

PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Nurse staffing and patient-perceived quality of nursing care: A cross-sectional analysis of survey and administrative data in German hospitals
AUTHORS	Winter, Vera; Dietermann, Karina; Schneider, Udo; Schreyögg, Jonas

VERSION 1 – REVIEW

REVIEWER	Wang, Ruoxi Huazhong University of Science and Technology Tongji Medical College, School of Medicine and Health Management
REVIEW RETURNED	26-May-2021

GENERAL COMMENTS	<p>Thank you very much for this opportunity to review the manuscript entitled “Nurse staffing and patient-perceived quality of nursing care: A cross-sectional analysis of survey and administrative data in German hospitals”. The authors explored the association between nurse staffing and patient-perceived quality of nursing care using a multi-dimensional survey design, which may shed light on hospital management practice. I hope the following comments may help improve the quality of this manuscript:</p> <ol style="list-style-type: none">1. Methods: the authors used two items to measure each of the three dimensions of quality of care. Why using the arithmetic mean rather than the total score to measure quality of care?2. Statistical model: the authors reported that they employed a fixed effects model to address the endogeneity issues. To the best of my knowledge, a fixed effect model is used in longitudinal designs, where it measures the relationship between Δy and Δx by using within individual comparison (Ref: https://en.wikipedia.org/wiki/Fixed_effects_model). In this case, the model should contain time variable. If the authors intended to measure the “group effect” of unite type, then I guess a mixed model is more appropriate, as the presented model measured both i and u and accounted for individual level effect.3. Statistical model: the authors conducted several subgroup analyses (patient severity, hospital size and type of department) considering the potential moderating effects. Why not conducting moderating analysis to measure how these factors moderated the association between nurse staffing and patient-perceived quality of care?4. Table 2: please present the results in a form of Mean (SD) for continuous variable that follows normal distribution. Table 3: please replace $p=0.000$ with $p<0.001$.
-------------------------	--

	5. Table 2, the results seemed to be contradictory to the statement “For instance, for dimensions one and two, the average responses for patients discharged from internal medicine were 0.24 and 0.46 scale points higher, respectively, compared to orthopedic patients.”(P10 Lines 20-23).
--	--

REVIEWER	Zhang, Yuxia
REVIEW RETURNED	Zhongshan Hospital Fudan University, Department of Nursing 14-Jun-2021

GENERAL COMMENTS	<p>1. Although the sample size was large, the response rate was only 14.2%, how did you control patients' reporting bias? we are not sure whether the data of quality of care you collected was representative, maybe only those patients who were satisfied or dissatisfied about the quality of care would participate the survey.</p> <p>2. How did you ensure the quality of online survey?</p> <p>3. It may seem more reasonable if you only use the data of 2019 to calculate occupation days since the information of patients' perceived quality of care was originated from the year of 2019.</p> <p>4. As you mentioned in the article, you aimed to analyze the non-linear relationship between the nursing staff and quality of care, could you provide the cutoff? I mean since the result showed that "if a nurse already has to care for a lot of patients, one additional patient has a less strong effect on quality perceptions", could you specify the number of patients, which would provide more practical and useful information to hospital managers.</p>
-------------------------	---

VERSION 1 – AUTHOR RESPONSE

Responses to Reviewers

Reviewer: 1

Thank you very much for this opportunity to review the manuscript entitled “Nurse staffing and patient-perceived quality of nursing care: A cross-sectional analysis of survey and administrative data in German hospitals”. The authors explored the association between nurse staffing and patient-perceived quality of nursing care using a multi-dimensional survey design, which may shed light on hospital management practice. I hope the following comments may help improve the quality of this manuscript:

1. Methods: the authors used two items to measure each of the three dimensions of quality of care. Why using the arithmetic mean rather than the total score to measure quality of care?

Response: The number of items per dimension varies; we reported two exemplary items per dimension and now added the exact number of items per dimension in the table. We decided to use the arithmetic mean to remain on the scale of the items (1-5 and 1-4, respectively) and to have a higher comparability across the dimensions. The scaling does not affect the significance of the results in any case.

2. Statistical model: the authors reported that they employed a fixed effects model to address the endogeneity issues. To the best of my knowledge, a fixed effect model is used in longitudinal designs, where it measures the relationship between Δy and Δx by using within individual comparison (Ref: https://en.wikipedia.org/wiki/Fixed_effects_model). In this case, the model should contain time variable. If the authors intended to measure the “group effect” of unite type, then I guess a mixed model is more appropriate, as the presented model measured both i and u and accounted for individual level effect.

Response:

Thank you for your remark. Deciding whether an effect should be fixed or random is not always straightforward, but the basic idea is that an effect is random if the levels/values the parameter takes correspond to a (preferably random) sample from the population, and we want to generalize our findings to that larger population. E.g., if we have a large sample of nursing units or hospitals thought to be representative of the larger population of nursing units/hospitals, we typically include a random nursing unit intercept and/or random hospital intercept. The value of the random intercept is then estimated for each unit or hospital; each gets its own, “custom” parameter value.

By contrast, if we have *characteristics* of patients, units, or hospitals, we treat those as fixed (or “population-averaged” effects) and don’t estimate a separate value of the parameter for each patient/unit/hospital. E.g., we treat hospital size, unit type, and patient age and sex as fixed effects because the levels/values they take don’t correspond to a random sample of such levels/values, and we don’t estimate a separate effect for each hospital’s bed size, each unit’s type, etc. With a categorical fixed effect like unit type we *do* estimate an effect for all the levels but one (the referent level), but the key is that we don’t consider these various levels to be a sample of such levels drawn from a larger population. In the case of this study, the 24 unit types are not a random sample of unit types, and we wouldn’t expect the values of their regression coefficients to be normally distributed (the assumption we make regarding random intercepts in generalized and linear mixed models). Rather, these are presumably 24 common unit types, and we have little interest in generalizing to a larger population of unit types.

For that reason, we opted for a fixed effects model. As we acknowledge that this is a theoretical reasoning subject to discussion, we ran a random effects model with unit types as sensitivity analysis and show that results are stable and independent of whether fixed or random effects are used. we added this in the manuscript and in the appendix (Table A.3).

3. Statistical model: the authors conducted several subgroup analyses (patient severity, hospital size and type of department) considering the potential moderating effects. Why not conducting moderating analysis to measure how these factors moderated the association between nurse staffing and patient-perceived quality of care?

Response: Thank you for your comment. We indeed performed moderating analyses as suggested. However, this pushed our data sets to its limits and we often experienced a sample size issue in these analyses. E.g., to investigate whether the effects of staffing depend on patient severity, we would need to include at least four, yet up until 20 interaction terms (PTN, PTN2, skill mix, and patient-to-physician ratio interacting with PPCL as continuous or as categorical variable with six levels). As some of those models did not converge and others were highly cumbersome to interpret, we decided for the sake of comprehensibility and interpretability of the models to keep the subgroup analyses and mention this in the limitations section.

4. Table 2: please present the results in a form of Mean (SD) for continuous variable that follows normal distribution. Table 3: please replace $p=0.000$ with $p<0.001$.

Response: Thank you for your comment. We added standard deviations for the QoNC dimensions in Table 2 and replaced $p=0.000$ with $p<0.001$ in Table 3 and all regression tables in the Appendix.

5. Table 2, the results seemed to be contradictory to the statement “For instance, for dimensions one and two, the average responses for patients discharged from internal medicine were 0.24 and 0.46 scale points higher, respectively, compared to orthopedic patients.”(P10 Lines 20-23).

Response: Thank you for pointing out this mistake. We corrected it to “lower” instead of “higher”.

Reviewer: 2

1. Although the sample size was large, the response rate was only 14.2%, how did you control patients' reporting bias? we are not sure whether the data of quality of care you collected was representative, maybe only those patients who were satisfied or dissatisfied about the quality of care would participate the survey.

Response: Thank you for your comment. The response rate is comparable to other large-scale patient surveys. We checked for representativeness of the study population. Compared to the general population of hospitalized patients in Germany, our sample is generally representative in observable characteristics. However, the share of patients older than 80 is lower in our sample compared to the general population of hospitalized patients in Germany. We also compared respondents to non-respondents and did not find any substantial deviations in observable characteristics. We report this in the methods section and mention it as a limitation of our study.

2. How did you ensure the quality of online survey?

Response: thank you for your comment. To ensure the quality of the survey, we followed the scientific standards for scale development. The entire development and validation of the survey is described elsewhere (Blume et al. 2021). To sum up our proceeding, we drew on a systematic literature search and expert interviews to derive our initial items. We conducted two pre-tests (one paper and pencil pre-test and one online pre-test) with different participants, collected, discussed, and reported all changes made to the survey. After data collection, we performed comprehensive exploratory and confirmatory factor analyses to ensure the validity and reliability of the survey. We added this information in the methods section of the manuscript.

3. It may seem more reasonable if you only use the data of 2019 to calculate occupation days since the information of patients' perceived quality of care was originated from the year of 2019.

Response: For the number of inpatients, we use the information of 2019 only from the Quality reports. However, for average length of stay, we use data from the years 2014-2019, because our underlying set of claims data covers only a subset (albeit a substantial one) of all hospital patients in Germany and this might induce bias. By using a five-years average, we reduce variance and achieve an improved model fit and increased robustness. We ensured that there are no systematic changes in the length of stay for individual hospital-unit combinations u_h across years (i.e., the average length of stay in hospital units is stable across the time period).

4. As you mentioned in the article, you aimed to analyze the non-linear relationship between the nursing staff and quality of care, could you provide the cutoff? I mean since the result showed that "if a nurse already has to care for a lot of patients, one additional patient has a less strong effect on quality perceptions", could you specify the number of patients, which would provide more practical and useful information to hospital managers.

Response: Thank you for your comment. We added a figure in the appendix (A.2) to illustrate the non-linear relationship in (for factor 1). As you can see, the threshold is around 16 patients per nurse, which is quite far away from our mean PTN and in an area where data points become very limited. Hence, it is rather unlikely that the effect of an additional patient per nurse will in fact turn positive. But we see that the effect of an additional patient per nurse if the PTN increases from 3 to 4 is substantially more detrimental to the QoNC rating than if PTN increases from 13 to 14.

REFERENCES

Blume K, Kirchner-Heklau U, Winter V, et al. Construction and validation of a scale to measure the quality of nurse delivered care in acute care settings. *Working Paper*. 2021

VERSION 2 – REVIEW

REVIEWER	Zhang, Yuxia
REVIEW RETURNED	Zhongshan Hospital Fudan University, Department of Nursing 20-Sep-2021
GENERAL COMMENTS	It seems more reasonable if authors adjust some other important variables, such as hospital types and clinical units types, because these variables may cause potentially important influence on the quality of care

VERSION 2 – AUTHOR RESPONSE

Reviewer: 2

Dr. Yuxia Zhang, Zhongshan Hospital Fudan University Comments to the Author:

It seems more reasonable if authors adjust some other important variables, such as hospital types and clinical units types, because these variables may cause potentially important influence on the quality of care

Response: Thank you for your comment. We control for several variables which potentially affect quality perceptions, including hospital types (by hospital size and location, i.e., rurality) and hospital unit types (the 24 different types shown in table 1), and mention this below the table: Fixed effects for the 24 unit types, months, rurality, and bed categories and between-unit effects for patient-to-nurse and skill mix included but not shown. We now write this more explicitly in our methods section and revised the information below the tables.