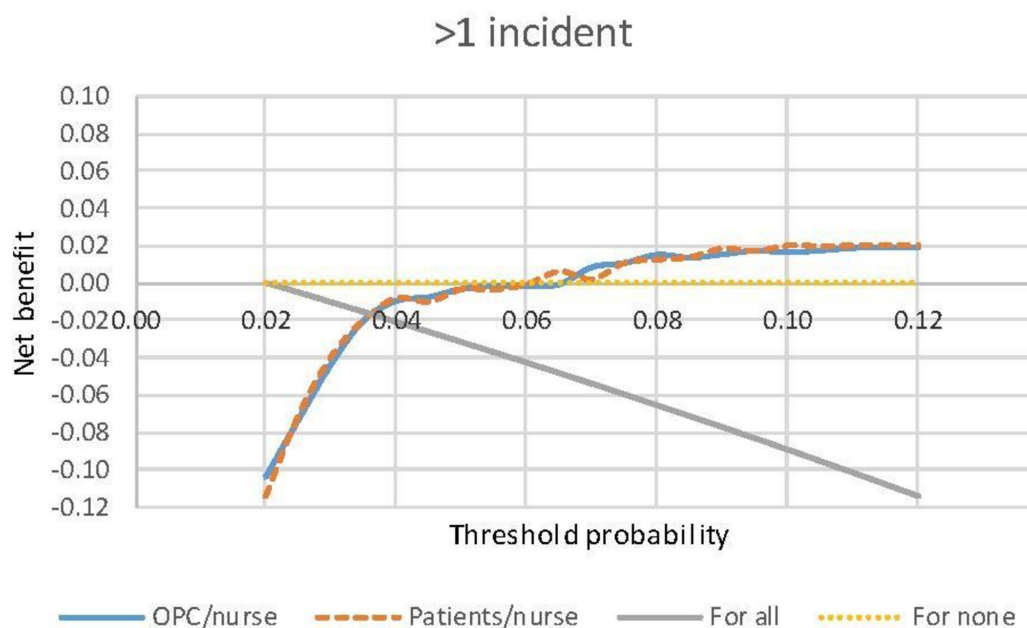


Figure 4 Decision curves for >1 incident according to the OPC/nurse measure and patients/nurse measure, respectively. OPC, Oulu Patient Classification.



Figure 5 Decision curves for death according to the OPC/nurse measure and patients/nurse measure, respectively. OPC, Oulu Patient Classification.

data on incidents collected from HaiPro can be targeted to certain days.²⁴ Units that underwent major organisational changes over the previous year were excluded from our study, because they may negatively influence the data quality including incident reports. The accuracy of the data used, in terms of NWL, incidents and mortality, is highly reliable and probably better than in previous studies on NWL and adverse outcomes. The staffing measurement determined by the RAFAELA system implicitly considers specific characteristics of each ward, such as organisational factors in terms of unit size, leadership and physical environment.^{16 19 21}

We found evidence that a staffing measure based on daily measurement of individual patient care needs and the recommended NWL (OPC/nurse) is slightly better in predicting incidents and mortality rates as compared the standard patient-to-nurse measure. Yet it needs to be stressed that, based on decision curve analysis, it was not clear which measure of NWL will produce higher net benefit in terms of avoiding patient safety incidents and patient mortality. Current findings therefore ought to be further investigated and the findings replicated in larger, longitudinal multicentre studies.

A strength of this study is that the analyses were conducted based on nurses' independent classifications of patients' nursing intensity. The data used was based on a scientifically tested NWL system, which enabled comparisons,¹⁶ since the patient-casemix and patient severity groups require different staff resources to maximise positive patient outcomes.^{4 8 18 28 29} NWL consequently ought to be monitored daily using reliable instruments to ensure good patient outcomes. Such optimal resource allocation is needed for successful leadership and clinical governance, and it is crucial for favourable outcomes, to preventing adverse events and to reducing patient mortality.

Our study nevertheless has certain limitations. The reliability of incident reports can always be questioned, despite that the HaiPro system has been in systematic use for almost 10 years. Although we could control for ward-specific effects and effects of day of the week, holiday and season, there might be other confounding factors. Hospital settings are characterised by complexity regarding factors that may affect total NWL.^{1 2 13 28-31} While a list of central organisational and contextual factors were included in the PAONCIL instrument, we did not address the effects of skill mix, competence level or work experience on patient outcomes. Physicians' patient-related direct time and healthcare support should also probably be included in further studies.³² Further analyses of other patient characteristics, such as age, sex or diagnoses, were not conducted because the OPC instrument takes these variables into account. Earlier studies have shown that the OPC instrument identifies patients' individual characteristics such as functional ability, symptoms of diseases and the effect on nursing intensity of the most central patient characteristics.^{16 22} Hence, the measurement by the OPC covers the actual patient casemix for each day. However, the contribution of these aspects, especially age and sex, may be analysed in more details in further studies. Another limitation was that a death or an incident caused by low staffing on a ward on 1 day may not always occur on that same day or at that same ward. This could be explored by analysing patient records around the critical days and at multiple wards. Although this study was the first about the relationship between the assumed optimal NWL and daily outcomes, a multicentre study with several hospitals is needed to further test the generalisability of the results.

CONCLUSIONS

This study has showed that a work situation above the assumed optimal level increases the risk for adverse events and patient mortality. However, the resources for nursing staff are limited in all organisations. Nurse managers, therefore, have to use available resources in the most optimal way. This study provided some new evidence to suggest that the traditional nurse staffing method, the patient-to-nurse ratio, is not necessarily preferable when it comes to controlling for patients' severity and casemix. The staffing measure based on the assumed optimal NWL may therefore be considered a novel attempt to fill a gap in the existing knowledge on leadership and clinical governance. Efficient resource allocation is needed for successful leadership and clinical governance and it is crucial for favourable outcomes, for preventing adverse events and for reducing the mortality risk. Future research is needed to ascertain whether good patient outcomes are ensured by daily monitoring of nurses' workload with instruments like the one studied here.

Contributors LF did the literature search. LF and JS designed the study. LF collected the data. JS prepared the data and performed the analyses. LF, MK and JS contributed to data interpretation, writing and revision of the report.

Funding The authors disclose receipt of the following financial support for the research and authorship for this article: The State Research Funding of Vaasa Hospital District based on the Health Care Act (1326/2010).

Competing interests None declared.

Patient consent Not required.

Ethics approval This study received approval from the chief administrative physicians of all four hospitals involved. No further ethical approval was therefore necessary, which is in accordance with the regulatory regime for conducting health research in Finland.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Full descriptions of all models estimated and their estimates are found in the supplementary electronic files.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2018. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

REFERENCES

- Aiken LH, Clarke SP, Sloane DM, *et al.* Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction. *JAMA* 2002;288:1987–93.
- Aiken LH, Cimmiotti JP, Sloane DM, *et al.* Effects of nurse staffing and nurse education on patient deaths in hospitals with different nurse work environments. *Med Care* 2011;49:1047–53.
- Aiken LH, Sloane DM, Bruyneel L, *et al.* Nurse staffing and education and hospital mortality in nine European countries: a retrospective observational study. *Lancet* 2014;383:1824–30.
- Seago JA, Williamson A, Atwood C. Longitudinal analyses of nurse staffing and patient outcomes: more about failure to rescue. *J Nurs Adm* 2006;36:13–21.
- Rafferty AM, Clarke SP, Coles J, *et al.* Outcomes of variation in hospital nurse staffing in English hospitals: cross-sectional analysis of survey data and discharge records. *Int J Nurs Stud* 2007;44:175–82.
- Zingg W, Holmes A, Dettenkofer M, *et al.* Hospital organisation, management, and structure for prevention of health-care-associated infection: a systematic review and expert consensus. *Lancet Infect Dis* 2015;15:212–24.
- Lang TA, Hodge M, Olson V, *et al.* Nurse-patient ratios: a systematic review on the effects of nurse staffing on patient, nurse employee, and hospital outcomes. *J Nurs Adm* 2004;34:326–37.
- Needleman J, Buerhaus P, Pankratz VS, *et al.* Nurse Staffing and Inpatient Hospital Mortality. *N Engl J Med Overseas Ed* 2011;364:1037–45.
- Donaldson N, Bolton LB, Aydin C, *et al.* Impact of California's licensed nurse-patient ratios on unit-level nurse staffing and patient outcomes. *Policy Polit Nurs Pract* 2005;6:198–210.
- Kane RL, Shamliyan TA, Mueller C, *et al.* The association of registered nurse staffing levels and patient outcomes: systematic review and meta-analysis. *Med Care* 2007;45:1195–04.
- Shuldhham C, Parkin C, Firouzi A, *et al.* The relationship between nurse staffing and patient outcomes: a case study. *Int J Nurs Stud* 2009;46:986–92.
- Van den Heede K, Sermeus W, Diya L, *et al.* Nurse staffing and patient outcomes in Belgian acute hospitals: cross-sectional analysis of administrative data. *Int J Nurs Stud* 2009;46:928–39.
- Tarnow-Mordi WO, Hau C, Warden A, *et al.* Hospital mortality in relation to staff workload: a 4-year study in an adult intensive-care unit. *Lancet* 2000;356:185–9.
- Lang TA, Hodge M, Olson V, *et al.* Nurse-patient ratios: a systematic review on the effects of nurse staffing on patient, nurse employee, and hospital outcomes. *J Nurs Adm* 2004;34:163–74.
- National Institute for Health and Care Excellence. Safe Staffing for nursing in adult inpatient wards in acute hospitals. Safe staffing guideline 1. 2014 <http://www.nice.org.uk/sg1>
- Rauhala A. *The validity and feasibility of measurement tools for human resources management in nursing case of the RAFAELA system*: Department of Health Policy and Management, University of Kuopio, 2008. Doctoral dissertation.
- Adomat R, Hewison A. Assessing patient category/dependence systems for determining the nurse/patient ratio in ICU and HDU: a review of approaches. *J Nurs Manag* 2004;12:299–308.
- Junttila JK, Koivu A, Fagerström L, *et al.* Hospital mortality and optimality of nursing workload: A study on the predictive validity of the RAFAELA Nursing Intensity and Staffing system. *Int J Nurs Stud* 2016;60:46–53.
- Rauhala A, Fagerström L. Determining optimal nursing intensity: the RAFAELA method. *J Adv Nurs* 2004;45:351–9.
- Rauhala A, Kivimäki M, Fagerström L, *et al.* What degree of work overload is likely to cause increased sickness absenteeism among nurses? Evidence from the RAFAELA patient classification system. *J Adv Nurs* 2007;57:286–95.
- Fagerström L, Rainio AK, Rauhala A, *et al.* Professional assessment of optimal nursing care intensity level. A new method for resource allocation as an alternative to classical time studies. *Scand J Caring Sci* 2000;14:97–104.
- Andersen MH, Lønning K, Bjørnelv GM, *et al.* Nursing intensity and costs of nurse staffing demonstrated by the RAFAELA system: liver vs. kidney transplant recipients. *J Nurs Manag* 2016;24:798–805.
- Kinnunen M. *Learning from errors in organization – Factors preventing or making it possible*: University of Vaasa, Acta Wasaensia:230. Doctoral thesis.
- Holmström A-R. *Learning from Medication errors in healthcare – how to make medication error reporting systems work? Doctoral dissertation*: Faculty of Pharmacy, University of Helsinki, 2017.
- Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making* 2006;26:565–74.
- Spetz J, Donaldson N, Aydin C, *et al.* How many nurses per patient? Measurements of nurse staffing in health services research. *Health Serv Res* 2008;43:1674–92.
- Classen DC, Resar R, Griffin F, *et al.* 'Global trigger tool' shows that adverse events in hospitals may be ten times greater than previously measured. *Health Aff* 2011;30:581–9.
- Needleman J, Buerhaus P, Mattke S, *et al.* Nurse-staffing levels and the quality of care in hospitals. *N Engl J Med* 2002;346:1715–22.
- Griffiths P, Ball J, Drennan J, *et al.* Nurse staffing and patient outcomes: Strengths and limitations of the evidence to inform policy and practice. A review and discussion paper based on evidence reviewed for the National Institute for Health and Care Excellence Safe Staffing guideline development. *Int J Nurs Stud* 2016;63:213–25.
- Fasoli DR, Fincke BG, Haddock KS. Going beyond patient classification systems to create an evidence-based staffing methodology. *J Nurs Adm* 2011;41:434–9.
- Cho SH, Hwang JH, Kim J. Nurse staffing and patient mortality in intensive care units. *Nurs Res* 2008;57:322–30.
- Kerlin MP, Small DS, Cooney E, *et al.* A randomized trial of nighttime physician staffing in an intensive care unit. *N Engl J Med* 2013;368:2201–9.