BMJ Open

Quality of Life of Patients with Dementia in Acute Hospitals: Comparing a regular ward to a special care ward with dementia care concept. A Bayesian Multilevel Model Analysis.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-030743
Article Type:	Research
Date Submitted by the Author:	01-Apr-2019
Complete List of Authors:	Lüdecke, Daniel; University Medical Center Hamburg, Department of Medical Sociology Poppele, Georg; Evangelisches Krankenhaus Alsterdorf Klein, Jens Kofahl, Christopher
Keywords:	Dementia < NEUROLOGY, Quality of Life, Acute Hospital, Quality of Care, INTERNAL MEDICINE

Title

Quality of Life of Patients with Dementia in Acute Hospitals: Comparing a regular ward to a special care ward with dementia care concept. A Bayesian Multilevel Model Analysis.

Authors

Daniel Lüdecke, Georg Poppele, Jens Klein, Christopher Kofahl

Corresponding author

Daniel Lüdecke
Department of Medical Sociology
University Medical Center Hamburg-Eppendorf
Martinistr. 52
20246 Hamburg
Germany
Email: d.luedecke@uke.de
ORCID iD: 0000-0002-8895-3206
Tel.: ++49 40 7410 52878

Co-authors

Georg Poppele, Evangelisches Krankenhaus Alsterdorf (Protestant Hospital Alsterdorf), Hamburg, Germany

Jens Klein, University Medical Center Hamburg-Eppendorf, Department of Medical Sociology, Hamburg, Germany

Christopher Kofahl, University Medical Center Hamburg-Eppendorf, Department of Medical Sociology, Hamburg, Germany

Keywords

Dementia, Quality of Life, Acute Hospital, Quality of Care, Internal Medicine

Word Count: 4.230 (excluding title page, abstract, strength and limitations, references, tables)

Abstract

Objectives: To identify factors that predict the quality of life (QoL) of patients with dementia in acute hospitals and to analyse if a special care concept can increase the patients' QoL.

Design: A non-randomised case-control-study, including two internal medicine wards from hospitals in Hamburg, Germany.

Setting and Participants: In all, 526 patients with dementia from two hospitals were included in the study (intervention: n=333; control: n=193). Inclusion criteria was an at least mild cognitive impairment or dementia. The intervention group was a hospital with a special care ward for internal medicine focussing on patients with dementia. The control group was from a hospital with a regular care ward without special dementia care concept.

Outcome Measures: Our main outcome was the QoL (range 0-100) from patients with dementia in two different hospitals. A Bayesian multilevel analysis was conducted to identify predictors like age, dementia, agitation, physical and chemical restraints or functional limitations that affect QoL.

Results: QoL differs significantly between the control (40.7) and intervention group (51.2), $p<0.001$. Regression analysis suggests that physical restraint (estimated effect: -5.0), psychotropic drugs use (-4.4) and agitation (-2.9) are negatively associated with QoL. After controlling for confounders, the positive effect of the special care concept remained (5.7).

Conclusions: A special care ward will improve the quality of care and has a positive impact on the QoL of patients with dementia. Health policies should consider the benefits of special care concepts and develop incentives for hospitals to improve the QoL and quality of care for these patients.

Article Summary

Strengths and Limitations

- This is one of the first studies to investigate quality of life of patients with dementia in acute internal medicine wards in Germany.
- Study results suggest that a special care concept leads to a clinically relevant improvement in quality of life for patients with dementia.
- The statistical method applied in this study explicitly incorporates and accounts for information from previous research.
- The structural differences between the hospitals from the control and intervention may limit the generalizability of the results regarding the benefit of special care concepts for regular internal medicine wards.

Introduction

Acute hospitals face the challenge of an increase in old-age patients, which particularly affects internal medicine wards [1]. The average age of patients in internal medicine wards is above 70 years and may even come close to 80 years [2,3], leading to an increasing prevalence of cognitive impairments [4]. Although data on the occurrence of dementia or cognitive impairments of patients in hospitals is inconsistent, the larger proportion of studies report prevalence rates of about 40% [5–7].

Many hospitals are insufficiently prepared for patients with cognitive impairments, especially in acute care units predominantly focussing on somatic diseases [8]. Patients with cognitive impairments or dementia do not fit into the typical routines and standardised workflows of hospitals as these patients need more resources for care and treatment [9,10]. These patients often become disoriented, anxious, agitated and challenge hospital staff with erratic behaviour when placed in regular care wards. This results in an increased likelihood of falls, complications during the hospital stay and post-operative complications [11,12]. Hence, patients with dementia are a vulnerable group with a higher risk for long-lasting functional impairments [13,14].

To control the challenging behaviour of patients with dementia, a common practice – at least in German acute hospitals, and especially for older patients – is to use physical restraints like side rails to keep patients in bed, or chemical restraints such as the on-demand-use of psychotropic medication or hypnotics and sedatives [15,16]. This practise is not limited to German hospitals; however, there seems to be large variations between countries [17]. Physical and chemical restraints as well as anxiety and challenging behaviour are associated with poorer outcomes in quality of life [QoL] and quality of care of patients with dementia [18,19]. Therefore, hospitals in general, and specifically internal medicine wards with their increasing proportion of patients with

dementia, need to address these issues in order to improve the quality of care for these patients.

At least in Germany, there were lately no care concepts that fully address the needs of patients with dementia in internal medicine [20]. The special care ward "DAVID" in the Protestant Hospital Alsterdorf in Hamburg was one of the first internal medicine wards in Germany that implemented a comprehensive care concept for patients with dementia, aiming to improve the patients' QoL during their hospital stay. QoL is an important indicator of quality of care and a major dimension when assessing patient reported outcomes, particularly in older people as global outcome measure for interventions [21,22]. The assumption of this care concept is that a special care ward for patients with dementia leads to better outcomes in QoL compared to regular internal medicine wards. A study ("DAVID 2") was conducted to investigate the impact of such a care concept. This paper shows the results of this study and addresses two research questions. First, which factors predict the QoL of patients with dementia in acute hospitals? Second, beyond these factors, can a special care concept for patients with dementia in acute hospitals increase the patients' QoL?

Methods

Study Design and Setting

The aim of this study was to compare the quality of care for patients with dementia within a specialised dementia care concept as opposed to regular care in acute hospitals. The present study was designed as a non-randomised case-control-study, including two internal medicine wards in two hospitals located in Hamburg, Germany. The intervention group was a hospital that implemented a special care ward for internal medicine focussing on patients with dementia. The control group was from a hospital

with a regular care ward for internal medicine, which had no special dementia care concept.

Intervention Group

The special care ward "DAVID" is an internal medicine ward in the Protestant Hospital Alsterdorf and has 14 beds. In the year of data collection (2016), 349 patients were treated. The ward employed nine care workers as nursing staff.

Key components of the special care concept are a) a specific architectonical design, including a homelike lounge, a specific colouring of doors and walls, and a light concept with minimum 500 lux at eye level; b) doctors, care workers and service staff are trained in coping with challenging behaviour and other dementia related issues; c) mobile devices for diagnostics, to perform as many treatments as possible in the different rooms of the special care ward; d) involvement of relatives into assessment, care and discharge planning; and e) regular therapeutic offers like occupational or speech therapy, and social offers like music, playing or spending more time than usual to care for the patients.

To fulfil these high standards of quality of care, the ward "DAVID" employs more care staff in relation to the number of patients as compared to other regular internal medicine wards.

Control Group

The regular care ward is part of a larger hospital with emergency hospitalisation. It has 80 beds and in the year of data collection, about 3.500 patients were treated in this internal medicine ward. Twenty-six employees worked as care staff in this ward. Trainees sometimes supported the care team. The regular care ward had no specific care

concept for dementia patients. The care staff was not particularly trained in dementia topics.

Data collection and participants

An assessment questionnaire was developed to obtain data from patients with dementia. Study nurses were trained in using this assessment questionnaire and then conducted the data collection in both hospitals. Two study nurses were responsible for the special care ward and one for the regular care ward. A pre-test of two months was conducted to test and revise the questionnaire. As a result, some items were removed and instructions for study nurses were defined more precisely. After the pre-test, data was collected over a period of 12 months. To detect small to medium effect sizes (Cohen's d ~ 0.1 to 0.2), a power analysis was performed prior to the data collection and yielded a sample size of at least 173 subjects per group. Patients were included when they showed at least mild cognitive impairments or memory problems. In the special care ward (intervention group) all patients were assessed because a diagnosed dementia was a requirement for admission to that hospital. Hence, the participation rate for the special care ward was about 94% and excluded only a few patients that were not responsive. For the regular care ward (control group), a short dementia screening was used to assess the severity of dementia, in order to identify patients who qualify for the study [23]. This was necessary because not all patients have had a clarified dementia diagnosis. The total sample size for the present analysis consists of N=526 patients (special care ward: n=333; regular care ward: n=193).

Prior to the study, a study protocol was developed and submitted to the ethical committee of the medical association of Hamburg. The ethical committee approved the proposal and attested that the study conforms to ethical and legal requirements

(approval code PV5102). Study participants were not able to give their informed consent due to their cognitive impairments. However, as data mostly derived from the hospitals' regular documentation and was completely anonymous, the ethics committee waived the need of an informed consent.

Patient and Public involvement

Patients and the public were not involved in the development of the research question nor study design.

Measures

Outcome: QoL in patients with dementia was assessed using the QUALIDEM [24,25]. After observing patients for about one week (depending on the length of stay), the study nurses rated their QoL. QUALIDEM comprises 37 items reflecting nine different subdomains of QoL: "care relationship" (7 items, 0-21 points), "positive affect" (6 items, 0-18 points), "negative affect" (3 items, 0-9 points), "restless and tense behaviour" (3 items, 0-9 points), "positive self-image" (3 items, 0-9 points), "social relations" (6 items, 0-18 points), "social isolation" (3 items, 0-9 points), "feeling at home" (4 items, 0-12 points) and "have something to do" (2 items, 0-6 points). For patients with very severe dementia (Minimental State Examination Test [26] [MMSE] < 7), only six of the nine subscales apply, where the dimensions "positive self-image", "feeling at home" and "have something to do" were omitted. The recommendation is to report descriptive results of the QUALIDEM separately for each subscale. For regression analyses, a QoL index was calculated, ranging from 0 to 100 points with a higher score indicating better QoL. The QUALIDEM total score applies to all severities of dementia, so all patients' scores are comparable [27].

Independent Variables: Age, gender, main diagnosis for admission to hospital and length of stay were recorded. A modified version of the Charlson's Comorbidity Index [CCI], which included depression and hypertension as new items, was built based upon the assessment of comorbidities and chronical diseases [28,29]. If patients had no chronic illnesses, the CCI had a score of zero points. Else, higher scores indicated more serious comorbid disease. Shortly after admission to hospital, the study nurses measured functional limitations and cognitive status of patients. Functional limitations in daily living were assessed with the Barthel-Index [30]. This score ranges from 0 (completely dependent) to 100 points (no basic functional limitations) and was recoded according to the classification of the ICD-10 [31] (German adaption) into a score from 1 to 6 points. The Minimental State Examination Test [26] [MMSE] measures the cognitive impairments of patients, ranging from 0 (very strong cognitive impairments) to 30 (very mild or no cognitive impairments) points. This score was recoded into three categories, also based on ICD-10 classification: severe dementia (0-16), moderate dementia (17-23 points) and mild dementia (24-27 points).

After about one week of hospital stay, the study nurses rated the patients' agitation and challenging behaviour and recorded psychotropic drug use (chemical restraint) and physical restraints. Agitation and challenging behaviour of patients was assessed using the Pittsburgh Agitation Scale [PAS] [32] ranging from 0 to 16 points (higher scores indicate stronger agitation).

Physical restraints were defined as the use of one the following measures: Side rails to keep a patient in bed, tying a patient to a bed, and use of "therapeutic" chairs that prevent patients to stand up. The variable was dichotomised, indicating whether patients (in the course of the hospital stay) were mechanically restrained by at least one of these measures or not.

Psychotropic drug use was defined as on-demand-use ("as-needed") of medication for the nervous system by means of the Anatomical Therapeutical Chemical (ATC) classification [33] and comprises antipsychotics, anxiolytics, hypnotics/sedatives and antidepressants (N05A-C, N06A). Indicated use of psychotropic drugs was defined as medications that were prescribed for regular, not on-demand-use and not only given to patients in order to control their challenging behaviour. Such use of psychotropic drugs was excluded from the analysis. The on-demand-use variable was dichotomised and shows whether, during the complete hospital stay, chemical restraints were applied to patients or not.

While these variables already cover many different aspects that have an effect on the QoL, a dummy variable for the hospitals used as proxy for the intervention estimates the impact of the special care concept. This should reflect how much of the change in QoL is attributable to the special care concept.

Missing Data

In total, 11% of individual items across all scales were missing (at random). The missing data pattern was analysed and missing data was imputed using the multivariate imputation by chained equations method [34], using 11 imputation steps corresponding to the proportion of missing data [35]. The method for imputing missing values depends on the variable's nature. For continuous variables, predictive mean matching was applied, while logistic regressions were used for binary variables.

Statistical Methods

Descriptive results for the total sample and each hospital are reported. Statistically significant differences of $p<0.05$ between the two hospital wards were tested using t-

tests, χ^2-tests or Mann-Whitney-U-tests, depending on the level of measurement and distribution of variables. Differences between the hospitals in the QUALIDEM subscales are presented as boxplots, showing the median value and upper and lower quartiles of the value distribution.

As multivariate analysis, a Bayesian linear mixed model was applied to analyse the associations between the independent variables and the outcome. Computations were based on Stan [36], a probabilistic programming language for specifying Bayesian models, using Markov Chain Monte Carlo sampling (in particular, Hamiltonian Monte Carlo) [37]. We assume that the patients' main diagnosis is associated with different degrees of physical impairments, which affect the QoL. Therefore, the variable 'main diagnosis' was used as level-2 unit (*random intercept*) in the multilevel model to control for the variation in the outcome. We used informative priors for the predictors age, female gender, severe dementia, psychotic drug use and physical restraints, based on information from former research [18,38,39]. Weakly informative priors were used for the remaining predictors. The prior and posterior distributions of the model are summarised in the supplemental material (see Supplementary File 1).

Continuous predictors were centred before entering the model. Age was divided by 10, so a 1-unit change in the predictor of age reflects a change of 10 years in patients. The median value of the posterior distribution is used as "Bayesian point estimate", which minimises the difference of estimates from true values over posterior samples, but there are many other plausible values (the "posterior distribution") to describe the association between predictors and outcome. Hence, 50% and 89% highest density intervals [40] [HDI] are shown to indicate the range of most credible values and to reflect the (un-)certainty of the estimates. The intraclass correlation coefficient [41] was calculated to see how much of the proportion of the variance in the outcome can be explained by the

grouping structure ('main diagnosis'). We developed post-hoc additional regression models with interaction terms for need predictors (Barthel-Index, physical and chemical restraints, PAS-Score) to check if the associations between the complexity of patients' needs and QoL differ between hospitals. We found no significant interaction terms and decided to present the most parsimonious model here and show further results in the appendix (see Supplementary File 2).

All analyses were conducted with the R statistical package [42], including the packages *mice* [34], *ggplot* [43], *brms* [44] and *sjPlot* [45]. The source code is available in the supplemental material (see Supplementary File 3). Data is available online [46].

Results

Sample Characteristics

Table 1 gives an overview of the sample characteristics. The proportion of female to male patients is similar in both groups. The mean age is 4 years higher in the control group. There are also significant group differences in the Barthel-Index indicating higher functional impairment in the control group, while the dementia severity was the same in both hospitals. Comorbid conditions are slightly higher in the control group. Patients stayed 9.4 days in hospital on average and nearly one day longer in the intervention group as compared to the control group. Large differences between the two hospitals can be seen in the use of medical and physical restraints with significantly less use in the intervention group. Agitation- and QoL-scores also show strong group differences to the disadvantage of the patients in the control group.

Table 1: Sample Characteristics

Characteristic	**Control Group (Regular Care Ward, n=193)**	**Intervention Group (Special Care Ward, n=333)**	**Total (N=526)**	**p-value of difference**
Female, %	59.1	61.6	60.6	.637
Mean Age (SD)	83.1 (7.2)	79.0 (11.9)	80.5 (10.6)	**<.001**
Mean_Barthel-Index (SD)	29.9 (27.9)	40.7 (30.4)	36.7 (29.9)	**<.001**
Mild Dementia, %	7.8	9.6	8.9	.580
Moderate Dementia, %	30.0	26.7	27.9	.473
Severe Dementia, %	62.2	63.7	63.2	.805
Mean_Length of Stay (SD)	8.9 (7.5)	9.7 (5.5)	9.4 (6.3)	**.002**
Physical Restraints (yes), %	54.4	28.2	37.8	**<.001**
Psychotropic Drug Use (yes, as-needed), %	25.9	14.1	18.4	**.001**
Mean-Score Pittsburg Agitation Scale (SD)	3.9 (3.1)	3.0 (3.2)	3.3 (3.2)	**<.001**
Mean Charlson' Comorbitiy Index (SD)	3.2 (3.0)	2.5 (2.0)	2.8 (1.6)	**<.001**
Mean Qualidem Total Score (SD)	40.7 (14.5)	51.2 (17.2)	47.3 (17.0)	**<.001**

Barthel-Index: 0-100 (higher = better functioning); Dementia (MMSE score): mild: 24-27; moderate: 17-23; severe: <= 16; Pittsburg Agitation Scale: 0-16 (higher = stronger agitation and anxiety); Charlson' Comorbitiy Index: 0-9 (higher = more comorbidities); QUALIDEM: 0-100 (higher = better QoL)

Quality of life

Looking at the QoL for patients with severe to mild cognitive impairments (these are the ones, in which all subdomains of the QUALIDEM could be applied), there is a consistent pattern across all QUALIDEM-domains: Patients in the control group have a lower QoL compared to the intervention group. Except for the last subdomain ('having something to do'), all differences are statistically significant (Figure 1).

The same consistent pattern can be found for patients with very severe dementia symptoms (MMSE score < 7). Here, only the second of the six applied subdomains ('positive affect') does not differ significantly between intervention and control (Figure 2).

Predictors of quality of life

Figure 3 shows the results from the Bayesian mixed model. Three predictors are clearly negatively associated with QoL: physical restraint, psychotropic drug use and agitation (PAS-score). Physical restraint is associated with a 4.9-point decrease in QoL. With 50% probability, the QoL decreases by 4.1 to 5.8 points and with a chance of 89% by -7 to -2.8 points respectively. The application of psychotropic drugs as-needed shows similar results, with a posterior median of -4.4. The third clearly negative associated predictor is agitation, which shows a decrease in QoL of about 2.9 points for each additional point in the PAS-score.

Dementia and gender are not clearly associated with QoL. Neither are the length of hospital stay and the CCI.

The age of the patient correlates slightly positive with QoL, where an increase of 10 years means an increase of about 1.2 points in the QoL. The posterior median of the Barthel-Index is 2.0, so for a one-category change in functional impairments the QoL changes by two points. This means that patients with severe functional impairments differ by about 10 points in QoL compared to patients with no functional impairments.

Controlling for all other predictors, the intervention (special care ward) shows the strongest association with our outcome of interest, the patients' QoL. The posterior median is 5.7, and with an 89% probability, the credible values describing the effect of the intervention on QoL are within the range from 3.8 to 7.6.

The intraclass correlation coefficient of the model is rather low (0.01). This means, the 'main diagnosis' does not explain much of the variance in the patients' QoL and there is almost no regularization ("shrinkage") of estimated model parameters and no larger differences between hospitals according to the patients' needs, as indicated by their main diagnosis.

Discussion

The study reported in this paper sought to understand those factors that influence the QoL in patients with dementia and whether a special care concept for these patients performs better in this regard as opposed to regular care wards.

One of our main findings is that QoL differs significantly between the control and intervention group. We found substantial differences between the two hospitals in the patients' total QoL score in favour of the special care ward. Beyond the statistical significance, this finding also has a clinical impact. Studies suggest a change in 3 points for the Quality of Life – Alzheimer's Disease Scale [47], which has a range of 40 points, to be clinically relevant [48,49]. Transferred to the range of the QUALIDEM scale, a difference of about 7.5 points would be considered as an important improvement in QoL. Another indicator to evaluate the clinical relevance of a change in QoL is an increase of the score of half a standard deviation [50], which would be about 8.5 points for our data. Taking these reference points as a basis, we found evidence for the clinical relevant improvement in QoL of patients in a special care ward.

A second key finding is the identification of those factors that are clearly associated with QoL. The use of physical and chemical restraints, both happening more frequently in the control group, are associated with lower outcomes in QoL. This finding is in line with other studies that suggest a negative association between physical and chemical restraints and QoL [18,39] and explains why the regular care ward performs less good in this regard than the special care ward. Agitation was also negatively associated with QoL. This is understandable as agitation is an expression of anxiety and indisposition of people with dementia and typically occurs after admission to hospital. Furthermore, agitation is often a reason for psychotropic drug use or physical restraint and, thus, also negatively affects QoL [51,52]. Independent from these factors, the special care ward

itself shows the strongest impact on QoL, indicating that patients with dementia explicitly benefit from specialised care concepts. Other studies also report these benefits, both in a nursing home or hospital setting [53,54]. Since we controlled for patient characteristics like main diagnosis, age, functional limitations, chronic comorbidities, agitation, length of stay etc. in our model, we do not assume that the positive effect of the special care ward is completely a result of a biased sample between intervention and control group. Although the two compared hospitals differ in their structures and size, patients' characteristics are largely comparable between the samples in the control and intervention group. For instance, there is no substantial difference between the two hospitals regarding the relationship between functional impairments and physical restraints. Moreover, to see if the complexity of patients' need affect our findings, we calculated regression models with interaction terms between need factors moderated by hospitals (see Supplementary File 2). The association between complexity of needs and QoL is not significantly different between the intervention and control group. Based on our results we suggest that the special care concept mainly explains the differences in the QoL. Although it is certainly difficult to determine the exact effect of the special care concept on the patients' QoL, our findings seem plausible in the light of the key elements of this intervention. A higher ratio of care staff as to patients, smaller facilities or systematically trained employees can be considered essential for health care provision to patients with dementia and are much better conditions for less physical or chemical restraints, independent of the functional limitations of patients. The special care ward provides a more dementia-friendly interior design, including orientation and navigation aids and the use of light and colours, which are considered as important components to reduce agitation for patients with dementia [55]. These findings and conclusions are in line with other studies on hospital care that

suggest that an increased staff ratio or the implementation of multiple components, which particularly address the needs of patients with dementia, lead to reduced use of physical restraints and psychotropic drug use and improve the quality of care [56,57]. The special care ward benefits from a higher staff ratio, i.e. nurses have to care for fewer patients with dementia compared to the control group. While this is an intentional element of the concept, the downside is higher personnel costs. Only few studies investigated the follow-up costs for patients with dementia in home care settings after hospitalization. Costa et al. predicted additional monthly costs in home care of about 445 Euros due to increased agitation of patients with dementia [58]. Thus, if patients with dementia benefit from special care concepts and perceive better outcomes in quality of life and care, the increased costs for more care personnel may be compensated by reducing follow-up costs for the ambulatory care. However, further research is needed to give more exact projections of the increased costs and potential of saving money.

Another finding is that the severity of cognitive impairments, measured with the MMSE, is a rather improper indicator to represent the underlying problems of and with the dementia disease, as these factors were not consistently associated with QoL. Direct measures of the problems associated with dementia, as agitation or challenging behaviour, should be considered as well when it comes to investigate the QoL of patients with dementia.

Our study has several limitations. One concerns the structural differences between the two hospitals. The hospital with the special care ward is much smaller than the hospital that hosted the control group. A second control group or an intervention group in a hospital of a similar size as the hospital with the regular care ward may have permitted a more distinct comparison. However, since we accounted for many different patient

characteristics including functional status, comorbidities and behavioural problems, we assume that a bias due to patient selection mechanisms is rather low. Another structural difference between the intervention and control group that certainly affects the results are the different staff ratios. In the special care ward, nurses have to care for fewer patients than in the regular care ward. Although we assume that this aspect probably has the highest impact on the outcomes in QoL, this is not a "selection bias" but a core component of the intervention. A further limitation is possibly the first and thus rather exploratory use of the QUALIDEM assessment in a hospital setting. Although studies show reliable results of the QUALIDEM in nursing homes even for a short observation period of about one week [59], there are no studies that evaluate the reliability and validity for use in hospitals. We have done checks of internal consistencies, which showed that most subdomains of the QUALIDEM perform well with our data and are comparable to results from other validation studies [60]. This indicates that the use of the QUALIDEM is feasible for hospital research. However, we cannot give evidence on the interrater reliability apart from the intense training of the study nurses.

Conclusions

On the whole, we think that a special care ward will improve the quality of care and is effective regarding the positive impact on the QoL of patients with dementia. Our study showed that after controlling for different predictors, the intervention still has a perceptible effect concerning clinical important differences in our outcome of interest, the patients' quality of life. However, such improvements can only be achieved by implementing a concept with multiple components that address the explicit needs of patients with dementia. The implementation of a special care concept usually increases the costs for hospitals because it requires a higher staff-patient-ratio, regular training of

employees or more therapeutic offers. On the other hand, costs that accumulate in informal care after hospital stay as a result of poorer quality of care in hospitals can be much higher than additional personnel costs and could probably be reduced [58,61]. Health policies should consider the benefits of special care concepts and develop incentives for hospitals to improve the QoL and quality of care for patients with dementia.

Acknowledgement

The authors are grateful to the clinical staff of both hospitals who supported this study by helping with the data collection. Furthermore, we are thankful to the Authority of Health and Consumer Protection and the Homann-foundation for funding this study.

Author Contributions

DL was involved in conception and design of the study, analysis and interpretation of data and drafting the article. GP was involved in conception and design of the study and interpretation of data. JK was involved in interpretation of data and drafting the article. CK was involved in conception and design of the study, interpretation of data and drafting the article.

Competing interests

The authors have declared that no competing interests exist.

Funding

The present study was funded by the Behörde für Gesundheit und Verbraucherschutz (Authority of Health and Consumer Protection) in Hamburg and the Homann-foundation (grant number Z/65454/2014/F500-02.10/10,002). No one of the funders was involved in the data analysis process and did not influence or hold back results of the study.

Ethics approval and consent to participate

Prior to the study, a study protocol was developed and submitted to the ethical committee of the medical association of Hamburg ("Ethik-Kommission der Ärztekammer Hamburg"). The IRB of the ethical committee approved the proposal and attested that the study conforms to ethical and legal requirements (approval code PV5102). Study participants were not able to give their informed consent due to their cognitive impairments. However, as data mostly derived from the hospitals' regular documentation and was completely anonymous, the IRB of the ethics committee waived the need of an informed consent.

Availability of data and material

All data generated or analysed during this study are included are available in the Zenodo repository (DOI: 10.5281/zenodo.1479677) at https://doi.org/10.5281/zenodo.1479676.

References

1 Recine U, Scotti E, Bruzzese V, *et al.* The change of hospital internal medicine: a study on patients admitted in internal medicine wards of 8 hospitals of the Lazio area, Italy. *Ital J Med* 2015;9:252. doi:10.4081/itjm.2015.523

2 Raveh D, Gratch L, Yinnon AM, *et al.* Demographic and clinical characteristics of patients admitted to medical departments. *J Eval Clin Pract* 2005;11:33–44. doi:10.1111/j.1365-2753.2004.00492.x

3 Sonnenblick M, Raveh D, Gratch L, *et al.* Clinical and demographic characteristics of elderly patients hospitalised in an internal medicine department in Israel: Characteristics of elderly patients in medical wards. *Int J Clin Pract* 2007;61:247–54. doi:10.1111/j.1742-1241.2006.00925.x

4 Mukadam N, Sampson EL. A systematic review of the prevalence, associations and outcomes of dementia in older general hospital inpatients. *Int Psychogeriatr* 2011;23:344–55. doi:10.1017/S1041610210001717

5 Zekry D, Herrmann FR, Grandjean R, *et al.* Does dementia predict adverse hospitalization outcomes? A prospective study in aged inpatients. *Int J Geriatr Psychiatry* 2009;24:283–91. doi:10.1002/gps.2104

6 Lyketsos CG, Sheppard JM, Rabins PV. Dementia in elderly persons in a general hospital. *Am J Psychiatry* 2000;157:704–7. doi:10.1176/appi.ajp.157.5.704

7 Sampson EL, Blanchard MR, Jones L, *et al.* Dementia in the acute hospital: prospective cohort study of prevalence and mortality. *Br J Psychiatry J Ment Sci* 2009;195:61–6. doi:10.1192/bjp.bp.108.055335

8 Pinkert C, Holle B. Menschen mit Demenz im Akutkrankenhaus: Literaturübersicht zu Prävalenz und Einweisungsgründen. *Z Für Gerontol Geriatr* 2012;45:728–34. doi:10.1007/s00391-012-0319-1

9 Dutzi I, Schwenk M, Micol W, *et al.* Patienten mit Begleitdiagnose Demenz. *Z Für Gerontol Geriatr* 2013;46:208–13. doi:10.1007/s00391-013-0483-y

10 Timmons S, O'Shea E, O'Neill D, *et al.* Acute hospital dementia care: results from a national audit. *BMC Geriatr* 2016;16:113. doi:10.1186/s12877-016-0293-3

11 Pi H-Y, Gao Y, Wang J, *et al.* Risk Factors for In-Hospital Complications of Fall-Related Fractures among Older Chinese: A Retrospective Study. *BioMed Res Int* 2016;2016:1–11. doi:10.1155/2016/8612143

12 Hu C-J, Liao C-C, Chang C-C, *et al.* Postoperative Adverse Outcomes in Surgical Patients with Dementia: A Retrospective Cohort Study. *World J Surg* 2012;36:2051–8. doi:10.1007/s00268-012-1609-x

13 Motzek T, Junge M, Marquardt G. Einfluss der Demenz auf Verweildauer und Erlöse im Akutkrankenhaus. *Z Für Gerontol Geriatr* 2017;50:59–66. doi:10.1007/s00391-016-1040-2

14 Gross AL, Jones RN, Habtemariam DA, *et al.* Delirium and Long-term Cognitive Trajectory Among Persons With Dementia. *Arch Intern Med* 2012;172:1324. doi:10.1001/archinternmed.2012.3203

15 Nakanishi M, Okumura Y, Ogawa A. Physical restraint to patients with dementia in acute physical care settings: effect of the financial incentive to acute care hospitals. *Int Psychogeriatr* 2017;:1–10. doi:10.1017/S104161021700240X

16 Krüger C, Mayer H, Haastert B, *et al.* Use of physical restraints in acute hospitals in Germany: a multi-centre cross-sectional study. *Int J Nurs Stud* 2013;50:1599–606. doi:10.1016/j.ijnurstu.2013.05.005

17 Heinze C, Dassen T, Grittner U. Use of physical restraints in nursing homes and hospitals and related factors: a cross-sectional study: Physical restraints in nursing homes and hospitals. *J Clin Nurs* 2012;21:1033–40. doi:10.1111/j.1365-2702.2011.03931.x

18 Beerens HC, Sutcliffe C, Renom-Guiteras A, *et al.* Quality of Life and Quality of Care for People With Dementia Receiving Long Term Institutional Care or Professional Home Care: The European RightTimePlaceCare Study. *J Am Med Dir Assoc* 2014;15:54–61. doi:10.1016/j.jamda.2013.09.010

19 Beerens HC, Zwakhalen SMG, Verbeek H, *et al.* Factors associated with quality of life of people with dementia in long-term care facilities: A systematic review. *Int J Nurs Stud* 2013;50:1259–70. doi:10.1016/j.ijnurstu.2013.02.005

20 Kricheldorff C, Hofmann W. Demenz im Akutkrankenhaus. *Z Für Gerontol Geriatr* 2013;46:196–7. doi:10.1007/s00391-013-0484-x

21 Treurniet HF, Essink-Bot M-L, Mackenbach JP, *et al.* Health-related quality of life: an indicator of quality of care? *Qual Life Res* 1997;6:363–9. doi:10.1023/A:1018435427116

22 Valderas JM, Alonso J. Patient reported outcome measures: a model-based classification system for research and clinical practice. *Qual Life Res* 2008;17:1125–35. doi:10.1007/s11136-008-9396-4

23 Kaiser AK, Hitzl W, Iglseder B. Three-question dementia screening: Development of the Salzburg Dementia Test Prediction (SDTP). *Z Für Gerontol Geriatr* 2014;47:577–82. doi:10.1007/s00391-013-0568-7

24 Ettema TP, Dröes R-M, de Lange J, *et al.* QUALIDEM: development and evaluation of a dementia specific quality of life instrument. Scalability, reliability and internal structure. *Int J Geriatr Psychiatry* 2007;22:549–56. doi:10.1002/gps.1713

25 Dichter MM, Ettema TP, Schwab CGG, *et al.* Benutzerhandbuch für die deutschsprachige QUALIDEM Version 2.0. Witten: 2016.

26 Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;12:189–98.

27 Dichter MN, Meyer G. Quality of Life of People with Dementia in Nursing Homes. In: Schüssler S, Lohrmann C, eds. *Dementia in Nursing Homes*. Cham: Springer International Publishing 2017. 139–57. doi:10.1007/978-3-319-49832-4_11

28 Charlson ME, Pompei P, Ales KL, *et al.* A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;40:373–83.

29 Charlson ME, Charlson RE, Peterson JC, *et al.* The Charlson comorbidity index is adapted to predict costs of chronic disease in primary care patients. *J Clin Epidemiol* 2008;61:1234–40. doi:10.1016/j.jclinepi.2008.01.006

30 Mahoney FI, Barthel DW. Functional Evaluation: The Barthel Index. *Md State Med J* 1965;14:61–5.

31 Deutsches Institut f. medizinische Dokumentation u. Information (DIMDI). *ICD-10-GM 2017 Systematisches Verzeichnis*. Lich, Hess: pictura Werbung 2016.

32 Rosen J, Burgio L, Kollar M, *et al.* The Pittsburg Agitation Scale: A user-friendly instrument for rating agitation in dementia patients. *Am J Geriatr Psychiatry Off J Am Assoc Geriatr Psychiatry* 1994;2:52–9.

33 WHO Collaborating Centre for Drug Statistics Methodology, editor. *Guidelines for ATC classification and DDD assignment 2013.* Oslo: 2012.

34 Buuren S van, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. *J Stat Softw* 2011;45. doi:10.18637/jss.v045.i03

35 White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med* 2011;30:377–99. doi:10.1002/sim.4067

36 Carpenter B, Gelman A, Hoffman MD, *et al.* Stan: A Probabilistic Programming Language. *J Stat Softw* 2017;76. doi:10.18637/jss.v076.i01

37 Gelman A, Carlin JB, Stern HS, *et al. Bayesian data analysis.* Third edition. Boca Raton: CRC Press 2014.

38 Barca ML, Engedal K, Laks J, *et al.* Quality of Life among Elderly Patients with Dementia in Institutions. *Dement Geriatr Cogn Disord* 2011;31:435–42. doi:10.1159/000328969

39 Beer C, Flicker L, Horner B, *et al.* Factors Associated with Self and Informant Ratings of the Quality of Life of People with Dementia Living in Care Facilities: A Cross Sectional Study. *PLoS ONE* 2010;5:e15621. doi:10.1371/journal.pone.0015621

40 McElreath R. *Statistical rethinking: a Bayesian course with examples in R and Stan*. Boca Raton London New York: CRC Press, Taylor & Francis Group 2016.

41 Hox JJ. *Multilevel analysis: techniques and applications*. 2nd ed. New York: Routledge 2010.

42 R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing 2018. https://www.R-project.org/

43 Wickham H. *ggplot2: Elegant Graphics for Data Analysis.* 2nd ed. New York: Springer 2016.

44 Bürkner P-C. brms: An R Package for Bayesian Multilevel Models Using Stan. *J Stat Softw* 2017;80. doi:10.18637/jss.v080.i01

45 Lüdecke D. *sjPlot: Data Visualization for Statistics in Social Science.* 2018. doi:10.5281/zenodo.1308157

46 Lüdecke D. Dataset: Quality Of Life Of Patients With Dementia In Acute Hospitals-Version 1.0. 2018. doi:10.5281/zenodo.1479677

47 Thorgrimsen L, Selwood A, Spector A, *et al.* Whose quality of life is it anyway? The validity and reliability of the Quality of Life-Alzheimer's Disease (QoL-AD) scale. *Alzheimer Dis Assoc Disord* 2003;17:201–8.

48 Selwood A, Thorgrimsen L, Orrell M. Quality of life in dementia--a one-year follow-up study. *Int J Geriatr Psychiatry* 2005;20:232–7. doi:10.1002/gps.1271

49 Hoe J, Hancock G, Livingston G, *et al.* Changes in the quality of life of people with dementia living in care homes. *Alzheimer Dis Assoc Disord* 2009;23:285–90. doi:10.1097/WAD.0b013e318194fc1e

50 Norman GR, Sloan JA, Wyrwich KW. Interpretation of Changes in Health-related Quality of Life: The Remarkable Universality of Half a Standard Deviation. *Med Care* 2003;41:582–92. doi:10.1097/01.MLR.0000062554.74615.4C

51 Lam K, Kwan JSK, Wai Kwan C, *et al.* Factors Associated With the Trend of Physical and Chemical Restraint Use Among Long-Term Care Facility Residents in Hong Kong: Data From an 11-Year Observational Study. *J Am Med Dir Assoc* 2017;18:1043–8. doi:10.1016/j.jamda.2017.06.018

52 Ó Flatharta T, Haugh J, Robinson SM, *et al.* Prevalence and predictors of bedrail use in an acute hospital. *Age Ageing* 2014;43:801–5. doi:10.1093/ageing/afu081

53 Kok JS, Berg IJ, Scherder EJA. Special Care Units and Traditional Care in Dementia: Relationship with Behavior, Cognition, Functional Status and Quality of Life - A Review. *Dement Geriatr Cogn Disord Extra* 2013;3:360–75. doi:10.1159/000353441

54 von Renteln-Kruse W, Neumann L, Klugmann B, *et al.* Geriatric Patients With Cognitive Impairment. *Dtsch Ärztebl Int* 2015;112:103–12. doi:10.3238/arztebl.2015.0103

55 Büter K, Motzek T, Dietz B, *et al.* Demenzsensible Krankenhausstationen: Expertenempfehlungen zu Planung und Gestaltung. *Z Für Gerontol Geriatr* 2017;50:67–72. doi:10.1007/s00391-016-1079-0

56 Blair A, Anderson K, Bateman C. The "Golden Angels": effects of trained volunteers on specialling and readmission rates for people with dementia and delirium in rural hospitals. *Int Psychogeriatr* 2018;:1–10. doi:10.1017/S1041610218000911

57 Sinvani L, Warner-Cohen J, Strunk A, *et al.* A Multicomponent Model to Improve Hospital Care of Older Adults with Cognitive Impairment: A Propensity Score-Matched Analysis. *J Am Geriatr Soc* Published Online First: 11 August 2018. doi:10.1111/jgs.15452

58 Costa N, Wübker A, De Mauléon A, *et al.* Costs of Care of Agitation Associated With Dementia in 8 European Countries: Results From the RightTimePlaceCare Study. *J Am Med Dir Assoc* 2018;19:95.e1-95.e10. doi:10.1016/j.jamda.2017.10.013

59 Dichter MN, Schwab CGG, Meyer G, *et al.* Item distribution, internal consistency and inter-rater reliability of the German version of the QUALIDEM for people with mild to severe and very severe dementia. *BMC Geriatr* 2016;16. doi:10.1186/s12877-016-0296-0

60 Dichter M, Dortmann O, Halek M, *et al.* Scalability and internal consistency of the German version of the dementia-specific quality of life instrument QUALIDEM in nursing homes – a secondary data analysis. *Health Qual Life Outcomes* 2013;11:91. doi:10.1186/1477-7525-11-91

61 Leon J, Cheng CK, Neumann PJ. Alzheimer's disease care: costs and potential savings. *Health Aff Proj Hope* 1998;17:206–16.

Figure Titles and Legends

Figure 1: QUALIDEM subdomain scores by care ward in patients with mild to severe dementia (MMSE score from 7 to 27, n = 400)

Figure 2: QUALIDEM subdomain scores by care ward, patients with very severe dementia (MMSE score < 7, n = 126)

Figure 3: Predictors of Health Related Quality of Life, Regression Coefficients, Bayesian Linear Mixed Model, Posterior Median (+50% and 89% High Density Interval)

Supplementary Files

Supplementary File 1: Methodological comments, Word document (docx-format).

Supplementary File 2: Regression Models with Interaction Terms, Word document (docx-format).

Supplementary File 3: R Souce Code (to use with R statistics, CC BY-NC 4.0 license)

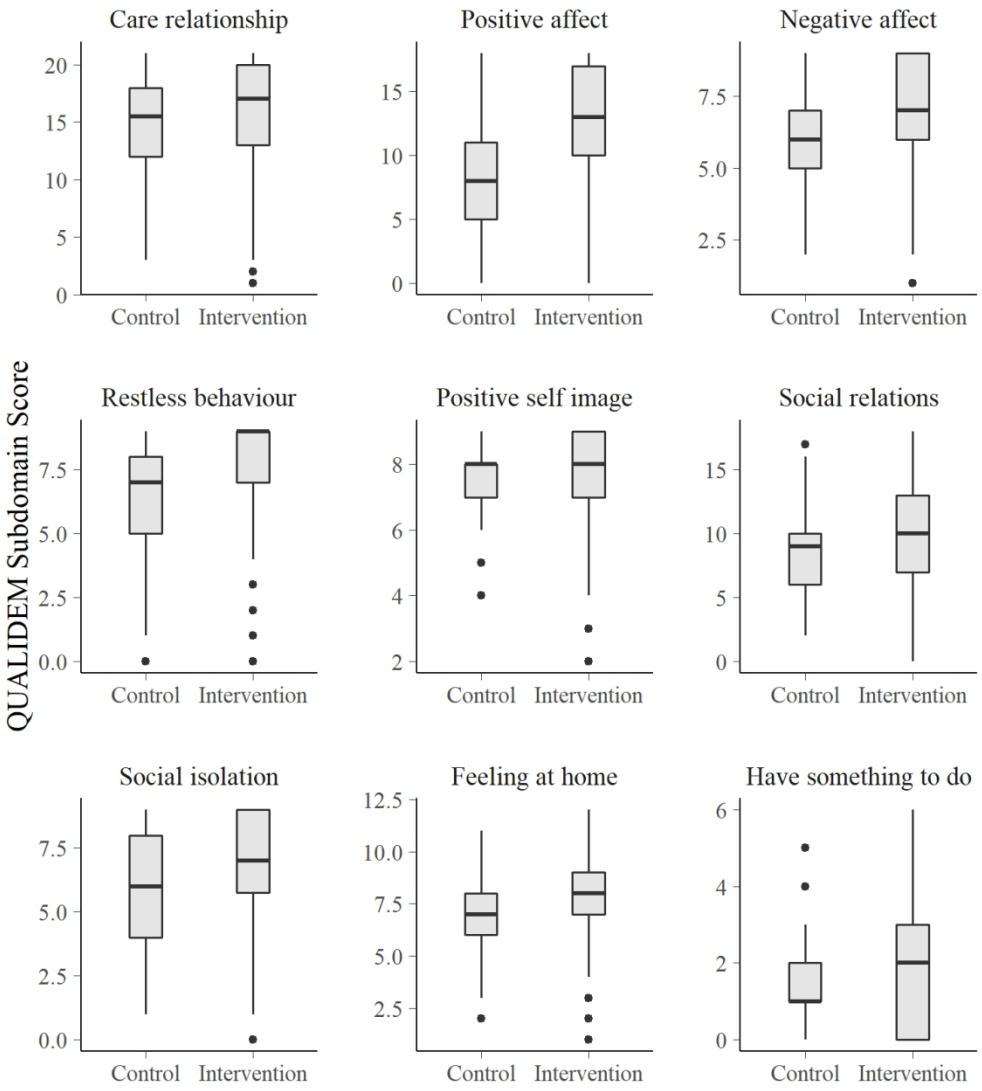


Figure 1: QUALIDEM subdomain scores by care ward in patients with mild to severe dementia (MMSE score from 7 to 27, n = 400)

169x189mm (220 x 220 DPI)

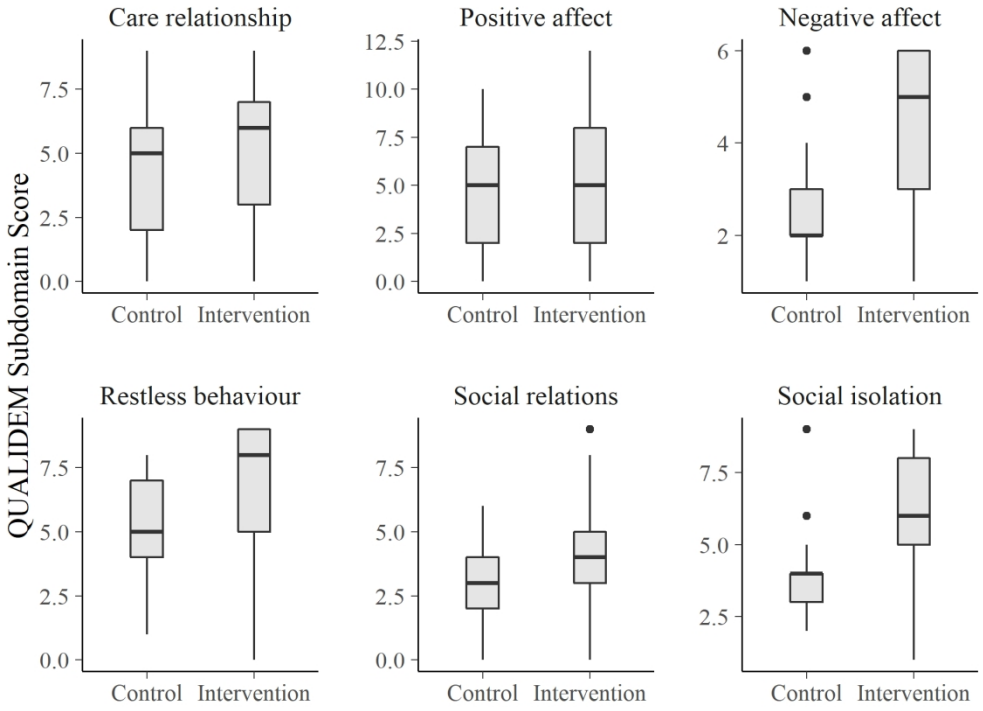


Figure 2: QUALIDEM subdomain scores by care ward, patients with very severe dementia (MMSE score < 7, n = 126)

169x125mm (300 x 300 DPI)

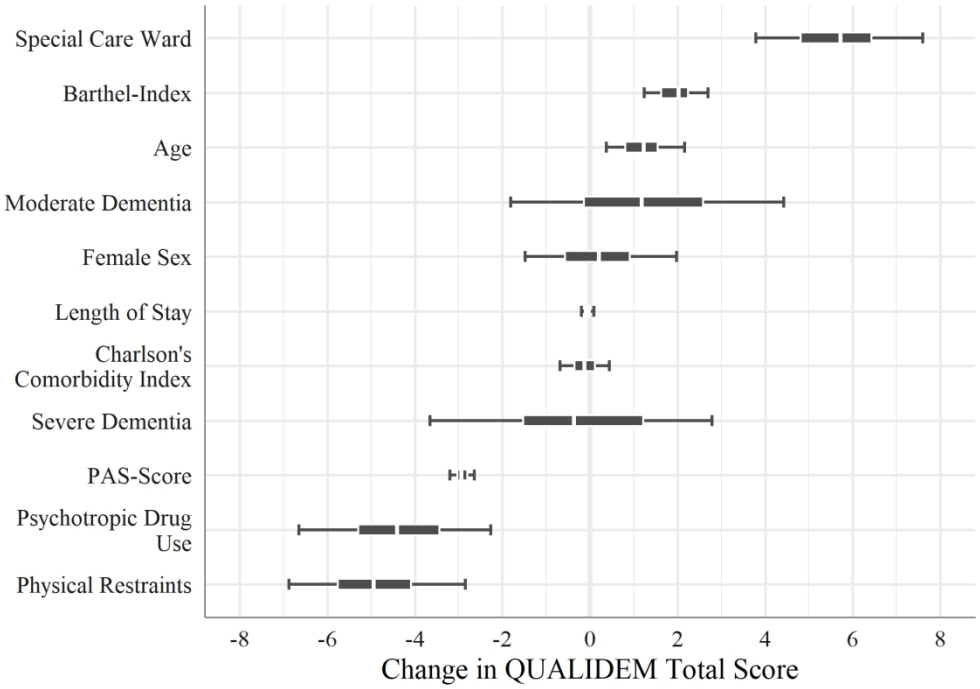


Figure 3: Predictors of Health Related Quality of Life, Regression Coefficients, Bayesian Linear Mixed Model, Posterior Median (+50% and 89% High Density Interval)

169x119mm (300 x 300 DPI)

Supplemental Material 1

Quality of Life of Patients with Dementia in Acute Hospitals: Comparing a regular ward to a special care ward with dementia care concept. A Bayesian Multilevel Model Analysis.

Methodological comments

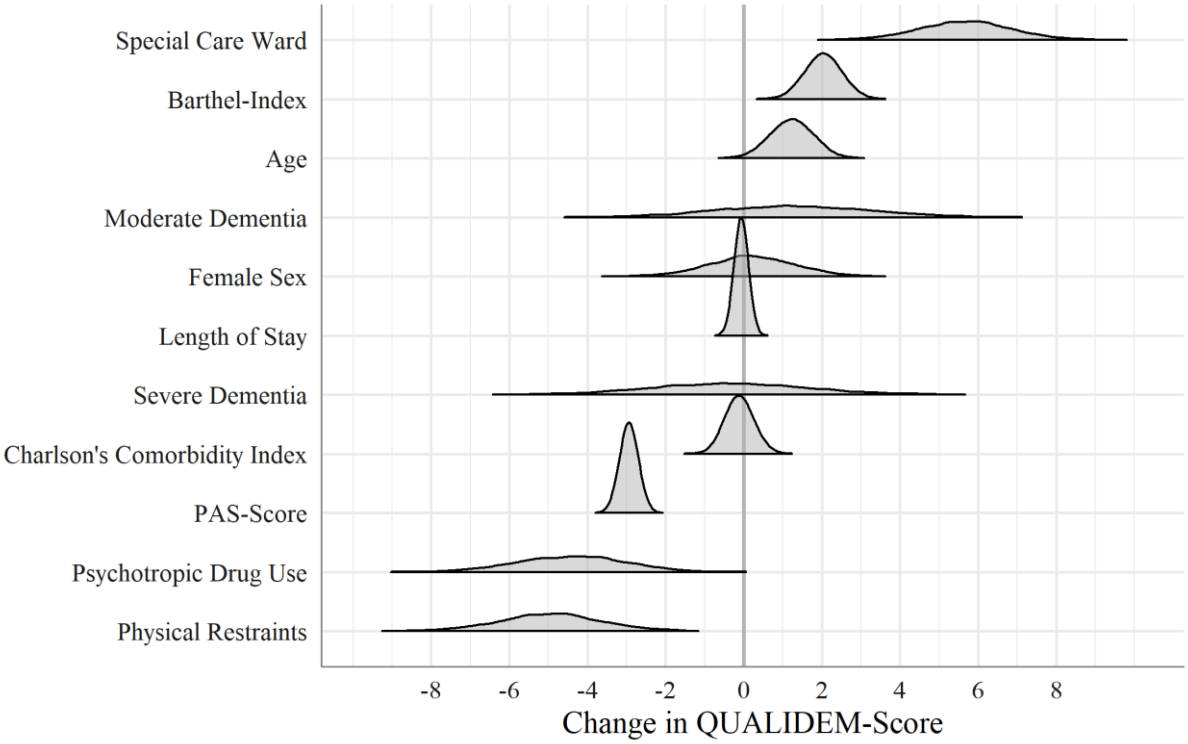
1. Short note on Bayesian Methods

Bayesian methods are probably not very common, and readers may ask "Why use Bayesian regression models?" Gelman et al.[1] give a well summarized answer to this question: "A pragmatic rationale for the use of Bayesian methods is the inherent flexibility introduced by their incorporation of multiple levels of randomness and the resultant ability to combine information from different sources, while incorporating all reasonable sources of uncertainty in inferential summaries. Such methods naturally lead to smoothed estimates in complicated data structures and consequently have the ability to obtain better real-world answers."[1]. Shortly speaking, advantages of Bayesian methods are, for instance, the ability to derive probability statements for every quantity of interest or explicitly incorporate prior knowledge about parameters into the model.

2. Distribution of Posterior Samples from the Regression Model

Contrary to a frequentist approach, there is no unique point estimate in Bayesian analysis. The posterior median minimises the expected absolute error (i.e. the difference of estimates from true values over posterior samples), but there are many other plausible values to describe the association between predictors and outcome, which is called the *distribution of posterior samples*. Credible or Highest Density Intervals are based on quantiles of this distribution and indicate the probability of an estimate being inside this interval. In terms of inference, this allows us to say that with, e.g. 50% probability, the true effect of a *Special Care Ward* on our outcome lies between 5.1 and 6.6. Figure S1 shows the distribution of posterior samples from our regression model.

Figure S1: *Distribution of Posterior Samples from Regression Model*



3. Tabular Summary from the Regression Model

Table S1 provides a summary of the regression model, including the data for two different Highest Density Intervals as well as the ratio of effective number of samples. The ratio should be 1.0 or close to 1.0 and indicates whether the samples were efficient, which means that the results for the related coefficient are reliable. Furthermore, all Rhat values of the models were approximately 1. The Rhat statistic[2] measures the ratio of the average variance of the draws within each chain to the variance of the pooled draws across chains. If the chains have not converged to a common distribution, the Rhat statistic will tend to be greater than one, indicating that more iterations may be required for reliable results.

Table S1: *Predictors of Quality of Life, Regression Coefficients, Bayesian Linear Mixed Model, Posterior Median (+50% and 89% Highest Density Intervals)*

Term	*Estimate*	*SE*	*50% HDI*	*89% HDI*	*Ratio of effective Number of Samples*
(Intercept)	46.2	2.2	45.2 – 48.2	42.7 – 49.8	1.00
Length of Stay	-0.1	0.1	-0.1 – 0.0	-0.2 – 0.1	1.00
Age	1.2	0.5	0.8 – 1.5	0.4 – 2.1	1.00
Moderate Dementia	1.2	2.0	-0.1 – 2.7	-1.8 – 4.6	1.00
Severe Dementia	-0.4	2.0	-1.6 – 1.1	-3.6 – 2.7	1.00
Female	0.2	1.1	-0.5 – 1.0	-1.6 – 1.9	1.00
Barthel Score	2.0	0.5	1.8 – 2.4	1.3 – 2.8	1.00
Physical Restraints (yes)	-4.9	1.2	-5.8 – -4.1	-7.0 – -2.8	1.00
Special Care Ward (Intervention)	5.7	1.2	4.9 – 6.5	3.8 – 7.6	1.00
PAS-Score	-2.9	0.2	-3.1 – -2.8	-3.2 – -2.7	1.00
Charlson's Comorbidity Index	-0.1	0.3	-0.4 – 0.1	-0.6 – 0.5	1.00
Psychotropic Drug Use (yes, as-needed)	-4.4	1.4	-5.3 – -3.5	-6.5 – -2.1	1.00
sigma	11.9	0.4	11.5 – 12.0	11.3 – 12.5	1.00

All Rhat values ~ 1, all mcse < 0.05. Results based on 4 chains each 1000 iterations for each chain. LOO-adjusted R^2: 0.500

4. Test for Practical Equivalence of Parameters

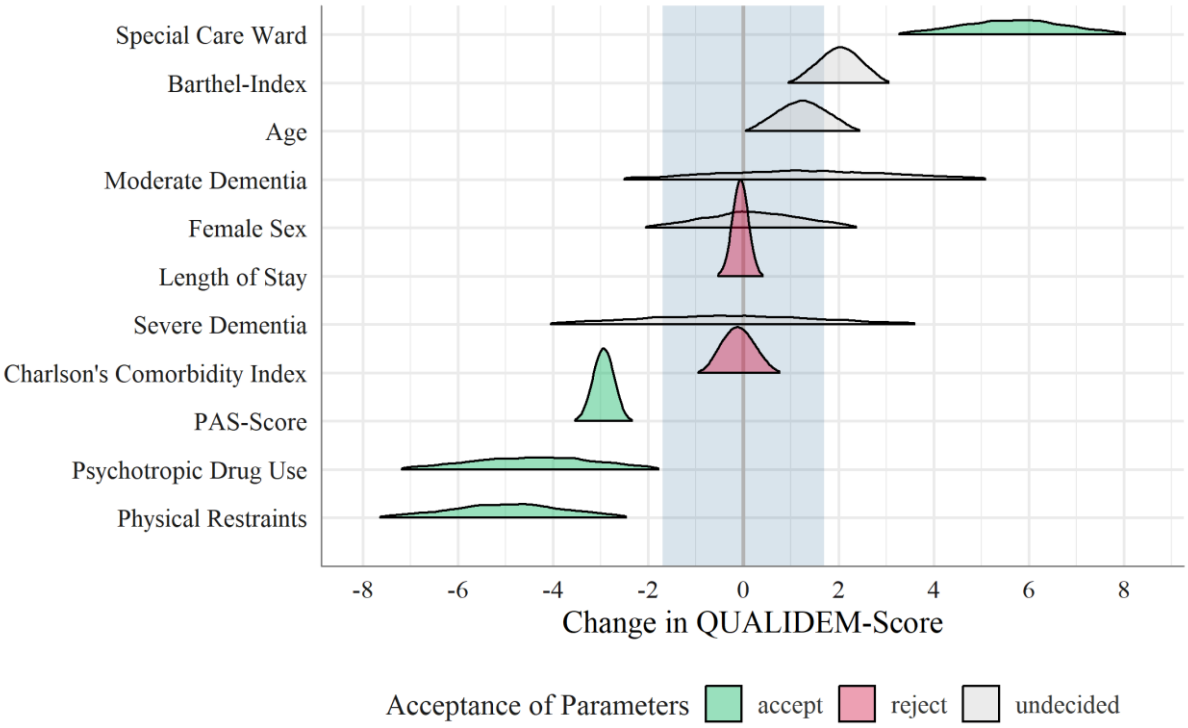
Bayesian methods do not perform classical "null hypothesis significance tests". Rather, to account for the uncertainty in estimation and the magnitude of parameter values, Kruschke[3] suggests checking whether parameter values lie inside a certain range that is considered as "practically no effect".

Figure S2 shows the distribution of the 95% Highest Density Intervals of posterior samples and indicates whether these lie completely outside, completely inside or partially inside the region of practical equivalence (ROPE). If the HDI is completely outside the ROPE, the "null hypothesis" for this parameter is "rejected". If the ROPE completely covers the HDI, i.e. all most credible values of a parameter are inside the region of practical equivalence, the null hypothesis is accepted. Else, it is undecided whether to accept or reject the null hypothesis.

The ROPE limits are specified following Kruschke's suggestion, which is defined as *0 +/- SD(dependent variable) * 0.1* for linear models.

Physical restraint, psychotropic drug use, agitation (PAS-score) and the intervention (special care ward) are those predictors that can be considered as relevant for our outcome.

Figure S2: *95%-Range of the Distribution of Posterior Samples from Regression Model; Region of Practical Equivalence emphasized in light-blue.*



5. Summary of the Distribution of Prior and Posterior Samples

Bayesian regression models can incorporate prior information about the parameters. Informative or weakly informative priors have the advantage of regularization of parameters to rule out large effects. This leads to less biased results and reductions of outliers in situations where the sample size in general, or for certain groups in the data, is small. Where we found information about our parameters, we included it by specifying *informative priors*. For instance, studies suggest that physical restraints are, on average, associated with an about 5-point reduction in quality of life, so the *location* of the prior distribution of our parameter physical restraint is at -5.0 (see Table S2). *Weakly informative priors* simply have a location of zero.

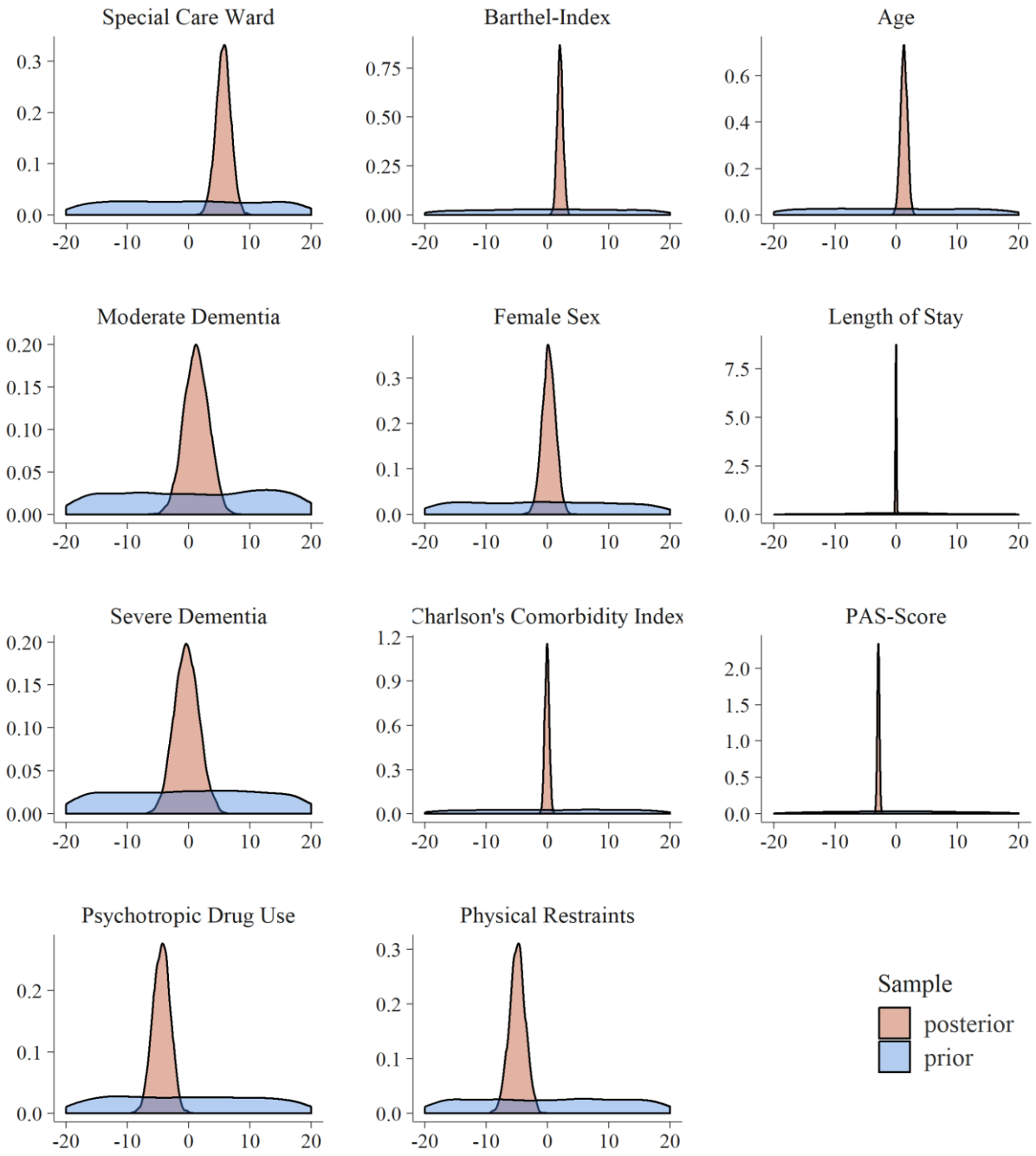
Table S2: *Prior Summary*

Predictor	Location	Scale
Length of Stay	0.00	6
Age	0.16	40
Moderate dementia	0.00	42
Severe dementia	-0.44	42
Sex, female	-3.22	42
Barthel-Score	0.00	29
Physical restraints	-5.00	42
Special Care Ward	0.00	42
PAS-Score	0.00	13
Charlson's Comorbidity Index	0.00	27
Psychotic Drug Use	0.29	42

To prevent a too strong shrinkage of uncertainty, the *scale* of the prior distribution should be large enough. If no assumptions about the scale of the effect are made, a recommended way (for normally distributed priors) is to use a scale of 2.5 and adjust it by dividing the scale by the standard deviation of the related predictor and multiply it with the standard deviation of the outcome (*2.5 * SD(y) / SD(x)*, see https://cran.r-project.org/web/packages/rstanarm/vignettes/priors.html).

When the evidence in the data is *strong*, the impact of the prior information is rather low. On the other hand, if the evidence of the data is *weak*, for instance due to small sample size, the impact of the prior information increases. Figure S3 shows the summary of the prior and posterior sample distributions and suggests that the evidence in the data was rather strong and results were not biased due to unrealistic prior information.

Figure S3: *Posterior versus Prior Summary*

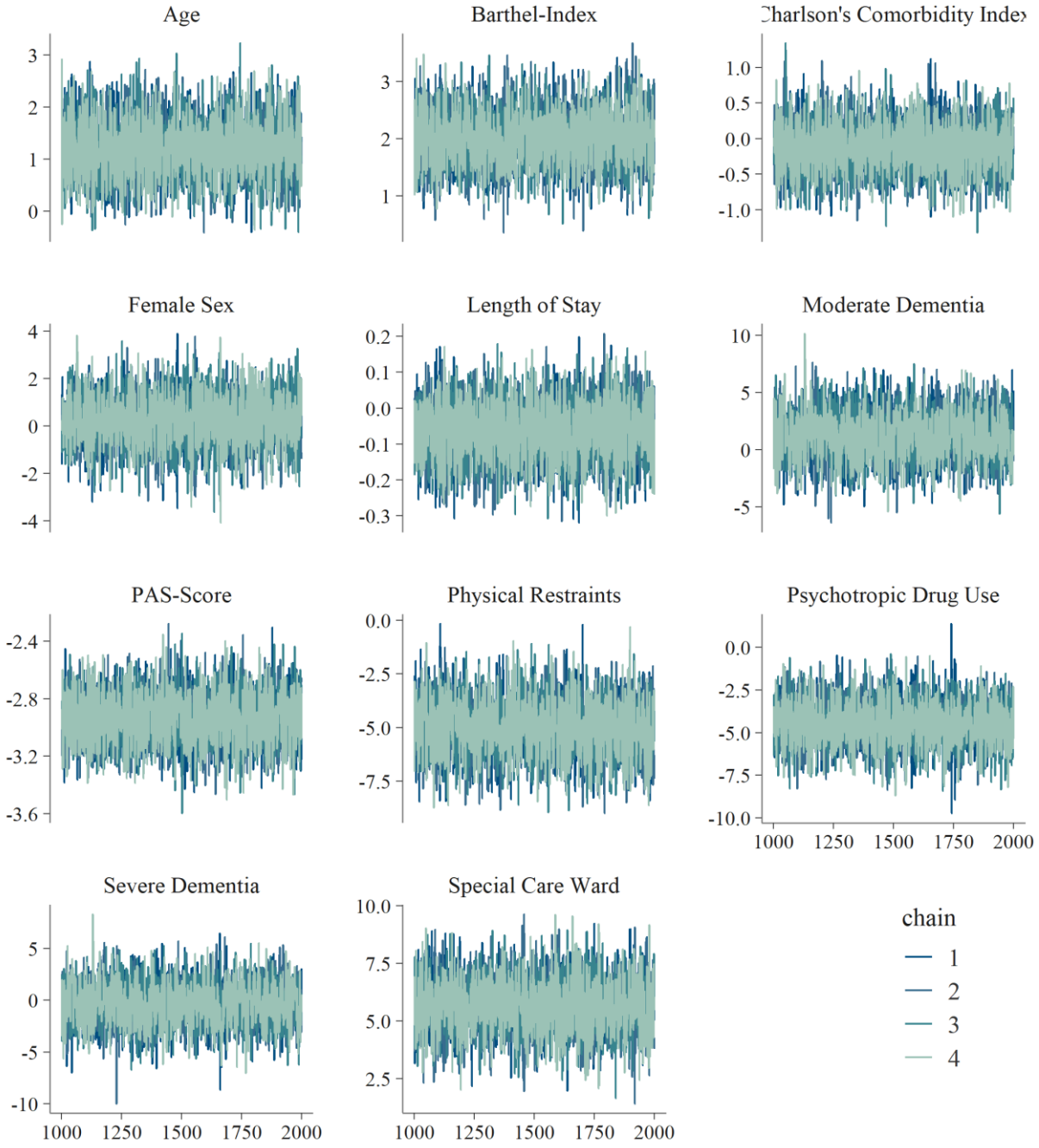


6. Trace plots

A trace plot provides a visual way to inspect sampling the behaviour and can be used, besides the Rhat-statistics, to assess convergence. The desired pattern is a "fat, hairy caterpillar", which shows no suspicious bends4, 5. If the chains converged to the posterior distribution, we can be confident in our model-based inferences.

The trace plot in Figure S4 indicates that all chains have converged well.

Figure S4: *Markov Chain trace plot, 4 chains with 1000 iterations each (excluding warmup-samples)*



7. References

1. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. Bayesian Data Analysis. Third edition. Boca Raton: CRC Press; 2014: p.24

2. Gelman A, Rubin DB. Inference from Iterative Simulation Using Multiple Sequences. Statistical Science. 1992;7(4):457-472. doi:10.1214/ss/1177011136

3. Kruschke JK. Rejecting or Accepting Parameter Values in Bayesian Estimation. Advances in Methods and Practices in Psychological Science. 2018; doi: 10.1177/2515245918771304

4. Sorensen T, Hohenstein S, Vasishth S. Bayesian linear mixed models using Stan: A tutorial for psychologists, linguists, and cognitive scientists. The Quantitative Methods for Psychology. 2016;12:175–200.

5. Lunn D, Jackson C, Best N, Thomas A, Spiegelhalter D. The BUGS book: a practical introduction to Bayesian analysis. Boca Raton, FL: CRC Press, Taylor & Francis Group; 2013.

Supplemental Material 2

To see if the complexity of patients' needs affect our findings, we calculated regression models with interaction terms between need factors moderated by hospitals. The figures S5 to S8 show the results of the interaction terms. Table S3 shows the coefficients in details. We found no "significant" associations for the interaction terms, concluding that the association between complexity of patients' needs and quality of life is not differing noticeably between the intervention and control group.

Plots were creates using the *ggeffects* package in R (Lüdecke D. ggeffects: Tidy Data Frames of Marginal Effects from Regression Models. Journal of Open Source Software. 2018;3: 772. doi:10.21105/joss.00772).

Figure S5: *Interaction between Barthel-Index and Intervention/Control-Group*

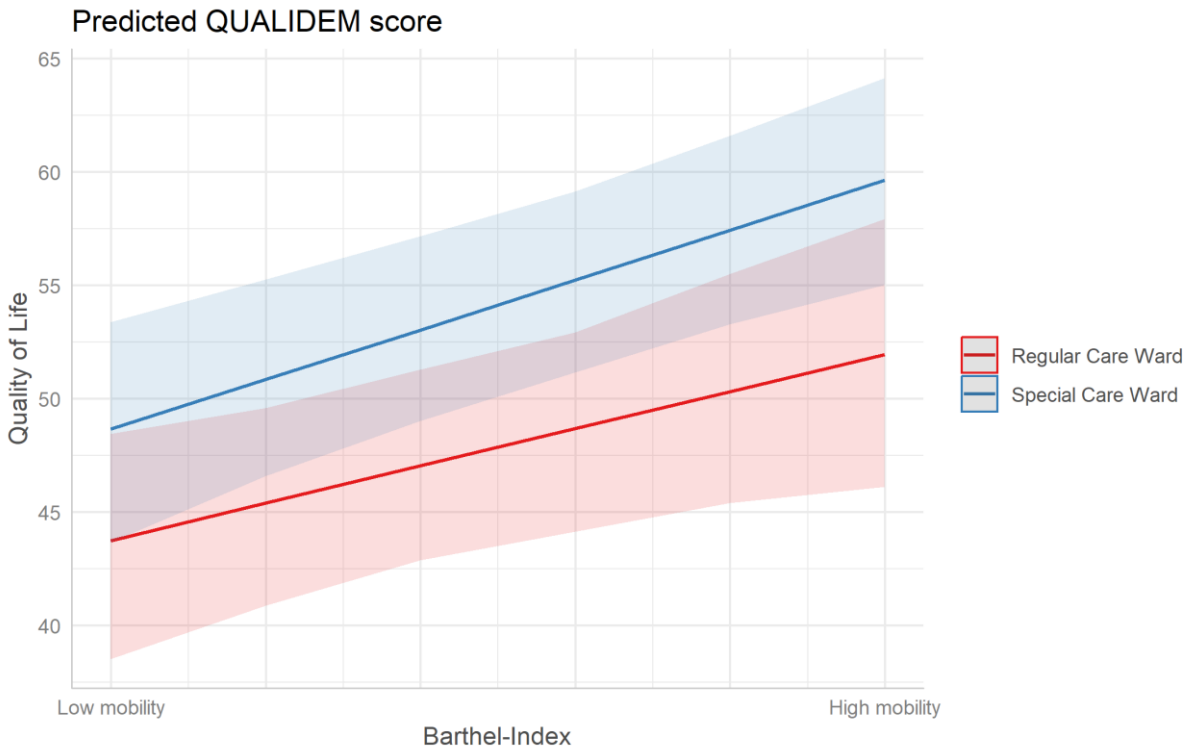


Figure S6: *Interaction between Physical Restraints and Intervention/Control-Group*

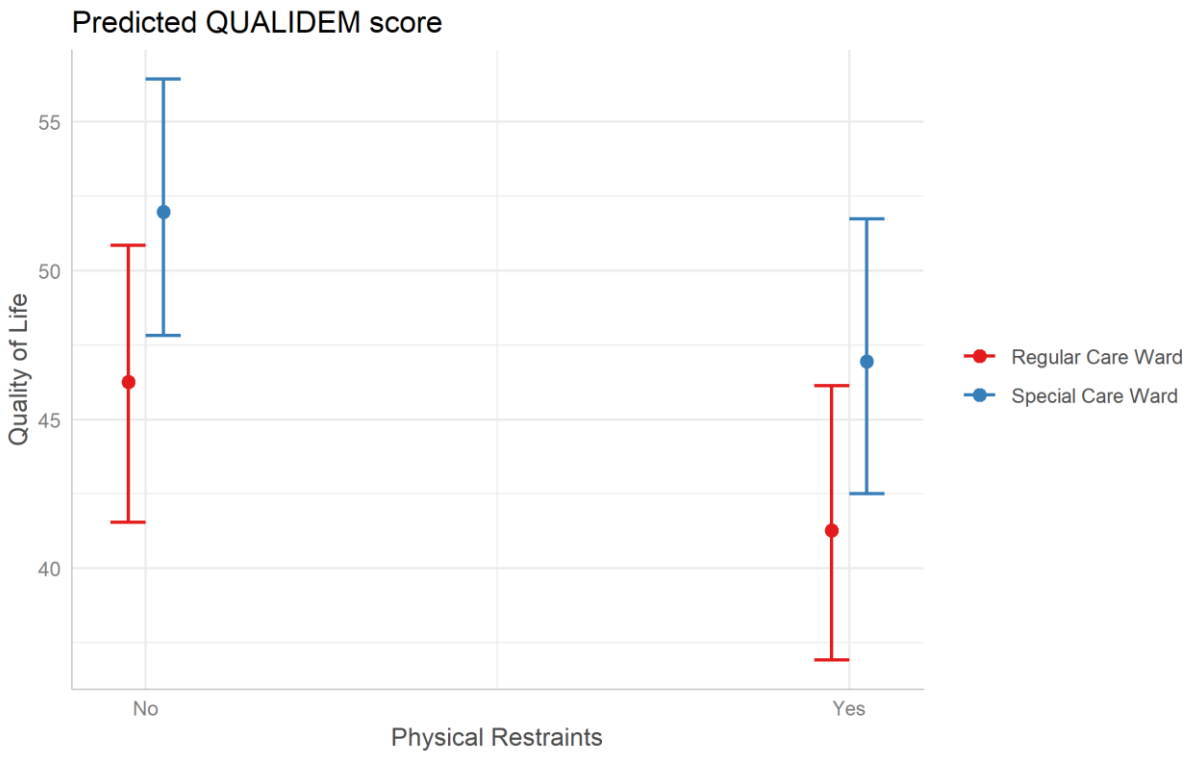


Figure S7: *Interaction between PAS and Intervention/Control-Group*

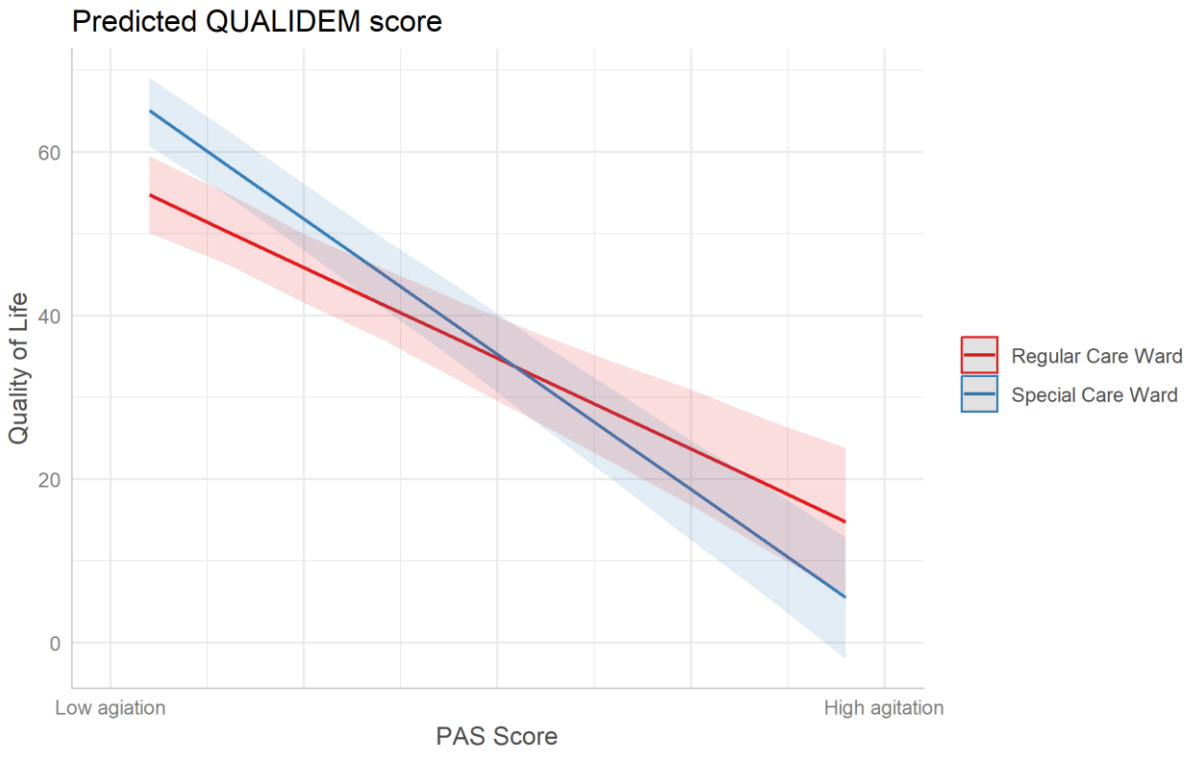


Figure S8: *Interaction between Chemical Restraints and Intervention/Control-Group*

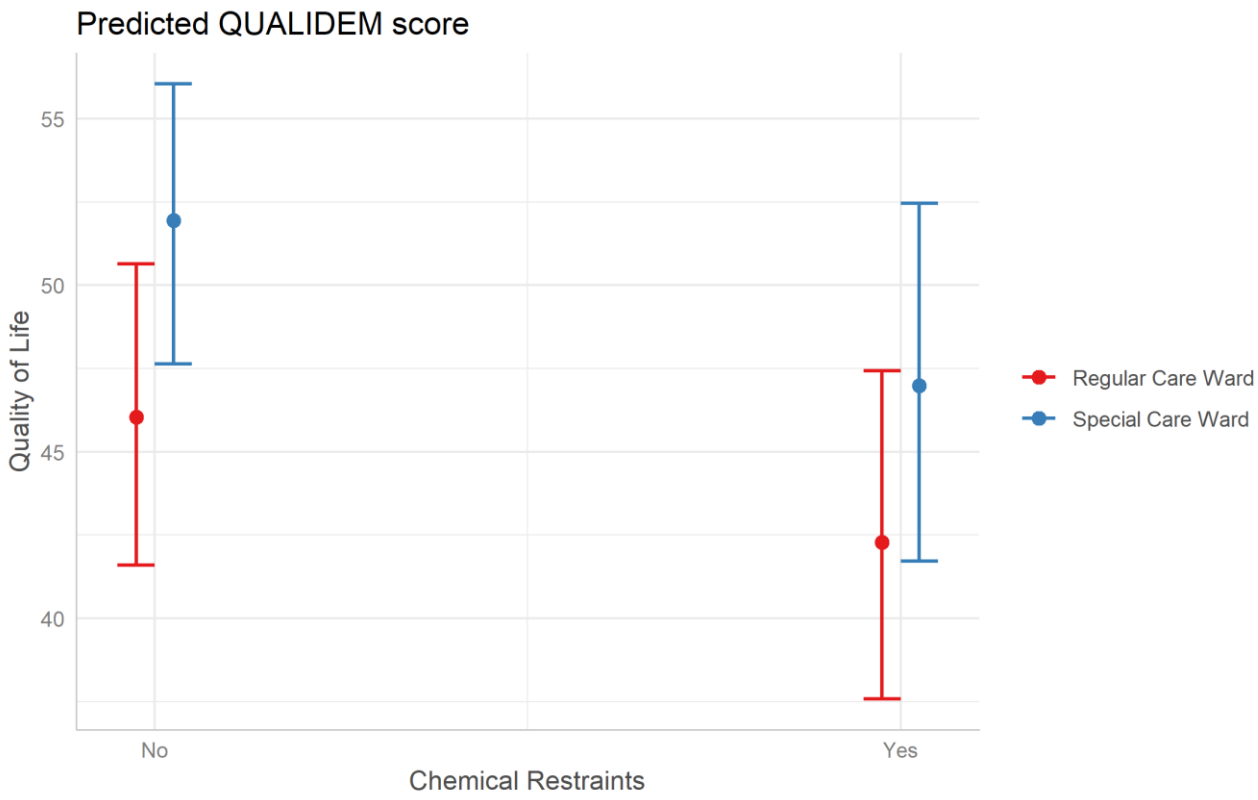


Table S3: *Coefficients for Models with Interaction Terms*

Terms	*Model: Interaction between Barthel and Hospital*		*Model: Interaction between Physical Restraints and Hospital*		*Model: Interaction between PAS and Hospital*		*Model: Interaction between Chemical Restraints and Hospital*	
	Estimates	*HDI (95%)*	*Estimates*	*HDI (95%)*	*Estimates*	*HDI (95%)*	*Estimates*	*HDI (95%)*
(Intercept)	36.18	26.19 – 45.67	36.31	27.00 – 46.10	36.57	27.11 – 45.71	36.16	26.45 – 45.69
Length of Stay	-0.07	-0.23 – 0.11	-0.06	-0.23 – 0.09	-0.07	-0.23 – 0.10	-0.06	-0.22 – 0.10
Age	0.12	0.01 – 0.23	0.12	0.02 – 0.23	0.12	0.01 – 0.22	0.12	0.02 – 0.23
Moderate Dementia	1.21	-2.79 – 5.13	1.19	-2.72 – 5.24	1.19	-2.83 – 4.72	1.25	-2.95 – 4.89
Severe Dementia	-0.34	-4.30 – 3.70	-0.44	-4.45 – 3.65	-0.74	-4.54 – 3.10	-0.30	-4.45 – 3.59
Female	0.16	-2.06 – 2.29	0.19	-1.98 – 2.29	0.52	-1.53 – 2.70	0.23	-1.76 – 2.53
Charlson's Comorbidity Index	-0.12	-0.81 – 0.59	-0.11	-0.77 – 0.64	-0.11	-0.77 – 0.59	-0.11	-0.82 – 0.59
Barthel Score	1.66	0.27 – 2.99	2.02	1.15 – 2.88	2.07	1.17 – 2.97	2.05	1.13 – 2.98
Physical Restraints (yes)	5.74	3.48 – 7.90	5.76	2.67 – 8.72	5.95	3.64 – 8.18	5.94	3.44 – 8.62
Special Care Ward (Intervention)	-5.02	-7.65 – -2.54	-4.92	-8.57 – -1.11	-4.94	-7.60 – -2.41	-4.96	-7.83 – -2.58
PAS-Score	-2.96	-3.30 – -2.62	-2.94	-3.30 – -2.61	-2.22	-2.77 – -1.66	-2.93	-3.27 – -2.59
Psychotropic Drug Use (yes, as-needed)	-4.43	-7.10 – -1.83	-4.40	-7.08 – -1.87	-4.30	-6.81 – -1.56	-3.76	-7.83 – -0.03
Barthel * Intervention	0.52	-1.00 – 2.06						
Phys. Restr. * Intervention			-0.09	-4.46 – 4.66				
PAS * Intervention					-1.09	-1.76 – -0.42		
Chem. Restr. * Intervention							-1.15	-6.63 – 4.27

Supplemental Material 3 – R Source Code

```
library(tidyverse)
library(ggridges)
library(sjmisc)
library(sjlabelled)
library(sjstats)
library(sjPlot)
library(brms)

# Data available at https://doi.org/10.5281/zenodo.1479676

# load data ----

load("Dataset.RData")

# divide age by 10

d$age10 <- d$age / 10


# Labels for final model ----

labs <-
  c(
    stay_c = "Length of Stay",
    age = "Age",
    age10 = "Age",
    mmse2 = "Moderate Dementia",
    mmse3 = "Severe Dementia",
    sex2 = "Female Sex",
    barthel_code = "Barthel-Index",
    groupintervention = "Special Care Ward",
    physres1 = "Physical Restraints",
    pas_c = "PAS-Score",
    cci_c = "Charlson's Comorbidity Index",
    chemicalres1 = "Psychotropic Drug Use", # oder as-needed
    b_stay_c = "Length of Stay",
    b_age = "Age",
    b_age10 = "Age",
    b_mmse2 = "Moderate Dementia",
    b_mmse3 = "Severe Dementia",
    b_sex2 = "Female Sex",
    b_barthel_code = "Barthel-Index",
    b_groupintervention = "Special Care Ward",
    b_physres1 = "Physical Restraints",
    b_pas_c = "PAS-Score",
    b_cci_c = "Charlson's Comorbidity Index",
    b_chemicalres1 = "Psychotropic Drug Use" # oder as-needed
  )

# prior-definition in brms ----

# scale is 2.5 * sd(y) / sd(x)

bprior <-
  prior(normal(0, 6), class = "b", coef = "stay_c") +
  prior(normal(.1554, 40), class = "b", coef = "age10") +
  prior(normal(0, 42), class = "b", coef = "mmse2") +
  prior(normal(-.444, 42), class = "b", coef = "mmse3") +
  prior(normal(-3.219, 42), class = "b", coef = "sex2") +
  prior(normal(0, 29), class = "b", coef = "barthel_code") +
  prior(normal(-5, 42), class = "b", coef = "physres1") +
  prior(normal(0, 42), class = "b", coef = "groupintervention") +
  prior(normal(0, 13), class = "b", coef = "pas_c") +
  prior(normal(.2925, 42), class = "b", coef = "chemicalres1") +
  prior(normal(0, 26.77), class = "b", coef = "cci_c")

# see:
# Beerens HC, Sutcliffe C, Renom-Guiteras A, et al. Quality of Life and
# Quality of Care for People With Dementia Receiving Long Term Institutional
# Care or Professional Home Care: The European RightTimePlaceCare Study.
# Journal of the American Medical Directors Association. 2014;15(1):54-61.
# doi:10.1016/j.jamda.2013.09.010
```

```
#
# QoL-Scale ranges from 13-52 (40 points). Effects from those study are
# multiplied by 2.5 (rescaling 40-points to 100-point-scale of QUALIDEM)

# see:
# Barca, M. L., Engedal, K., Laks, J., & Selbæk, G. (2011). Quality of Life
# among Elderly Patients with Dementia in Institutions. Dementia and Geriatric
# Cognitive Disorders, 31(6), 435-442. https://doi.org/10.1159/000328969

# QoL-scale ranges from 11-55 (45 points). Effects from those study are
# multiplied by 2.2 (rescaling 45-points to 100-point-scale of QUALIDEM)


# model formula ----

mf <-
  formula(
    QoL ~ stay_c + age10 + mmse + sex + cci_c +
      barthel_code + physres + group + pas_c +
      chemicalres + (1 | maindiag)
  )


# brms-model ----

set.seed(1207)

m2a <- brm(
  formula = mf,
  data = d,
  prior = bprior,
  sample_prior = TRUE
)


# Figure 3 ----

theme_set(theme_sjplot2(base_size = 14, base_family = "serif"))

p <- plot_model(
  m2a,
  title = "",
  axis.labels = labs,
  sort.est = T,
  colors = c("grey30"),
  axis.title = "Change in QUALIDEM-Score",
  wrap.title = 100,
  wrap.labels = 20,
  width = .2,
  grid.breaks = 2,
  size.inner = .1
) +
  ylab("Change in QUALIDEM Total Score") +
  theme_sjplot2(base_size = 14, base_family = "serif")

p_pdf <- p + theme_sjplot2(base_size = 28, base_family = "serif")

ggsave(
  filename = "Fig3.tiff", plot = p, width = 170, height = 120, units = "mm",
  dpi = 300, compression = "lzw"
)
ggsave(
  filename = "Fig3.pdf", scale = 2, plot = p_pdf, width = 170, height = 120,
  units = "mm", dpi = 300
)


# Appendix S1: Test for practical equivalence ----

rope(m2a, rope = c(-6, 6))
rope(m2a, rope = c(-7.5, 7.5))

equi_test(m2a)
```

```
# Appendix S1, Table Regression Coefficients ----

tab_df(tidy_stan(m2a, prob = c(.5, .89), digits = 1))


# Appendix S1, Prior Adjustement ----

ps <- prior_summary(m2a)
ps


# Appendix S1, Figure distribution Posterior Samples ----

tmp <- m2a %>%
  as_tibble() %>%
  select(2:12) %>%
  gather(key = "predictor", value = "estimate") %>%
  to_factor(predictor)

tmp$predictor <- lvls_reorder(tmp$predictor, c(9, 4, 8, 3, 7, 11, 10, 6, 1, 2, 5))

p2 <- ggplot(tmp, aes(x = estimate, y = predictor)) +
  geom_vline(xintercept = 0, colour = "grey70", size = .8) +
  geom_density_ridges2(rel_min_height = 0.001, scale = 2, alpha = .5) +
  scale_x_continuous(breaks = seq(-8, 8, 2)) +
  scale_y_discrete(labels = labs) +
  labs(x = "Change in QUALIDEM-Score", y = NULL) +
  theme_sjplot2(base_size = 14, base_family = "serif")

ggsave(filename = "FigS1.tiff", plot = p2, width = 190, height = 120, units = "mm", dpi = 300)


# Appendix S1, test for practical equivalence ----

## Short version

equi_test(m2a, out = "plot")

## More beautiful tweaked version

tmp.hdi <- hdi(m2a, prob = .95) %>%
  slice(c(-1, -13))

tmp2 <- m2a %>%
  as_tibble() %>%
  select(2:12) %>%
  map2_df(tmp.hdi$hdi.low, function(x, y) {
    x[x < y] <- NA
    x
  }) %>%
  map2_df(tmp.hdi$hdi.high, function(x, y) {
    x[x > y] <- NA
    x
  }) %>%
  gather(key = "predictor", value = "estimate") %>%
  to_factor(predictor)

tmp2$predictor <- lvls_reorder(tmp2$predictor, c(9, 4, 8, 3, 7, 11, 10, 6, 1, 2, 5))

tmp2$grp <- dplyr::case_when(
  tmp2$predictor %in% c("b_stay_c", "b_cci_c") ~ "reject",
  tmp2$predictor %in% c("b_age10", "b_mmse2", "b_mmse3", "b_sex2", "b_barthel_code") ~
"undecided",
  TRUE ~ "accept"
)

p3 <- ggplot(tmp2, aes(x = estimate, y = predictor, fill = grp)) +
  # rope based on "equi_test(model)".
  annotate("rect", xmin = -1.7, xmax = 1.7, ymin = 0, ymax = Inf, fill = sjplot_pal(palette =
"us")[1], alpha = 0.15) +
  geom_vline(xintercept = 0, colour = "grey70", size = .8) +
  geom_density_ridges2(rel_min_height = 0.01, scale = 2, alpha = .4) +
  scale_x_continuous(breaks = seq(-8, 8, 2)) +
  scale_y_discrete(labels = labs) +
  scale_fill_manual(values = sjplot_pal()[c(3, 1, 7)]) +
```

```
  labs(x = "Change in QUALIDEM-Score", y = NULL, fill = "Acceptance of Parameters") +
  theme_sjplot2(base_size = 14, base_family = "serif") +
  theme(
    legend.title = element_text(size = 13),
    legend.position = "bottom",
    axis.line.x = element_line(colour = "grey50"),
    axis.line.y = element_line(colour = "grey50")
  )

ggsave(filename = "FigS2.tiff", plot = p3, width = 190, height = 120, units = "mm", dpi = 300)


# Appendix S1, Posterior-Prior-Check ----

## Short version

plot_model(m2a, type = "diag", axis.lim = c(-20, 20))

## More beautiful tweaked version

pr_samp <- prior_samples(m2a) %>%
  select(starts_with("b_")) %>%
  gather(key = "Term", value = "Estimate") %>%
  mutate(Sample = "prior")

ps_samp <- posterior_samples(m2a) %>%
  select(starts_with("b_"), -b_Intercept) %>%
  gather(key = "Term", value = "Estimate") %>%
  mutate(Sample = "posterior")

m_pp_data <- bind_rows(pr_samp, ps_samp) %>% to_factor(Term)
m_pp_data$Term <- lvls_reorder(m_pp_data$Term, rev(c(8, 3, 10, 4, 7, 5, 11, 6, 1, 2, 9)))

p4 <- ggplot(m_pp_data, aes(x = Estimate, fill = Sample)) +
  geom_density(alpha = .4) +
  scale_x_continuous(limits = c(-20, 20)) +
  facet_wrap(
    ~ Term,
    scales = "free",
    labeller = labeller(Term = labs),
    nrow = 4
  ) +
  labs(x = NULL, y = NULL) +
  bayesplot::theme_default(base_size = 13) +
  theme(
    axis.line.x      = element_line(colour = "grey50"),
    axis.line.y      = element_line(colour = "grey50"),
    axis.text        = element_text(colour = "grey10"),
    axis.title       = element_text(colour = "black"),
    # strip.background = element_rect(colour = "grey50", fill = "grey90"),
    # strip.text       = element_text(colour = "grey20"),
    legend.title     = element_text(colour = "grey10"),
    legend.text      = element_text(colour = "grey20"),
    legend.position  = c(.5, .15),
    legend.justification = c(-2, 1)
  ) +
  scale_fill_manual(values = sjplot_pal("breakfast club")[c(1, 3)])

ggsave(filename = "FigS3.tiff", plot = p4, width = 190, height = 214, units = "mm", dpi = 300)


# Appendix S1, Traceplot ----

p <- rstan::traceplot(m2a$fit, pars = colnames(as.data.frame(m2a))[1:12], inc_warmup = F)
p$data$parameter <- as.character(p$data$parameter)
tmp <- p$data %>%
  filter(parameter != "b_Intercept")

for (i in 1:length(labs)) {
  if (names(labs)[i] %in% tmp$parameter) {
    r <- which(tmp$parameter == names(labs)[i])
    tmp$parameter[r] <- labs[i]
  }
}
```

```
p5 <- ggplot(tmp, aes(x = iteration, y = value, colour = chain)) +
  geom_line() +
  facet_wrap(~parameter, scales = "free_y", ncol = 3) +
  scale_color_manual(values = sjplot_pal("us", n = 4)) +
  labs(x = NULL, y = NULL) +
  bayesplot::theme_default(base_size = 13) +
  theme(
    axis.line.x      = element_line(colour = "grey50"),
    axis.line.y      = element_line(colour = "grey50"),
    axis.text        = element_text(colour = "grey10"),
    axis.title       = element_text(colour = "black"),
    # strip.background = element_rect(colour = "grey50", fill = "grey90"),
    # strip.text       = element_text(colour = "grey20"),
    legend.title     = element_text(colour = "grey10"),
    legend.text      = element_text(colour = "grey20"),
    legend.position  = c(.5, .15),
    legend.justification = c(-4.2, 0.7)
  )

ggsave(filename = "FigS4.tiff", compress = "lzw", plot = p5, width = 190, height = 214, units =
"mm", dpi = 300)
```

STROBE 2007 (v4) Statement—Checklist of items that should be included in reports of *case-control studies*

Section/Topic	Item #	Recommendation	Reported on page #
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	4-5
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	6-7
Participants	6	*(*a) Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls	6-7
		(*b*) For matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	8-10
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	8-10
Bias	9	Describe any efforts to address potential sources of bias	10
Study size	10	Explain how the study size was arrived at	7
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	10-11
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	10-11
		(*b*) Describe any methods used to examine subgroups and interactions	10-11
		(*c*) Explain how missing data were addressed	10
		(*d*) If applicable, explain how matching of cases and controls was addressed	N/A
		(*e*) Describe any sensitivity analyses	N/A
Results			

Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	12
		(b) Give reasons for non-participation at each stage	N/A
		(c) Consider use of a flow diagram	N/A
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	12-13
		(b) Indicate number of participants with missing data for each variable of interest	12-13
Outcome data	15*	Report numbers in each exposure category, or summary measures of exposure	12-13
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	13-14
		(*b*) Report category boundaries when continuous variables were categorized	N/A
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	N/A
Discussion			
Key results	18	Summarise key results with reference to study objectives	15-16
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	17-18
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	15-18
Generalisability	21	Discuss the generalisability (external validity) of the study results	18-19
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	20

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Quality of Life of Patients with Dementia in Acute Hospitals: A non-randomised case-control-study comparing a regular ward to a special care ward with dementia care concept.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-030743.R1
Article Type:	Research
Date Submitted by the Author:	26-Jun-2019
Complete List of Authors:	Lüdecke, Daniel; University Medical Center Hamburg-Eppendorf, Department of Medical Sociology Poppele, Georg; Evangelisches Krankenhaus Alsterdorf Klein, Jens; University Medical Center Hamburg-Eppendorf, Department of Medical Sociology Kofahl, Christopher; University Medical Center Hamburg-Eppendorf, Department of Medical Sociology
Primary Subject Heading:	Patient-centred medicine
Secondary Subject Heading:	Health services research, Nursing
Keywords:	Dementia < NEUROLOGY, Quality of Life, Acute Hospital, Quality of Care, INTERNAL MEDICINE

Title

Quality of Life of Patients with Dementia in Acute Hospitals: A non-randomised case-control-study comparing a regular ward to a special care ward with dementia care concept.

Authors

Daniel Lüdecke, Georg Poppele, Jens Klein, Christopher Kofahl

Corresponding author

Daniel Lüdecke
Department of Medical Sociology
University Medical Center Hamburg-Eppendorf
Martinistr. 52
20246 Hamburg
Germany
Email: d.luedecke@uke.de
ORCID iD: 0000-0002-8895-3206
Tel.: ++49 40 7410 52878

Co-authors

Georg Poppele, Evangelisches Krankenhaus Alsterdorf (Protestant Hospital Alsterdorf), Hamburg, Germany

Jens Klein, University Medical Center Hamburg-Eppendorf, Department of Medical Sociology, Hamburg, Germany

Christopher Kofahl, University Medical Center Hamburg-Eppendorf, Department of Medical Sociology, Hamburg, Germany

Keywords

Dementia, Quality of Life, Acute Hospital, Quality of Care, Internal Medicine

Word Count: 5.095 (excluding title page, abstract, strength and limitations, references, tables)

Abstract

Objectives: To identify factors that predict the quality of life (QoL) of patients with dementia in acute hospitals and to analyse if a special care concept can increase the patients' QoL.

Design: A non-randomised case-control-study, including two internal medicine wards from hospitals in Hamburg, Germany.

Setting and Participants: In all, 526 patients with dementia from two hospitals were included in the study (intervention: n=333; control: n=193). Inclusion criteria was an at least mild cognitive impairment or dementia. The intervention group was a hospital with a special care ward for internal medicine focussing on patients with dementia. The control group was from a hospital with a regular care ward without special dementia care concept.

Outcome Measures: Our main outcome was the QoL (range 0-100) from patients with dementia in two different hospitals. A Bayesian multilevel analysis was conducted to identify predictors like age, dementia, agitation, physical and chemical restraints or functional limitations that affect QoL.

Results: QoL differs significantly between the control (40.7) and intervention group (51.2), $p<0.001$. Regression analysis suggests that physical restraint (estimated effect: -5.0), psychotropic drugs use (-4.4) and agitation (-2.9) are negatively associated with QoL. After controlling for confounders, the positive effect of the special care concept remained (5.7).

Conclusions: A special care ward will improve the quality of care and has a positive impact on the QoL of patients with dementia. Health policies should consider the benefits of special care concepts and develop incentives for hospitals to improve the QoL and quality of care for these patients.

Article Summary

Strengths and Limitations

- This is one of the first studies to investigate quality of life of patients with dementia in acute internal medicine wards in Germany.
- The statistical method applied in this study explicitly incorporates and accounts for information and knowledge from previous research.
- There are no studies which have evaluated the reliability and validity for the use of the assessment instrument for our main outcome (quality of life) in hospitals settings.
- The structural differences between the hospitals from the control and intervention may limit the generalizability of the results regarding the benefit of special care concepts for regular internal medicine wards.

Introduction

Acute hospitals face the challenge of an increase in old-age patients, which particularly affects internal medicine wards [1]. The average age of patients in internal medicine wards is above 70 years and may even come close to 80 years [2,3], leading to an increasing prevalence of cognitive impairments [4]. Although data on the occurrence of dementia or cognitive impairments of patients in hospitals is inconsistent, the larger proportion of studies report prevalence rates of about 40% [5–7].

Many hospitals are insufficiently prepared for patients with cognitive impairments, especially in acute care units predominantly focussing on somatic diseases [8]. Patients with cognitive impairments or dementia do not fit into the typical routines and standardised workflows of hospitals as these patients need more resources for care and treatment [9,10]. These patients often become disoriented, anxious, agitated and challenge hospital staff with erratic behaviour when placed in regular care wards. This results in an increased likelihood of falls, complications during the hospital stay and post-operative complications [11,12]. Hence, patients with dementia are a vulnerable group with a higher risk for long-lasting functional impairments [13,14].

To control the challenging behaviour of patients with dementia, a common practice – at least in German acute hospitals, and especially for older patients – is to use physical restraints like side rails to keep patients in bed, or chemical restraints such as the on-demand-use of psychotropic medication or hypnotics and sedatives [15,16]. This practise is not limited to German hospitals; however, there seems to be large variations between countries [17]. Physical and chemical restraints as well as anxiety and challenging behaviour are associated with poorer outcomes in quality of life [QoL] and quality of care of patients with dementia [18,19]. Therefore, hospitals in general, and specifically internal medicine wards with their increasing proportion of patients with

dementia, need to address these issues in order to improve the quality of care for these patients.

At least in Germany, there were lately no care concepts that fully address the needs of patients with dementia in internal medicine [20]. The special care ward "DAVID" in the Protestant Hospital Alsterdorf in Hamburg was one of the first internal medicine wards in Germany that implemented a comprehensive care concept for patients with dementia, aiming to improve the patients' QoL during their hospital stay. QoL is an important indicator of quality of care and a major dimension when assessing patient reported outcomes, particularly in older people as global outcome measure for interventions [21,22]. The assumption of this care concept is that a special care ward for patients with dementia leads to better outcomes in QoL compared to regular internal medicine wards. A study ("DAVID 2") was conducted to investigate the impact of such a care concept. This paper shows the results of this study and addresses two research questions. First, which factors predict the QoL of patients with dementia in acute hospitals? Second, beyond these factors, can a special care concept for patients with dementia in acute hospitals increase the patients' QoL?

Methods

Study Design and Setting

The aim of this study was to compare the quality of care for patients with dementia within a specialised dementia care concept as opposed to regular care in acute hospitals. The present study was designed as a non-randomised case-control-study, including two internal medicine wards in two hospitals located in Hamburg, Germany. The intervention group was a hospital that implemented a special care ward for internal medicine focussing on patients with dementia. The control group was from a hospital

with a regular care ward for internal medicine, which had no special dementia care concept.

Intervention Group

The special care ward "DAVID" is an internal medicine ward in the Protestant Hospital Alsterdorf, a not-for-profit organization, and has 14 beds. In the year of data collection (2016), 349 patients were treated. The ward employed nine care workers as nursing staff.

Key components of the special care concept are a) a specific architectonical design, including a homelike lounge, a specific colouring of doors and walls, and a light concept with minimum 500 lux at eye level; b) doctors, nurses and service staff are trained in coping with challenging behaviour and other dementia related issues, like basal stimulation or validation therapy, but also included case conferences to discuss issues with current patients [23]; duration of training courses and case conferences was about one hour and were provided on a monthly basis by external instructors; additionally, twice per year, an internal training course was offered for employees, lasting for half a day; c) mobile devices for diagnostics, to perform as many treatments as possible in the different rooms of the special care ward; d) involvement of relatives into assessment, care and discharge planning; and e) regular therapeutic offers like occupational or speech therapy, and social offers like music, playing or spending more time than usual to care for the patients.

To fulfil these high standards of quality of care, the ward "DAVID" employs more care staff in relation to the number of patients as compared to other regular internal medicine wards in Germany. With respect to the total number of full-time equivalents [FTE] nurses, the staff-patient-ratio is one FTE nurse per 39 patients.

The Protestant Hospital Alsterdorf has a second ward for internal medicine, however, patients with dementia were usually immediately transferred to the special care ward after admission to hospital. Thus, as almost no patients with dementia were treated in the second internal medicine ward, the control group was taken from another hospital.

Control Group

The regular care ward is part of a larger private-company hospital with emergency hospitalisation. It has 80 beds and in the year of data collection, about 3.500 patients were treated in this internal medicine ward. Twenty-six employees worked as care staff in this ward. Trainees sometimes supported the care team. The staff-patient-ratio in the regular care ward is approximately one FTE nurse per 130 patients. However, since the internal medicine ward in this hospital also treats patients from the emergency ambulance, the staff-patient-ratio related to the number of patients who actually stayed longer in hospital (three days and more) is lower. Unfortunately, the hospital management was not willing to provide more detailed information beside the publicly available quality reports, so we cannot quantify the staff-patient-ratio exactly.
The regular care ward had no specific care concept for dementia patients. The care staff was not particularly trained in dementia topics.

Data collection and participants

An assessment questionnaire was developed to obtain data from patients with dementia. Study nurses were trained in using this assessment questionnaire and then conducted the data collection in both hospitals. Two study nurses were responsible for the special care ward and one for the regular care ward. A pre-test of two months was conducted to test and revise the questionnaire. As a result, some items were removed and instructions

for study nurses were defined more precisely. After the pre-test, data was collected over a period of about 12 months (from July 2015 to June 2016 in the special care ward and from August 2015 to September 2016 in the regular care ward). To detect small to medium effect sizes (Cohen's d ~ 0.1 to 0.2), a power analysis was performed prior to the data collection and yielded a sample size of at least 173 subjects per group. Patients were included when they showed at least mild cognitive impairments or memory problems. In the special care ward (intervention group) all patients were assessed because a diagnosed dementia was a requirement for admission to that hospital. Hence, the participation rate for the special care ward was about 94% and excluded only a few patients that were not responsive. For the regular care ward (control group), patients who already had a diagnosed dementia or cognitive impairments were included in the study. A short dementia screening was carried out by the study nurse to assess the severity of dementia of patients who had no clarified dementia diagnosis, and to identify further patients who qualify for the study [24]. The total sample size for the present analysis consists of N=526 patients (special care ward: n=333; regular care ward: n=193). For both the intervention and control group, patients were excluded from the study when they were completely confined to bed due to severe health-related dependency. As both care wards had no particular selection criteria for patients such as age, mobility, or the main diagnosis that lead to hospital admission, no further exclusion criteria for the study were defined.

Prior to the study, a study protocol was developed and submitted to the ethical committee of the medical association of Hamburg. The ethical committee approved the proposal and attested that the study conforms to ethical and legal requirements (approval code PV5102). Study participants were not able to give their informed consent due to their cognitive impairments. However, as data mostly derived from the

hospitals' regular documentation and was completely anonymous, the ethics committee waived the need of an informed consent.

Patient and Public involvement

Patients and the public were not involved in the development of the research question nor study design.

Measures

Outcome: QoL in patients with dementia was assessed using the QUALIDEM [25,26]. After observing patients for about one week (depending on the length of stay), the study nurses rated their QoL. QUALIDEM comprises 37 items reflecting nine different subdomains of QoL: "care relationship" (7 items, 0-21 points), "positive affect" (6 items, 0-18 points), "negative affect" (3 items, 0-9 points), "restless and tense behaviour" (3 items, 0-9 points), "positive self-image" (3 items, 0-9 points), "social relations" (6 items, 0-18 points), "social isolation" (3 items, 0-9 points), "feeling at home" (4 items, 0-12 points) and "have something to do" (2 items, 0-6 points). For patients with very severe dementia (Minimental State Examination Test [27] [MMSE] < 7), only six of the nine subscales apply, where the dimensions "positive self-image", "feeling at home" and "have something to do" were omitted. The recommendation is to report descriptive results of the QUALIDEM separately for each subscale. For regression analyses, a QoL index was calculated by summing up and normalizing the QUALIDEM subscales (six subscales for patients with very severe dementia, nine subscales for the remaining patients) to a range from 0 to 100 points. A higher score indicates better QoL. Due to normalization of the QUALIDEM total score for all severities of dementia, all patients' scores are consistent and comparable [28].

Independent Variables: Age, gender, main diagnosis for admission to hospital and length of stay were recorded. Details about the distribution of the main diagnoses among patients and by hospitals are shown in the Supplementary File 1. If a main diagnosis was mentioned no more than one time in both hospital wards, it was recoded into the category "other". The final variable "main diagnosis" comprised 20 different diagnoses. A modified version of the Charlson's Comorbidity Index [CCI], which included depression and hypertension as new items, was built based upon the assessment of comorbidities and chronical diseases [29,30]. If patients had no chronic illnesses, the CCI had a score of zero points. Else, higher scores indicated more serious comorbid disease. Shortly after admission to hospital, the study nurses measured functional limitations and cognitive status of patients. Functional limitations in daily living were assessed with the Barthel-Index [31]. This score ranges from 0 (completely dependent) to 100 points (no basic functional limitations) and was recoded according to the classification of the ICD-10 [32] (German adaption) into a score from 1 to 6 points. The Minimental State Examination Test [27] [MMSE] measures the cognitive impairments of patients, ranging from 0 (very strong cognitive impairments) to 30 (very mild or no cognitive impairments) points. This score was recoded into three categories, also based on ICD-10 classification: severe dementia (0-16), moderate dementia (17-23 points) and mild dementia (24-27 points). After about one week of hospital stay, the study nurses rated the patients' agitation and challenging behaviour and recorded psychotropic drug use (chemical restraint) and physical restraints. Agitation and challenging behaviour of patients was assessed using the Pittsburgh Agitation Scale [PAS] [33] ranging from 0 to 16 points (higher scores indicate stronger agitation).

Physical restraints were defined as the use of one the following measures: Side rails to keep a patient in bed, tying a patient to a bed, and use of "therapeutic" chairs that

prevent patients to stand up. The variable was dichotomised, indicating whether patients (in the course of the hospital stay) were mechanically restrained by at least one of these measures or not.

Psychotropic drug use was defined as on-demand-use ("as-needed") of medication for the nervous system by means of the Anatomical Therapeutical Chemical (ATC) classification [34] and comprises antipsychotics, anxiolytics, hypnotics/sedatives and antidepressants (N05A-C, N06A). Indicated use of psychotropic drugs was defined as medications that were prescribed for regular, not on-demand-use and not only given to patients in order to control their challenging behaviour. Such use of psychotropic drugs was excluded from the analysis. The on-demand-use variable was dichotomised and shows whether, during the complete hospital stay, chemical restraints were applied to patients or not.

While these variables already cover many different aspects that have an effect on the QoL, we decided to add a further predictor as proxy for the intervention to the model. Therefore, we included a binary variable with two categories ("control" as reference and "intervention") representing the two hospitals, to estimate the impact of the special care concept. This should reflect how much of the change in QoL is attributable to the special care concept.

Missing Data

In total, 11% of individual items across all scales were missing (at random), 6% of individual items when looking at the QUALIDEM only. The missing data pattern was analysed and missing data was imputed using the multivariate imputation by chained equations method [35], using 11 imputation steps corresponding to the proportion of missing data [36]. The method for imputing missing values depends on the variable's

nature. For continuous variables, predictive mean matching was applied, while logistic regressions were used for binary variables.

Statistical Methods

Descriptive results for the total sample and each hospital are reported. Statistically significant differences of $p<0.05$ between the two hospital wards were tested using t-tests, χ^2-tests or Mann-Whitney-U-tests, depending on the level of measurement and distribution of variables. Differences between the hospitals in the QUALIDEM subscales are presented as boxplots, showing the median value and upper and lower quartiles of the value distribution.

As multivariate analysis, a Bayesian linear mixed model was applied to analyse the associations between the independent variables and the outcome. Computations were based on Stan [37], a probabilistic programming language for specifying Bayesian models, using Markov Chain Monte Carlo sampling (in particular, Hamiltonian Monte Carlo) [38]. We assume that the patients' main diagnosis is associated with different degrees of physical impairments, which affect the QoL. Therefore, the variable 'main diagnosis' was used as level-2 unit (*random intercept*) in the multilevel model to control for the variation in the outcome. We used informative priors for the predictors age, female gender, severe dementia, psychotic drug use and physical restraints, based on information from former research [18,39,40]. Weakly informative priors were used for the remaining predictors. The prior and posterior distributions of the model are summarised in the supplemental material (see Supplementary File 2).

Continuous predictors were centred before entering the model. Age was divided by 10, so a 1-unit change in the predictor of age reflects a change of 10 years in patients. The median value of the posterior distribution is used as "Bayesian point estimate", which

minimises the difference of estimates from true values over posterior samples, but there are many other plausible values (the "posterior distribution") to describe the association between predictors and outcome. Hence, 50% and 89% highest density intervals [41] [HDI] are shown to indicate the range of most credible values and to reflect the (un-)certainty of the estimates. The intraclass correlation coefficient [42] was calculated to see how much of the proportion of the variance in the outcome can be explained by the grouping structure ('main diagnosis'). We developed post-hoc additional regression models with interaction terms for need predictors (Barthel-Index, physical and chemical restraints, PAS-Score) to check if the associations between the complexity of patients' needs and QoL differ between hospitals. We found no significant interaction terms and decided to present the most parsimonious model here and show further results in the appendix (see Supplementary File 3).

All analyses were conducted with the R statistical package [43], including the packages *mice* [35], *ggplot* [44], *brms* [45] and *sjPlot* [46]. The source code is available in the supplemental material (see Supplementary File 4). Data is available online [47].

Results

Sample Characteristics

Table 1 gives an overview of the sample characteristics. The proportion of female to male patients is similar in both groups. The mean age is 4 years higher in the control group. There are also significant group differences in the Barthel-Index indicating higher functional impairment in the control group, while the dementia severity was the same in both hospitals. Comorbid conditions are slightly higher in the control group. Patients stayed 9.4 days in hospital on average and nearly one day longer in the intervention group as compared to the control group. Large differences between the two hospitals

can be seen in the use of medical and physical restraints with significantly less use in the intervention group. Agitation- and QoL-scores also show strong group differences to the disadvantage of the patients in the control group.

In most cases, the distribution of main diagnoses of patients were comparable between the two hospital wards (see Supplemental File 1). Most frequent were pneumonia (13.5% in the intervention group and 11.9% in the control group), a worsening medical condition of patients (8.7% and 7.2%) or exsiccosis (4.8% and 6.7%). Noticeable differences between the two wards were found in urinary tract infections (UTI) (9.9% in the intervention group and 3.1% in the control group) or dyspnoea (1.2% and 7.8%).

Table 1: Sample Characteristics

Characteristic	**Control Group (Regular Care Ward, n=193)**	**Intervention Group (Special Care Ward, n=333)**	**Total (N=526)**	**p-value of difference**
Female, %	59.1	61.6	60.6	.637
Mean Age (SD)	83.1 (7.2)	79.0 (11.9)	80.5 (10.6)	**<.001**
Mean Barthel-Index (SD)	29.9 (27.9)	40.7 (30.4)	36.7 (29.9)	**<.001**
Mild Dementia, %	7.8	9.6	8.9	.580
Moderate Dementia, %	30.0	26.7	27.9	.473
Severe Dementia, %	62.2	63.7	63.2	.805
Mean Length of Stay, in Days (SD)	8.9 (7.5)	9.7 (5.5)	9.4 (6.3)	**.002**
Physical Restraints (yes), %	54.4	28.2	37.8	**<.001**
Psychotropic Drug Use (yes, as-needed), %	25.9	14.1	18.4	**.001**
Mean-Score Pittsburg Agitation Scale (SD)	3.9 (3.1)	3.0 (3.2)	3.3 (3.2)	**<.001**
Mean Charlson' Comorbitiy Index (SD)	3.2 (3.0)	2.5 (2.0)	2.8 (1.6)	**<.001**
Mean Qualidem Total Score (SD)	40.7 (14.5)	51.2 (17.2)	47.3 (17.0)	**<.001**

Barthel-Index: 0-100 (higher = better functioning); Dementia (MMSE score): mild: 24-27; moderate: 17-23; severe: <= 16; Pittsburg Agitation Scale: 0-16 (higher = stronger agitation and anxiety); Charlson' Comorbitiy Index: 0-9 (higher = more comorbidities); QUALIDEM: 0-100 (higher = better QoL)

Quality of life

Looking at the QoL for patients with severe to mild cognitive impairments (these are the ones, in which all subdomains of the QUALIDEM could be applied), there is a consistent pattern across all QUALIDEM-domains: Patients in the control group have a lower QoL compared to the intervention group. Except for the last subdomain ('having something to do'), all differences are statistically significant (Figure 1).

The same consistent pattern can be found for patients with very severe dementia symptoms (MMSE score < 7). Here, only the second of the six applied subdomains ('positive affect') does not differ significantly between intervention and control (Figure 2).

Predictors of quality of life

Figure 3 shows the results from the Bayesian mixed model. Three predictors are clearly negatively associated with QoL: physical restraint, psychotropic drug use and agitation (PAS-score). Physical restraint is associated with a 4.9-point decrease in QoL. With 50% probability, the QoL decreases by 4.1 to 5.8 points and with a chance of 89% by -7 to -2.8 points respectively. The application of psychotropic drugs as-needed shows similar results, with a posterior median of -4.4. The third clearly negative associated predictor is agitation, which shows a decrease in QoL of about 2.9 points for each additional point in the PAS-score.

Dementia and gender are not clearly associated with QoL. Neither are the length of hospital stay and the CCI.

The age of the patient correlates slightly positive with QoL, where an increase of 10 years means an increase of about 1.2 points in the QoL. The posterior median of the Barthel-Index is 2.0, so for a one-category change in functional impairments the QoL

changes by two points. This means that patients with severe functional impairments differ by about 10 points in QoL compared to patients with no functional impairments. Controlling for all other predictors, the intervention (special care ward) shows the strongest association with our outcome of interest, the patients' QoL. The posterior median is 5.7, and with an 89% probability, the credible values describing the effect of the intervention on QoL are within the range from 3.8 to 7.6.

The intraclass correlation coefficient of the model is rather low (0.01). This means, the 'main diagnosis' does not explain much of the variance in the patients' QoL and there is almost no regularization ("shrinkage") of estimated model parameters and no larger differences between hospitals according to the patients' needs, as indicated by their main diagnosis.

Discussion

The study reported in this paper sought to understand those factors that influence the QoL in patients with dementia and whether a special care concept for these patients performs better in this regard as opposed to regular care wards.

One of our main findings is that QoL differs significantly between the control and intervention group. We found substantial differences between the two hospitals in the patients' total QoL score in favour of the special care ward. Beyond the statistical significance, this finding also has a clinical impact. Studies suggest a change in 3 points for the Quality of Life – Alzheimer's Disease Scale [48], which has a range of 40 points, to be clinically relevant [49,50]. Transferred to the range of the QUALIDEM scale, a difference of about 7.5 points would be considered as an important improvement in QoL. Another indicator to evaluate the clinical relevance of a change in QoL is an increase of the score of half a standard deviation [51], which would be about 8.5 points

for our data. Taking these reference points as a basis, we found evidence for the clinical relevant improvement in QoL of patients in a special care ward.

A second key finding is the identification of those factors that are clearly associated with QoL. The use of physical and chemical restraints, both happening more frequently in the control group, are associated with lower outcomes in QoL. This finding is in line with other studies that suggest a negative association between physical and chemical restraints and QoL [18,40] and explains why the regular care ward performs less good in this regard than the special care ward. Agitation was also negatively associated with QoL. This is understandable as agitation is an expression of anxiety and indisposition of people with dementia and typically occurs after admission to hospital. Furthermore, agitation is often a reason for psychotropic drug use or physical restraint and, thus, also negatively affects QoL [52,53]. Independent from these factors, the special care ward itself shows the strongest impact on QoL, indicating that patients with dementia explicitly benefit from specialised care concepts. Other studies also report these benefits, both in a nursing home or hospital setting [54,55]. Since we controlled for patient characteristics like main diagnosis, age, functional limitations, chronic comorbidities, agitation, length of stay etc. in our model, we do not assume that the positive effect of the special care ward is completely a result of a biased sample between intervention and control group. Although the two compared hospitals differ in their structures and size, patients' characteristics are largely comparable between the samples in the control and intervention group. For instance, there is no substantial difference between the two hospitals regarding the relationship between functional impairments and physical restraints. Moreover, to see if the complexity of patients' need affects our findings, we calculated regression models with interaction terms between need factors moderated by hospitals (see Supplementary File 3). The association

between complexity of needs and QoL is not significantly different between the intervention and control group. Based on our results we suggest that the special care concept mainly explains the differences in the QoL. Although it is certainly difficult to determine the exact effect of the special care concept on the patients' QoL, our findings seem plausible in the light of the key elements of this intervention. A higher ratio of care staff as to patients, smaller facilities or systematically trained employees can be considered essential for health care provision to patients with dementia and are much better conditions for less physical or chemical restraints, independent of the functional limitations of patients. The special care ward provides a more dementia-friendly interior design, including orientation and navigation aids and the use of light and colours, which are considered as important components to reduce agitation for patients with dementia [56]. These findings and conclusions are in line with other studies on hospital care that suggest that an increased staff ratio or the implementation of multiple components, which particularly address the needs of patients with dementia, lead to reduced use of physical restraints and psychotropic drug use and improve the quality of care [57,58]. Furthermore, dementia-specific educational programmes, as implemented in the special care ward, have positive effects on nurses regarding their interaction with patients with dementia. Trained nurses can improve their coping skills in handling challenging behaviour of these patients, and better attend to the patients' unmet physical and psychological needs [59]. Studies suggest that the use of both physical and chemical restraints is reduced for nurses who completed a dementia-specific training as opposed to nurses who did not complete such an educational programme. Trained nurses had better skills in providing patient-centred care and thus improving the QoL for patients with dementia [59–61]. The special care ward benefits from a higher staff ratio, i.e. nurses have to care for fewer patients with dementia compared to the control group.

While this is an intentional element of the concept, the downside is higher personnel costs. Only few studies investigated the follow-up costs for patients with dementia in home care settings after hospitalization. Costa et al. predicted additional monthly costs in home care of about 445 Euros due to increased agitation of patients with dementia [62]. Thus, if patients with dementia benefit from special care concepts and perceive better outcomes in quality of life and care, the increased costs for more care personnel may be compensated by reducing follow-up costs for the ambulatory care. However, further research is needed to give more exact projections of the increased costs and potential of saving money.

Another finding is that the severity of cognitive impairments, measured with the MMSE, is a rather improper indicator to represent the underlying problems of and with the dementia disease, as these factors were not consistently associated with QoL. Direct measures of the problems associated with dementia, as agitation or challenging behaviour, should be considered as well when it comes to investigate the QoL of patients with dementia.

Our study has several limitations. One concerns the structural differences between the two hospitals. The hospital with the special care ward is much smaller than the hospital that hosted the control group. A second control group or an intervention group in a hospital of a similar size as the hospital with the regular care ward may have permitted a more distinct comparison. We tried to keep the impact of the structural differences as minimal as possible, for instance by accounting for many different patient characteristics including functional status, comorbidities and behavioural problems. Furthermore, the main diagnoses of patients were also considered in the analysis. We assume that we could at least partly adjust our analysis for a bias due to patient selection mechanisms. To validate our assumptions, we investigated to which extent the

association between patient characteristics and QoL is affected by differences between the control and intervention group (details shown in Supplementary File 3). Results suggest that our data provides no strong evidence for noticeably differences between the intervention and control group regarding the association between complexity of patients' needs and QoL. However, although we adjusted our analysis for many patient characteristics, we cannot eliminate a potential bias due to different hospital structures. In particular, the higher mean age and stronger functional limitations in the control group may indicate a selection bias in our sample. We suggest that further studies should take a second control group or a more comparable intervention group into account to gain more insight into potential biases due to structural differences of the control and intervention group. Another structural difference between the intervention and control group that certainly affects the results are the different staff-patient-ratios. In the special care ward, nurses have to care for fewer patients than in the regular care ward. Although we assume that this aspect probably has the highest impact on the outcomes in QoL, this is not a "selection bias" per se rather than a core component of the intervention. A higher staff-patient-ratio, dementia-specific training programmes, or a specific architectonical design are key elements of the special care concept, which, in their entirety, are reflected in the resulting differences between hospitals. A further limitation is possibly the first and thus rather exploratory use of the QUALIDEM assessment in a hospital setting. Although studies show reliable results of the QUALIDEM in nursing homes even for a short observation period of about one week [63], there are no studies that evaluate the reliability and validity for use in hospitals. We have done checks of internal consistencies, which showed that most subdomains of the QUALIDEM perform well with our data and are comparable to results from other validation studies [64]. This indicates that the use of the QUALIDEM is feasible for

hospital research. However, due to financial and logistic limitations, it was not possible to monitor the complete data collection and accurate completion of questionnaires. Hence, we cannot give evidence on the interrater reliability apart from the intense training of the study nurses. Finally, due to the nature of the study design, it was not possible that study nurses in the intervention and control group were blinded. This might affect the results insofar as study nurses may have generated more generous responses for the assessment scales [65].

Conclusions

On the whole, we think that a special care ward will improve the quality of care and is effective regarding the positive impact on the QoL of patients with dementia. Our study showed that after controlling for different predictors, the intervention still has a perceptible effect concerning clinical important differences in our outcome of interest, the patients' quality of life. However, such improvements can only be achieved by implementing a concept with multiple components that address the explicit needs of patients with dementia. The implementation of a special care concept usually increases the costs for hospitals because it requires a higher staff-patient-ratio, regular training of employees or more therapeutic offers. On the other hand, costs that accumulate in informal care after hospital stay as a result of poorer quality of care in hospitals can be much higher than additional personnel costs and could probably be reduced [62,66]. Health policies should consider the benefits of special care concepts and develop incentives for hospitals to improve the QoL and quality of care for patients with dementia.

Acknowledgement

The authors are grateful to the clinical staff of both hospitals who supported this study by helping with the data collection. Furthermore, we are thankful to the Authority of Health and Consumer Protection and the Homann-foundation for funding this study.

Author Contributions

DL was involved in conception and design of the study, analysis and interpretation of data and drafting the article. GP was involved in conception and design of the study and interpretation of data. JK was involved in interpretation of data and drafting the article. CK was involved in conception and design of the study, interpretation of data and drafting the article.

Competing interests

The authors have declared that no competing interests exist.

Funding

The present study was funded by the Behörde für Gesundheit und Verbraucherschutz (Authority of Health and Consumer Protection) in Hamburg and the Homann-foundation (grant number Z/65454/2014/F500-02.10/10,002). No one of the funders was involved in the data analysis process and did not influence or hold back results of the study.

Ethics approval and consent to participate

Prior to the study, a study protocol was developed and submitted to the ethical committee of the medical association of Hamburg ("Ethik-Kommission der Ärztekammer Hamburg"). The IRB of the ethical committee approved the proposal and attested that the study conforms to ethical and legal requirements (approval code PV5102). Study participants were not able to give their informed consent due to their cognitive impairments. However, as data mostly derived from the hospitals' regular documentation and was completely anonymous, the IRB of the ethics committee waived the need of an informed consent.

Availability of data and material

All data generated or analysed during this study are licensed under CC BY-NC 4.0 and available in the Zenodo repository (DOI: 10.5281/zenodo.1479677) at https://doi.org/10.5281/zenodo.1479676.

References

1 Recine U, Scotti E, Bruzzese V, *et al.* The change of hospital internal medicine: a study on patients admitted in internal medicine wards of 8 hospitals of the Lazio area, Italy. *Ital J Med* 2015;**9**:252. doi:10.4081/itjm.2015.523

2 Raveh D, Gratch L, Yinnon AM, *et al.* Demographic and clinical characteristics of patients admitted to medical departments. *J Eval Clin Pract* 2005;**11**:33–44. doi:10.1111/j.1365-2753.2004.00492.x

3 Sonnenblick M, Raveh D, Gratch L, *et al.* Clinical and demographic characteristics of elderly patients hospitalised in an internal medicine department in Israel: Characteristics of elderly patients in medical wards. *Int J Clin Pract* 2007;**61**:247–54. doi:10.1111/j.1742-1241.2006.00925.x

4 Mukadam N, Sampson EL. A systematic review of the prevalence, associations and outcomes of dementia in older general hospital inpatients. *Int Psychogeriatr* 2011;**23**:344–55. doi:10.1017/S1041610210001717

5 Zekry D, Herrmann FR, Grandjean R, *et al.* Does dementia predict adverse hospitalization outcomes? A prospective study in aged inpatients. *Int J Geriatr Psychiatry* 2009;**24**:283–91. doi:10.1002/gps.2104

6 Lyketsos CG, Sheppard JM, Rabins PV. Dementia in elderly persons in a general hospital. *Am J Psychiatry* 2000;**157**:704–7. doi:10.1176/appi.ajp.157.5.704

7 Sampson EL, Blanchard MR, Jones L, *et al.* Dementia in the acute hospital: prospective cohort study of prevalence and mortality. *Br J Psychiatry J Ment Sci* 2009;**195**:61–6. doi:10.1192/bjp.bp.108.055335

8 Pinkert C, Holle B. Menschen mit Demenz im Akutkrankenhaus: Literaturübersicht zu Prävalenz und Einweisungsgründen. *Z Für Gerontol Geriatr* 2012;**45**:728–34. doi:10.1007/s00391-012-0319-1

9 Dutzi I, Schwenk M, Micol W, *et al.* Patienten mit Begleitdiagnose Demenz. *Z Für Gerontol Geriatr* 2013;**46**:208–13. doi:10.1007/s00391-013-0483-y

10 Timmons S, O'Shea E, O'Neill D, *et al.* Acute hospital dementia care: results from a national audit. *BMC Geriatr* 2016;**16**:113. doi:10.1186/s12877-016-0293-3

11 Pi H-Y, Gao Y, Wang J, *et al.* Risk Factors for In-Hospital Complications of Fall-Related Fractures among Older Chinese: A Retrospective Study. *BioMed Res Int* 2016;**2016**:1–11. doi:10.1155/2016/8612143

12 Hu C-J, Liao C-C, Chang C-C, *et al.* Postoperative Adverse Outcomes in Surgical Patients with Dementia: A Retrospective Cohort Study. *World J Surg* 2012;**36**:2051–8. doi:10.1007/s00268-012-1609-x

13 Motzek T, Junge M, Marquardt G. Einfluss der Demenz auf Verweildauer und Erlöse im Akutkrankenhaus. *Z Für Gerontol Geriatr* 2017;**50**:59–66. doi:10.1007/s00391-016-1040-2

14 Gross AL, Jones RN, Habtemariam DA, *et al.* Delirium and Long-term Cognitive Trajectory Among Persons With Dementia. *Arch Intern Med* 2012;**172**:1324. doi:10.1001/archinternmed.2012.3203

15 Nakanishi M, Okumura Y, Ogawa A. Physical restraint to patients with dementia in acute physical care settings: effect of the financial incentive to acute care hospitals. *Int Psychogeriatr* 2017;:1–10. doi:10.1017/S104161021700240X

16 Krüger C, Mayer H, Haastert B, *et al.* Use of physical restraints in acute hospitals in Germany: a multi-centre cross-sectional study. *Int J Nurs Stud* 2013;**50**:1599–606. doi:10.1016/j.ijnurstu.2013.05.005

17 Heinze C, Dassen T, Grittner U. Use of physical restraints in nursing homes and hospitals and related factors: a cross-sectional study: Physical restraints in nursing homes and hospitals. *J Clin Nurs* 2012;**21**:1033–40. doi:10.1111/j.1365-2702.2011.03931.x

18 Beerens HC, Sutcliffe C, Renom-Guiteras A, *et al.* Quality of Life and Quality of Care for People With Dementia Receiving Long Term Institutional Care or Professional Home Care: The European RightTimePlaceCare Study. *J Am Med Dir Assoc* 2014;**15**:54–61. doi:10.1016/j.jamda.2013.09.010

19 Beerens HC, Zwakhalen SMG, Verbeek H, *et al.* Factors associated with quality of life of people with dementia in long-term care facilities: A systematic review. *Int J Nurs Stud* 2013;**50**:1259–70. doi:10.1016/j.ijnurstu.2013.02.005

20 Kricheldorff C, Hofmann W. Demenz im Akutkrankenhaus. *Z Für Gerontol Geriatr* 2013;**46**:196–7. doi:10.1007/s00391-013-0484-x

21 Treurniet HF, Essink-Bot M-L, Mackenbach JP, *et al.* Health-related quality of life: an indicator of quality of care? *Qual Life Res* 1997;**6**:363–9. doi:10.1023/A:1018435427116

22 Valderas JM, Alonso J. Patient reported outcome measures: a model-based classification system for research and clinical practice. *Qual Life Res* 2008;**17**:1125–35. doi:10.1007/s11136-008-9396-4

23 Tondi L, Ribani L, Bottazzi M, *et al.* Validation therapy (VT) in nursing home: A case-control study. *Arch Gerontol Geriatr* 2007;**44**:407–11. doi:10.1016/j.archger.2007.01.057

24 Kaiser AK, Hitzl W, Iglseder B. Three-question dementia screening: Development of the Salzburg Dementia Test Prediction (SDTP). *Z Für Gerontol Geriatr* 2014;**47**:577–82. doi:10.1007/s00391-013-0568-7

25 Ettema TP, Dröes R-M, de Lange J, *et al.* QUALIDEM: development and evaluation of a dementia specific quality of life instrument. Scalability, reliability and internal structure. *Int J Geriatr Psychiatry* 2007;**22**:549–56. doi:10.1002/gps.1713

26 Dichter MM, Ettema TP, Schwab CGG, *et al.* Benutzerhandbuch für die deutschsprachige QUALIDEM Version 2.0. Witten: 2016.

27 Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;**12**:189–98.

28 Dichter MN, Meyer G. Quality of Life of People with Dementia in Nursing Homes. In: Schüssler S, Lohrmann C, eds. *Dementia in Nursing Homes*. Cham: Springer 2017. 139–57. doi:10.1007/978-3-319-49832-4_11

29 Charlson ME, Pompei P, Ales KL, *et al.* A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;**40**:373–83.

30 Charlson ME, Charlson RE, Peterson JC, *et al.* The Charlson comorbidity index is adapted to predict costs of chronic disease in primary care patients. *J Clin Epidemiol* 2008;**61**:1234–40. doi:10.1016/j.jclinepi.2008.01.006

31 Mahoney FI, Barthel DW. Functional Evaluation: The Barthel Index. *Md State Med J* 1965;**14**:61–5.

32 Deutsches Institut f. medizinische Dokumentation u. Information (DIMDI). *ICD-10-GM 2017 Systematisches Verzeichnis*. Lich, Hess: pictura Werbung 2016.

33 Rosen J, Burgio L, Kollar M, *et al.* The Pittsburg Agitation Scale: A user-friendly instrument for rating agitation in dementia patients. *Am J Geriatr Psychiatry Off J Am Assoc Geriatr Psychiatry* 1994;**2**:52–9.

34 WHO Collaborating Centre for Drug Statistics Methodology, editor. *Guidelines for ATC classification and DDD assignment 2013*. Oslo: 2012.

35 Buuren S van, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. *J Stat Softw* 2011;**45**. doi:10.18637/jss.v045.i03

36 White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med* 2011;**30**:377–99. doi:10.1002/sim.4067

37 Carpenter B, Gelman A, Hoffman MD, *et al.* Stan: A Probabilistic Programming Language. *J Stat Softw* 2017;**76**. doi:10.18637/jss.v076.i01

38 Gelman A, Carlin JB, Stern HS, *et al. Bayesian data analysis*. Third edition. Boca Raton: CRC Press 2014.

39 Barca ML, Engedal K, Laks J, *et al.* Quality of Life among Elderly Patients with Dementia in Institutions. *Dement Geriatr Cogn Disord* 2011;**31**:435–42. doi:10.1159/000328969

40 Beer C, Flicker L, Horner B, *et al.* Factors Associated with Self and Informant Ratings of the Quality of Life of People with Dementia Living in Care Facilities: A Cross Sectional Study. *PLoS ONE* 2010;**5**:e15621. doi:10.1371/journal.pone.0015621

41 McElreath R. *Statistical rethinking: a Bayesian course with examples in R and Stan*. Boca Raton London New York: CRC Press, Taylor & Francis Group 2016.

42 Hox JJ. *Multilevel analysis: techniques and applications*. 2nd ed. New York: Routledge 2010.

43 R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing 2018. https://www.R-project.org/

44 Wickham H. *ggplot2: Elegant Graphics for Data Analysis.* 2nd ed. New York: Springer 2016.

45 Bürkner P-C. brms: An R Package for Bayesian Multilevel Models Using Stan. *J Stat Softw* 2017;**80**. doi:10.18637/jss.v080.i01

46 Lüdecke D. *sjPlot: Data Visualization for Statistics in Social Science.* 2018. doi:10.5281/zenodo.1308157

47 Lüdecke D. Dataset: Quality Of Life Of Patients With Dementia In Acute Hospitals-Version 1.0. 2018. doi:10.5281/zenodo.1479677

48 Thorgrimsen L, Selwood A, Spector A, *et al.* Whose quality of life is it anyway? The validity and reliability of the Quality of Life-Alzheimer's Disease (QoL-AD) scale. *Alzheimer Dis Assoc Disord* 2003;**17**:201–8.

49 Selwood A, Thorgrimsen L, Orrell M. Quality of life in dementia--a one-year follow-up study. *Int J Geriatr Psychiatry* 2005;**20**:232–7. doi:10.1002/gps.1271

50 Hoe J, Hancock G, Livingston G, *et al.* Changes in the quality of life of people with dementia living in care homes. *Alzheimer Dis Assoc Disord* 2009;**23**:285–90. doi:10.1097/WAD.0b013e318194fc1e

51 Norman GR, Sloan JA, Wyrwich KW. Interpretation of Changes in Health-related Quality of Life: The Remarkable Universality of Half a Standard Deviation. *Med Care* 2003;**41**:582–92. doi:10.1097/01.MLR.0000062554.74615.4C

52 Lam K, Kwan JSK, Wai Kwan C, *et al.* Factors Associated With the Trend of Physical and Chemical Restraint Use Among Long-Term Care Facility Residents in Hong Kong: Data From an 11-Year Observational Study. *J Am Med Dir Assoc* 2017;**18**:1043–8. doi:10.1016/j.jamda.2017.06.018

53 Ó Flatharta T, Haugh J, Robinson SM, *et al.* Prevalence and predictors of bedrail use in an acute hospital. *Age Ageing* 2014;**43**:801–5. doi:10.1093/ageing/afu081

54 Kok JS, Berg IJ, Scherder EJA. Special Care Units and Traditional Care in Dementia: Relationship with Behavior, Cognition, Functional Status and Quality of Life - A Review. *Dement Geriatr Cogn Disord Extra* 2013;**3**:360–75. doi:10.1159/000353441

55 von Renteln-Kruse W, Neumann L, Klugmann B, *et al.* Geriatric Patients With Cognitive Impairment. *Dtsch Ärztebl Int* 2015;**112**:103–12. doi:10.3238/arztebl.2015.0103

56 Büter K, Motzek T, Dietz B, *et al.* Demenzsensible Krankenhausstationen: Expertenempfehlungen zu Planung und Gestaltung. *Z Für Gerontol Geriatr* 2017;**50**:67–72. doi:10.1007/s00391-016-1079-0

57 Blair A, Anderson K, Bateman C. The "Golden Angels": effects of trained volunteers on specialling and readmission rates for people with dementia and delirium in rural hospitals. *Int Psychogeriatr* 2018;:1–10. doi:10.1017/S1041610218000911

58 Sinvani L, Warner-Cohen J, Strunk A, *et al.* A Multicomponent Model to Improve Hospital Care of Older Adults with Cognitive Impairment: A Propensity Score-Matched

Analysis. *J Am Geriatr Soc* Published Online First: 11 August 2018. doi:10.1111/jgs.15452

59 Schindel Martin L, Gillies L, Coker E, *et al.* An Education Intervention to Enhance Staff Self-Efficacy to Provide Dementia Care in an Acute Care Hospital in Canada: A Nonrandomized Controlled Study. *Am J Alzheimers Dis Dementiasr* 2016;**31**:664–77. doi:10.1177/1533317516668574

60 Matsuda Y, Hashimoto R, Takemoto S, *et al.* Educational benefits for nurses and nursing students of the dementia supporter training program in Japan. *PLOS ONE* 2018;**13**:e0200586. doi:10.1371/journal.pone.0200586

61 Galvin JE, Kuntemeier B, Al-Hammadi N, *et al.* "Dementia-friendly Hospitals: Care not Crisis." *Alzheimer Dis Assoc Disord* 2010;**24**:372–9. doi:10.1097/WAD.0b013e3181e9f829

62 Costa N, Wübker A, De Mauléon A, *et al.* Costs of Care of Agitation Associated With Dementia in 8 European Countries: Results From the RightTimePlaceCare Study. *J Am Med Dir Assoc* 2018;**19**:95.e1-95.e10. doi:10.1016/j.jamda.2017.10.013

63 Dichter MN, Schwab CGG, Meyer G, *et al.* Item distribution, internal consistency and inter-rater reliability of the German version of the QUALIDEM for people with mild to severe and very severe dementia. *BMC Geriatr* 2016;**16**. doi:10.1186/s12877-016-0296-0

64 Dichter M, Dortmann O, Halek M, *et al.* Scalability and internal consistency of the German version of the dementia-specific quality of life instrument QUALIDEM in nursing homes – a secondary data analysis. *Health Qual Life Outcomes* 2013;**11**:91. doi:10.1186/1477-7525-11-91

65 Schulz KF, Grimes DA. Blinding in randomised trials: hiding who got what. *The Lancet* 2002;**359**:696–700. doi:10.1016/S0140-6736(02)07816-9

66 Leon J, Cheng CK, Neumann PJ. Alzheimer's disease care: costs and potential savings. *Health Aff Proj Hope* 1998;**17**:206–16.

Figure Titles and Legends

Figure 1: QUALIDEM subdomain scores by care ward in patients with mild to severe dementia (MMSE score from 7 to 27, n = 400)

Figure 2: QUALIDEM subdomain scores by care ward, patients with very severe dementia (MMSE score < 7, n = 126)

Figure 3: Predictors of Health Related Quality of Life, Regression Coefficients, Bayesian Linear Mixed Model, Posterior Median (+50% and 89% High Density Interval)

Supplementary Files

Supplementary File 1: Figure, Distribution of Patients' Main Diagnosis by Hospital

Supplementary File 2: Methodological comments, Word document (docx-format).

Supplementary File 3: Regression Models with Interaction Terms, Word document (docx-format).

Supplementary File 4: R Source Code (to use with R statistics, CC BY-NC 4.0 license)

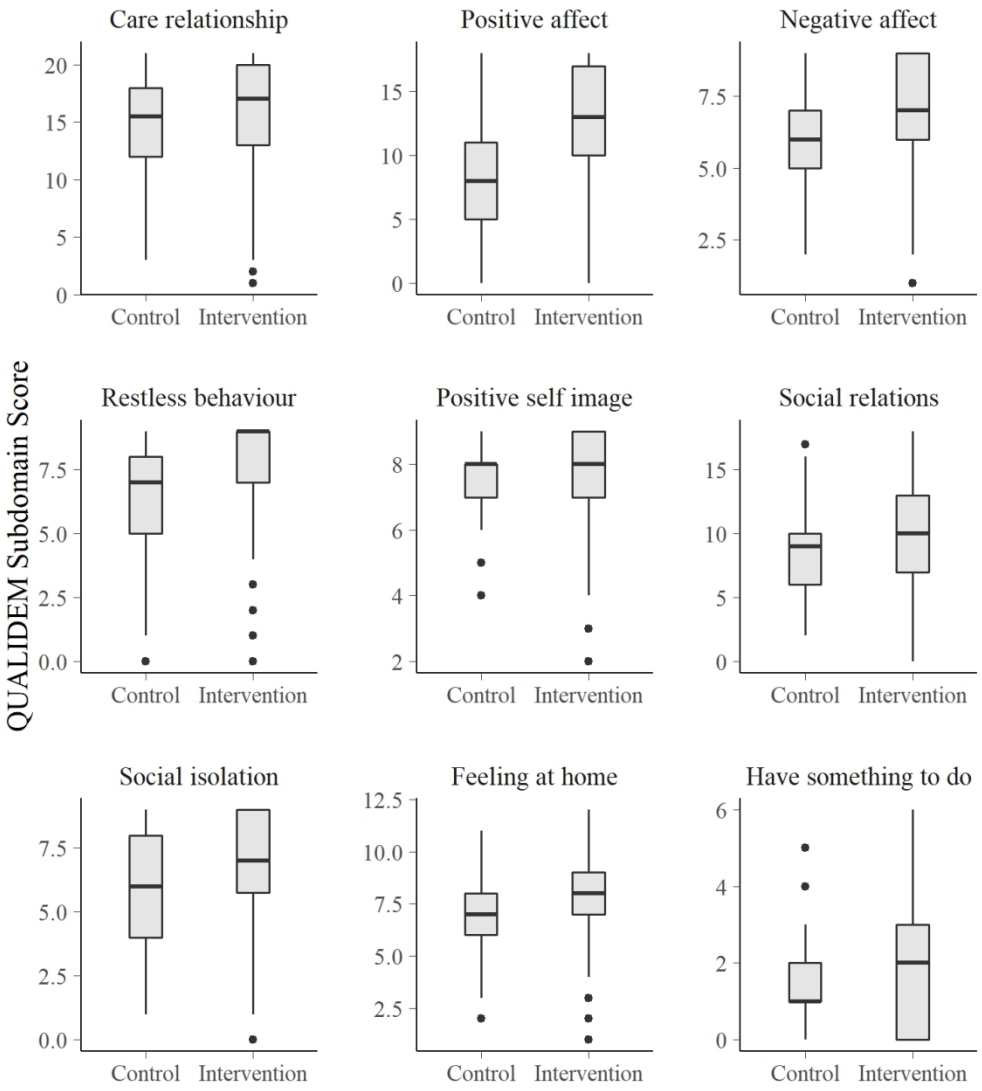


Figure 1: QUALIDEM subdomain scores by care ward in patients with mild to severe dementia (MMSE score from 7 to 27, n = 400)

169x189mm (220 x 220 DPI)

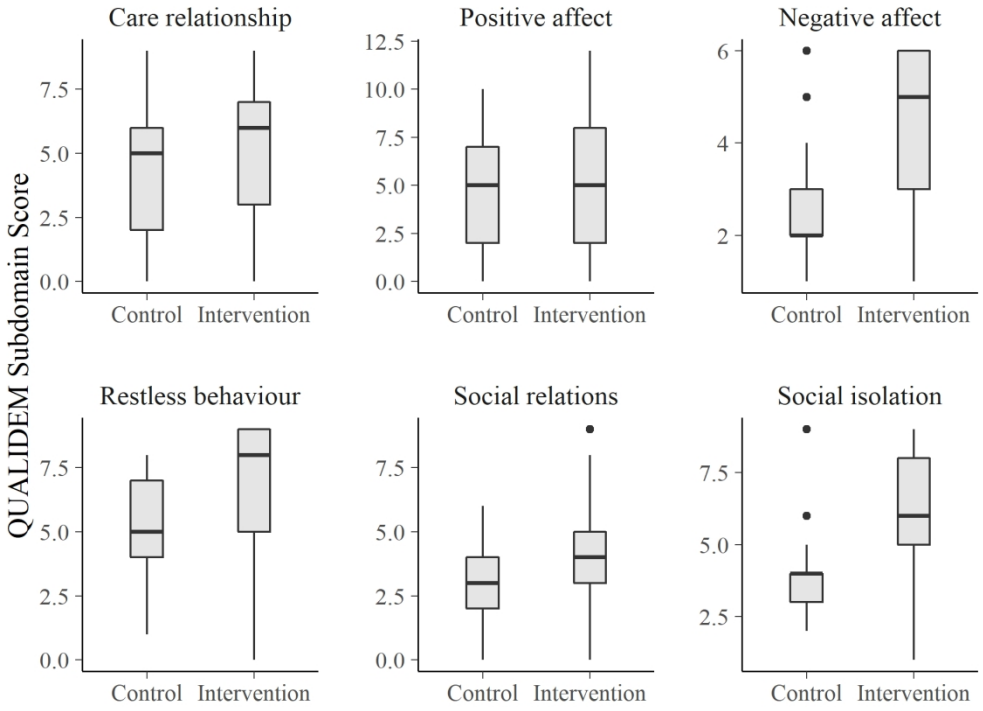


Figure 2: QUALIDEM subdomain scores by care ward, patients with very severe dementia (MMSE score < 7, n = 126)

169x125mm (300 x 300 DPI)

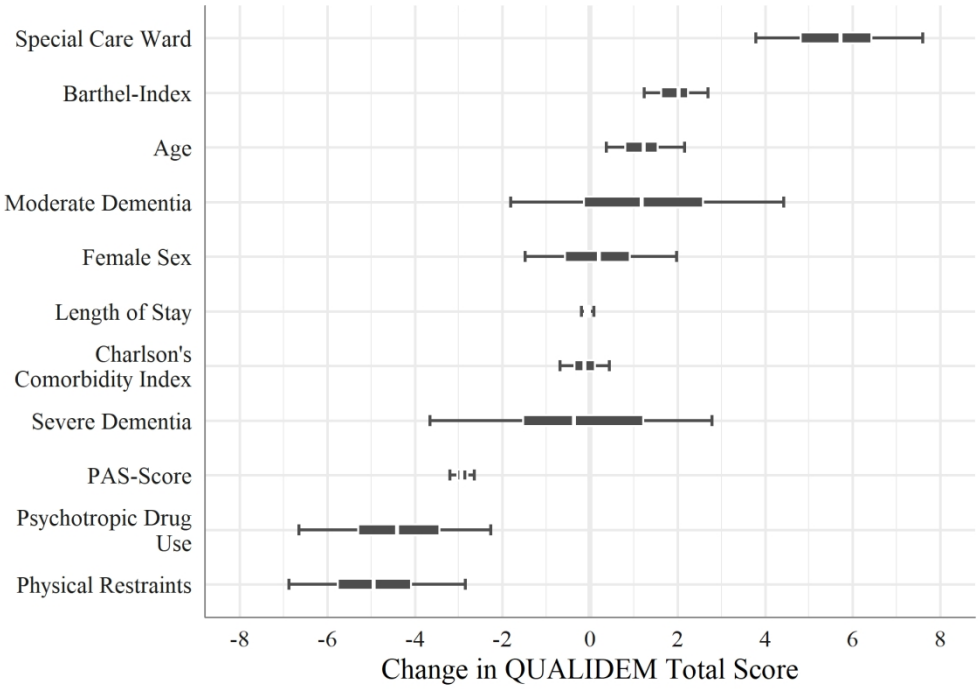
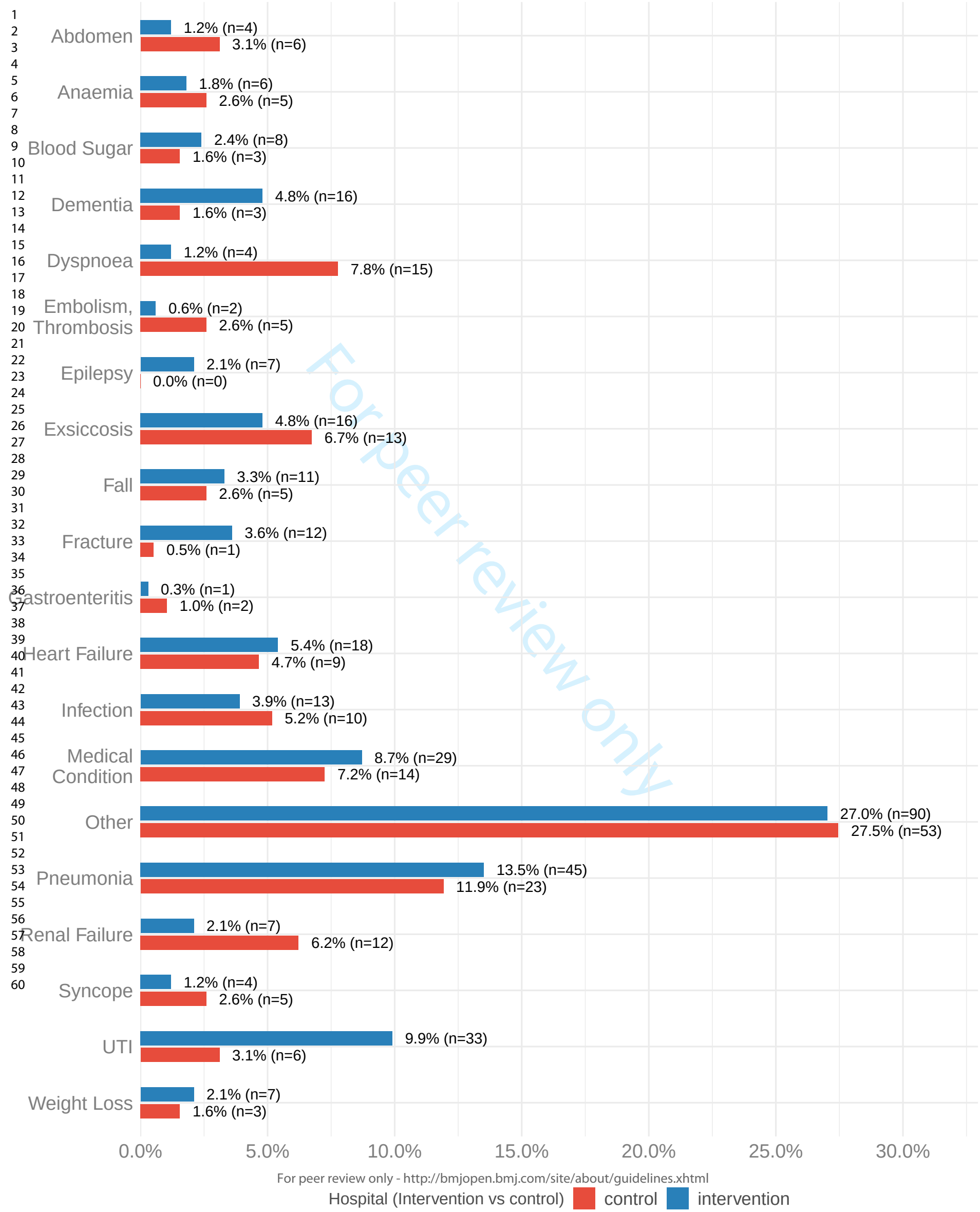


Figure 3: Predictors of Health Related Quality of Life, Regression Coefficients, Bayesian Linear Mixed Model, Posterior Median (+50% and 89% High Density Interval)

169x119mm (300 x 300 DPI)

Supplemental Material 1

Main Diagnosis by Hospital Ward



Supplemental Material 2

Quality of Life of Patients with Dementia in Acute Hospitals: Comparing a regular ward to a special care ward with dementia care concept. A Bayesian Multilevel Model Analysis.

Methodological comments

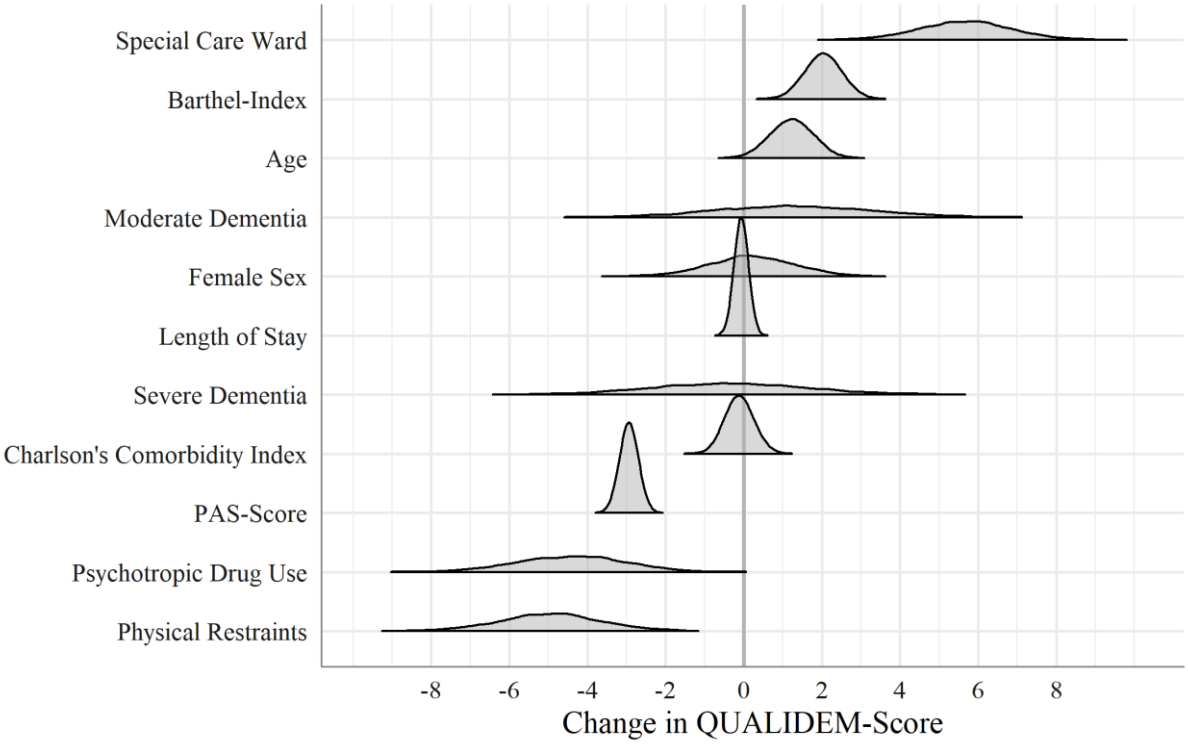
1. Short note on Bayesian Methods

Bayesian methods are probably not very common, and readers may ask "Why use Bayesian regression models?" Gelman et al.[1] give a well summarized answer to this question: "A pragmatic rationale for the use of Bayesian methods is the inherent flexibility introduced by their incorporation of multiple levels of randomness and the resultant ability to combine information from different sources, while incorporating all reasonable sources of uncertainty in inferential summaries. Such methods naturally lead to smoothed estimates in complicated data structures and consequently have the ability to obtain better real-world answers."[1]. Shortly speaking, advantages of Bayesian methods are, for instance, the ability to derive probability statements for every quantity of interest or explicitly incorporate prior knowledge about parameters into the model.

2. Distribution of Posterior Samples from the Regression Model

Contrary to a frequentist approach, there is no unique point estimate in Bayesian analysis. The posterior median minimises the expected absolute error (i.e. the difference of estimates from true values over posterior samples), but there are many other plausible values to describe the association between predictors and outcome, which is called the *distribution of posterior samples*. Credible or Highest Density Intervals are based on quantiles of this distribution and indicate the probability of an estimate being inside this interval. In terms of inference, this allows us to say that with, e.g. 50% probability, the true effect of a *Special Care Ward* on our outcome lies between 5.1 and 6.6. Figure S1 shows the distribution of posterior samples from our regression model.

Figure S1: *Distribution of Posterior Samples from Regression Model*



3. Tabular Summary from the Regression Model

Table S1 provides a summary of the regression model, including the data for two different Highest Density Intervals as well as the ratio of effective number of samples. The ratio should be 1.0 or close to 1.0 and indicates whether the samples were efficient, which means that the results for the related coefficient are reliable. Furthermore, all Rhat values of the models were approximately 1. The Rhat statistic[2] measures the ratio of the average variance of the draws within each chain to the variance of the pooled draws across chains. If the chains have not converged to a common distribution, the Rhat statistic will tend to be greater than one, indicating that more iterations may be required for reliable results.

Table S1: *Predictors of Quality of Life, Regression Coefficients, Bayesian Linear Mixed Model, Posterior Median (+50% and 89% Highest Density Intervals)*

Term	*Estimate*	*SE*	*50% HDI*	*89% HDI*	*Ratio of effective Number of Samples*
(Intercept)	46.2	2.2	45.2 – 48.2	42.7 – 49.8	1.00
Length of Stay	-0.1	0.1	-0.1 – 0.0	-0.2 – 0.1	1.00
Age	1.2	0.5	0.8 – 1.5	0.4 – 2.1	1.00
Moderate Dementia	1.2	2.0	-0.1 – 2.7	-1.8 – 4.6	1.00
Severe Dementia	-0.4	2.0	-1.6 – 1.1	-3.6 – 2.7	1.00
Female	0.2	1.1	-0.5 – 1.0	-1.6 – 1.9	1.00
Barthel Score	2.0	0.5	1.8 – 2.4	1.3 – 2.8	1.00
Physical Restraints (yes)	-4.9	1.2	-5.8 – -4.1	-7.0 – -2.8	1.00
Special Care Ward (Intervention)	5.7	1.2	4.9 – 6.5	3.8 – 7.6	1.00
PAS-Score	-2.9	0.2	-3.1 – -2.8	-3.2 – -2.7	1.00
Charlson's Comorbidity Index	-0.1	0.3	-0.4 – 0.1	-0.6 – 0.5	1.00
Psychotropic Drug Use (yes, as-needed)	-4.4	1.4	-5.3 – -3.5	-6.5 – -2.1	1.00
sigma	11.9	0.4	11.5 – 12.0	11.3 – 12.5	1.00

All Rhat values ~ 1, all mcse < 0.05. Results based on 4 chains each 1000 iterations for each chain. LOO-adjusted R^2: 0.500

4. Test for Practical Equivalence of Parameters

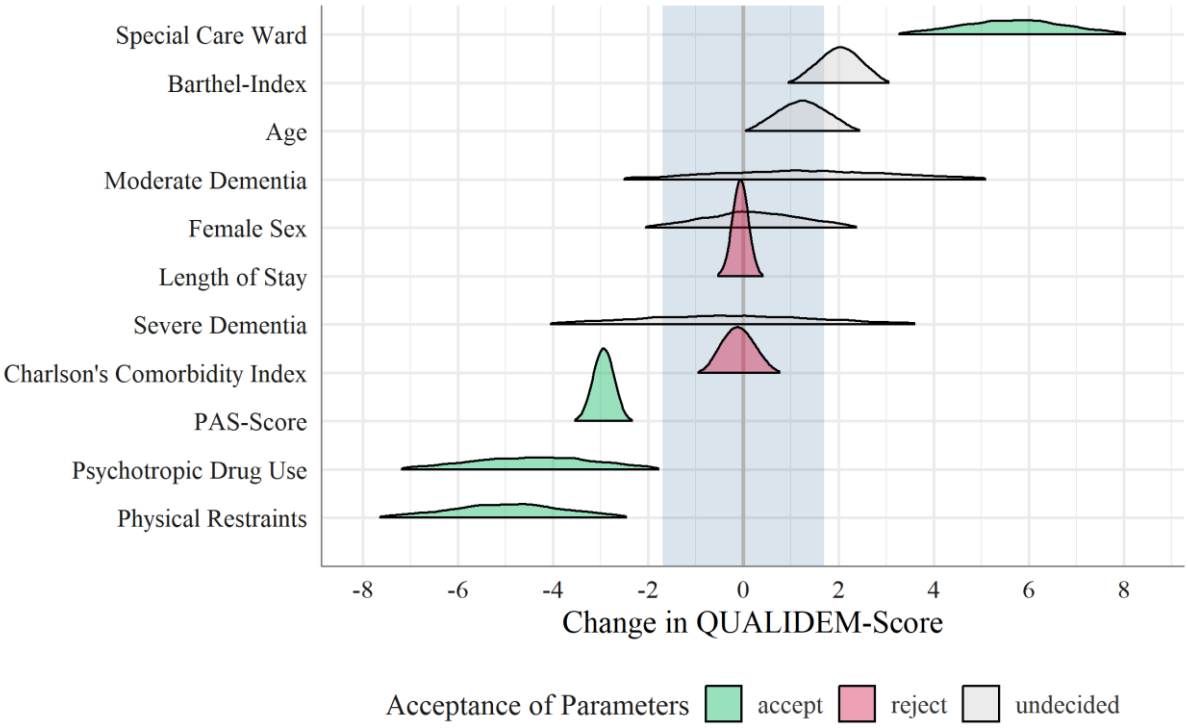
Bayesian methods do not perform classical "null hypothesis significance tests". Rather, to account for the uncertainty in estimation and the magnitude of parameter values, Kruschke[3] suggests checking whether parameter values lie inside a certain range that is considered as "practically no effect".

Figure S2 shows the distribution of the 95% Highest Density Intervals of posterior samples and indicates whether these lie completely outside, completely inside or partially inside the region of practical equivalence (ROPE). If the HDI is completely outside the ROPE, the "null hypothesis" for this parameter is "rejected". If the ROPE completely covers the HDI, i.e. all most credible values of a parameter are inside the region of practical equivalence, the null hypothesis is accepted. Else, it is undecided whether to accept or reject the null hypothesis.

The ROPE limits are specified following Kruschke's suggestion, which is defined as *0 +/- SD(dependent variable) * 0.1* for linear models.

Physical restraint, psychotropic drug use, agitation (PAS-score) and the intervention (special care ward) are those predictors that can be considered as relevant for our outcome.

Figure S2: *95%-Range of the Distribution of Posterior Samples from Regression Model; Region of Practical Equivalence emphasized in light-blue.*



5. Summary of the Distribution of Prior and Posterior Samples

Bayesian regression models can incorporate prior information about the parameters. Informative or weakly informative priors have the advantage of regularization of parameters to rule out large effects. This leads to less biased results and reductions of outliers in situations where the sample size in general, or for certain groups in the data, is small. Where we found information about our parameters, we included it by specifying *informative priors*. For instance, studies suggest that physical restraints are, on average, associated with an about 5-point reduction in quality of life, so the *location* of the prior distribution of our parameter physical restraint is at -5.0 (see Table S2). *Weakly informative priors* simply have a location of zero.

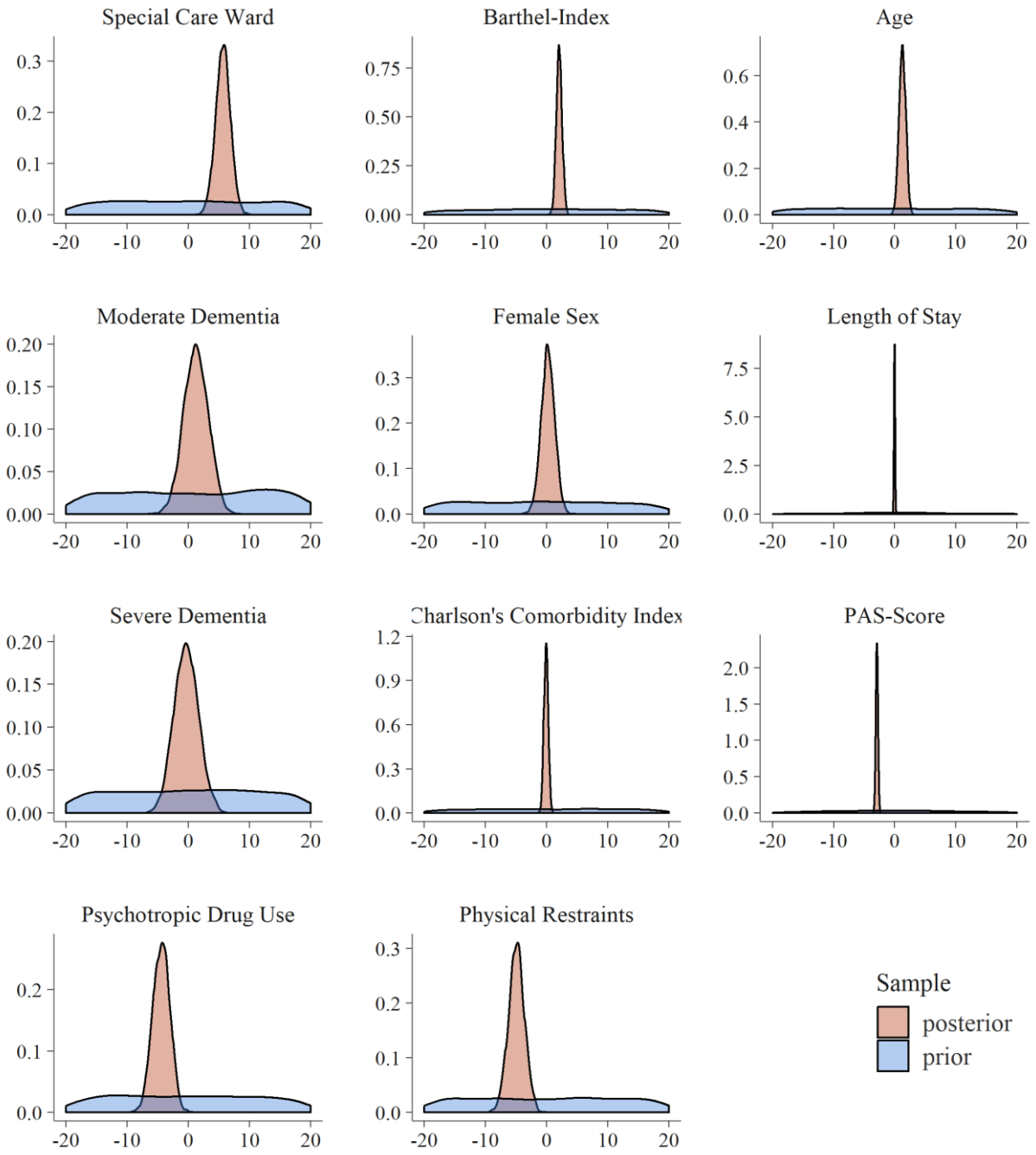
Table S2: *Prior Summary*

Predictor	Location	Scale
Length of Stay	0.00	6
Age	0.16	40
Moderate dementia	0.00	42
Severe dementia	-0.44	42
Sex, female	-3.22	42
Barthel-Score	0.00	29
Physical restraints	-5.00	42
Special Care Ward	0.00	42
PAS-Score	0.00	13
Charlson's Comorbidity Index	0.00	27
Psychotic Drug Use	0.29	42

To prevent a too strong shrinkage of uncertainty, the *scale* of the prior distribution should be large enough. If no assumptions about the scale of the effect are made, a recommended way (for normally distributed priors) is to use a scale of 2.5 and adjust it by dividing the scale by the standard deviation of the related predictor and multiply it with the standard deviation of the outcome (*2.5 * SD(y) / SD(x)*, see https://cran.r-project.org/web/packages/rstanarm/vignettes/priors.html).

When the evidence in the data is *strong*, the impact of the prior information is rather low. On the other hand, if the evidence of the data is *weak*, for instance due to small sample size, the impact of the prior information increases. Figure S3 shows the summary of the prior and posterior sample distributions and suggests that the evidence in the data was rather strong and results were not biased due to unrealistic prior information.

Figure S3: *Posterior versus Prior Summary*

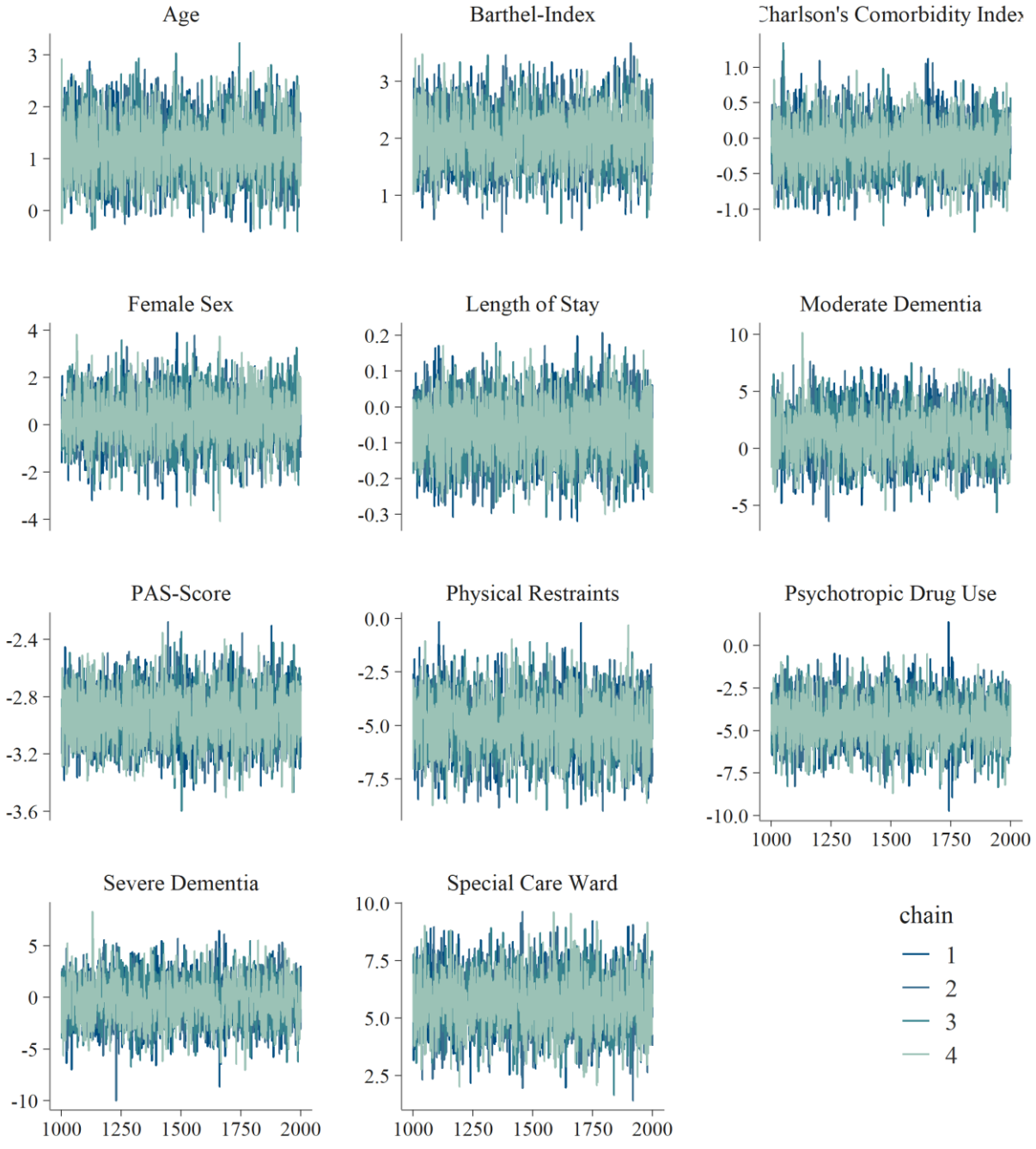


6. Trace plots

A trace plot provides a visual way to inspect sampling the behaviour and can be used, besides the Rhat-statistics, to assess convergence. The desired pattern is a "fat, hairy caterpillar", which shows no suspicious bends4, 5. If the chains converged to the posterior distribution, we can be confident in our model-based inferences.

The trace plot in Figure S4 indicates that all chains have converged well.

Figure S4: *Markov Chain trace plot, 4 chains with 1000 iterations each (excluding warmup-samples)*



7. References

1. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. Bayesian Data Analysis. Third edition. Boca Raton: CRC Press; 2014: p.24

2. Gelman A, Rubin DB. Inference from Iterative Simulation Using Multiple Sequences. Statistical Science. 1992;7(4):457-472. doi:10.1214/ss/1177011136

3. Kruschke JK. Rejecting or Accepting Parameter Values in Bayesian Estimation. Advances in Methods and Practices in Psychological Science. 2018; doi: 10.1177/2515245918771304

4. Sorensen T, Hohenstein S, Vasishth S. Bayesian linear mixed models using Stan: A tutorial for psychologists, linguists, and cognitive scientists. The Quantitative Methods for Psychology. 2016;12:175–200.

5. Lunn D, Jackson C, Best N, Thomas A, Spiegelhalter D. The BUGS book: a practical introduction to Bayesian analysis. Boca Raton, FL: CRC Press, Taylor & Francis Group; 2013.

Supplemental Material 3

To see if the complexity of patients' needs affect our findings, we calculated regression models with interaction terms between need factors moderated by hospitals. The figures S5 to S8 show the results of the interaction terms. Table S3 shows the coefficients in details. We found no "significant" associations for the interaction terms, concluding that the association between complexity of patients' needs and quality of life is not differing noticeably between the intervention and control group.

Plots were creates using the *ggeffects* package in R (Lüdecke D. ggeffects: Tidy Data Frames of Marginal Effects from Regression Models. Journal of Open Source Software. 2018;3: 772. doi:10.21105/joss.00772).

Figure S5: *Interaction between Barthel-Index and Intervention/Control-Group*

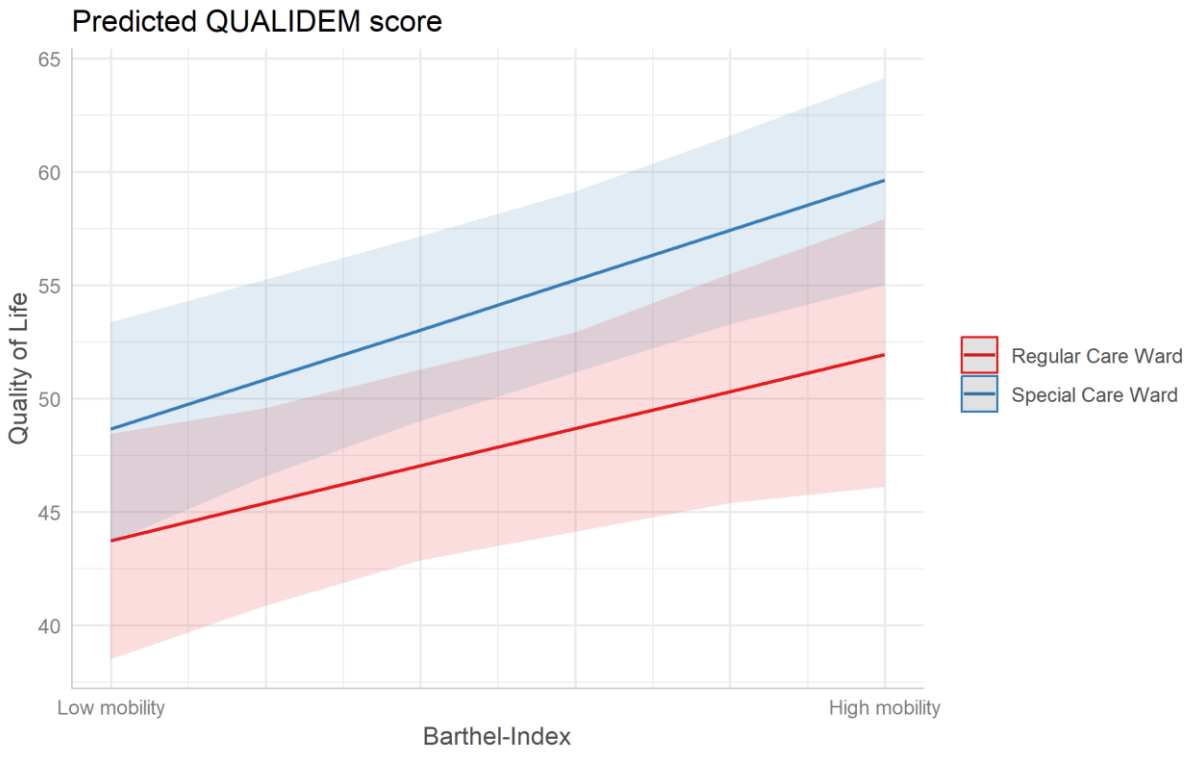


Figure S6: *Interaction between Physical Restraints and Intervention/Control-Group*

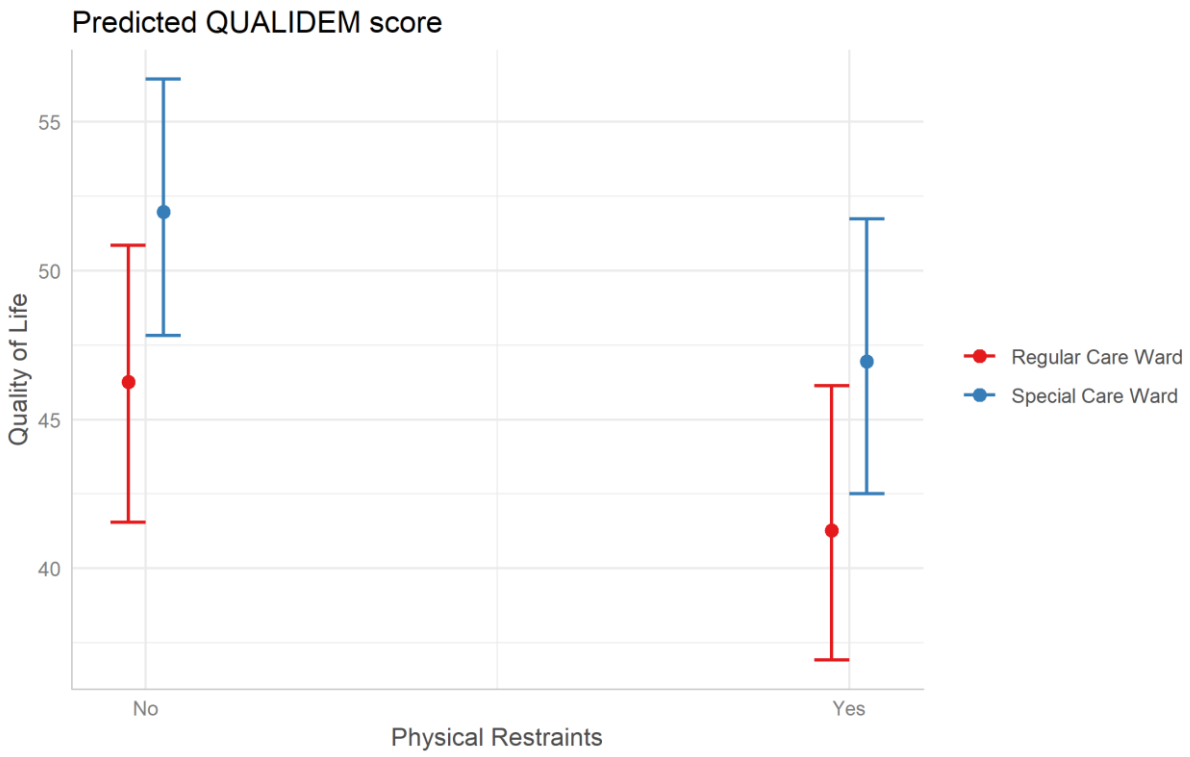


Figure S7: *Interaction between PAS and Intervention/Control-Group*

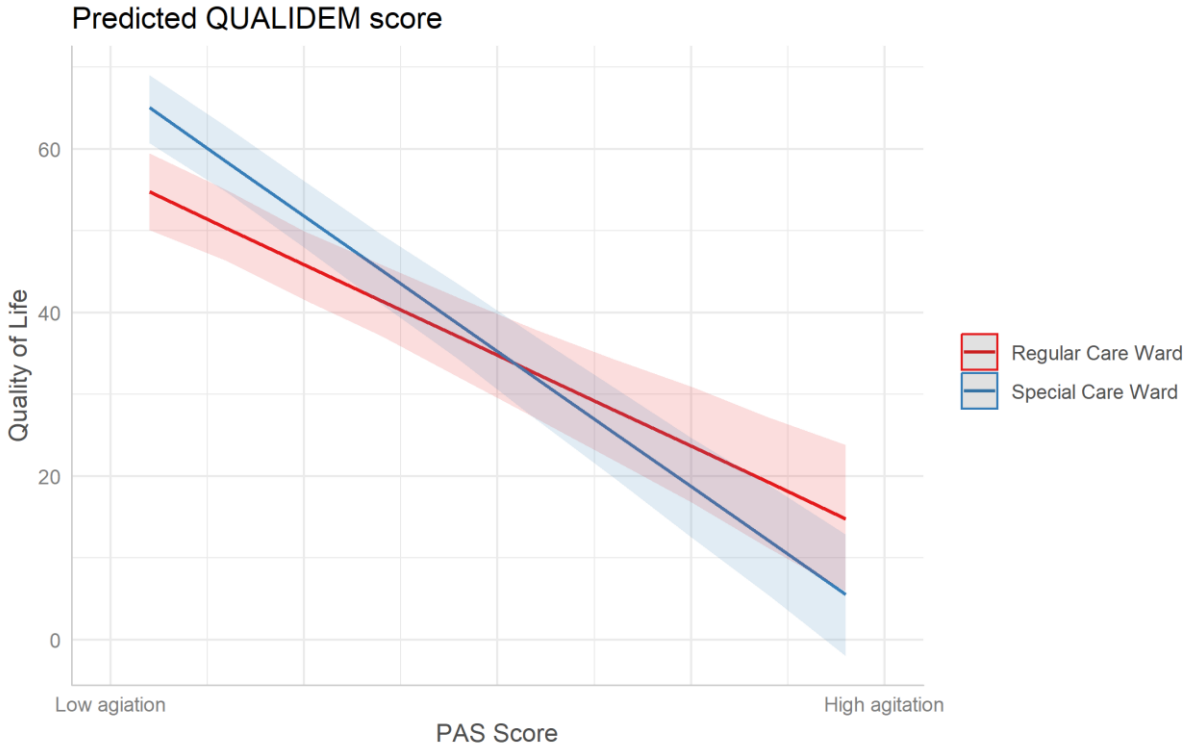


Figure S8: *Interaction between Chemical Restraints and Intervention/Control-Group*

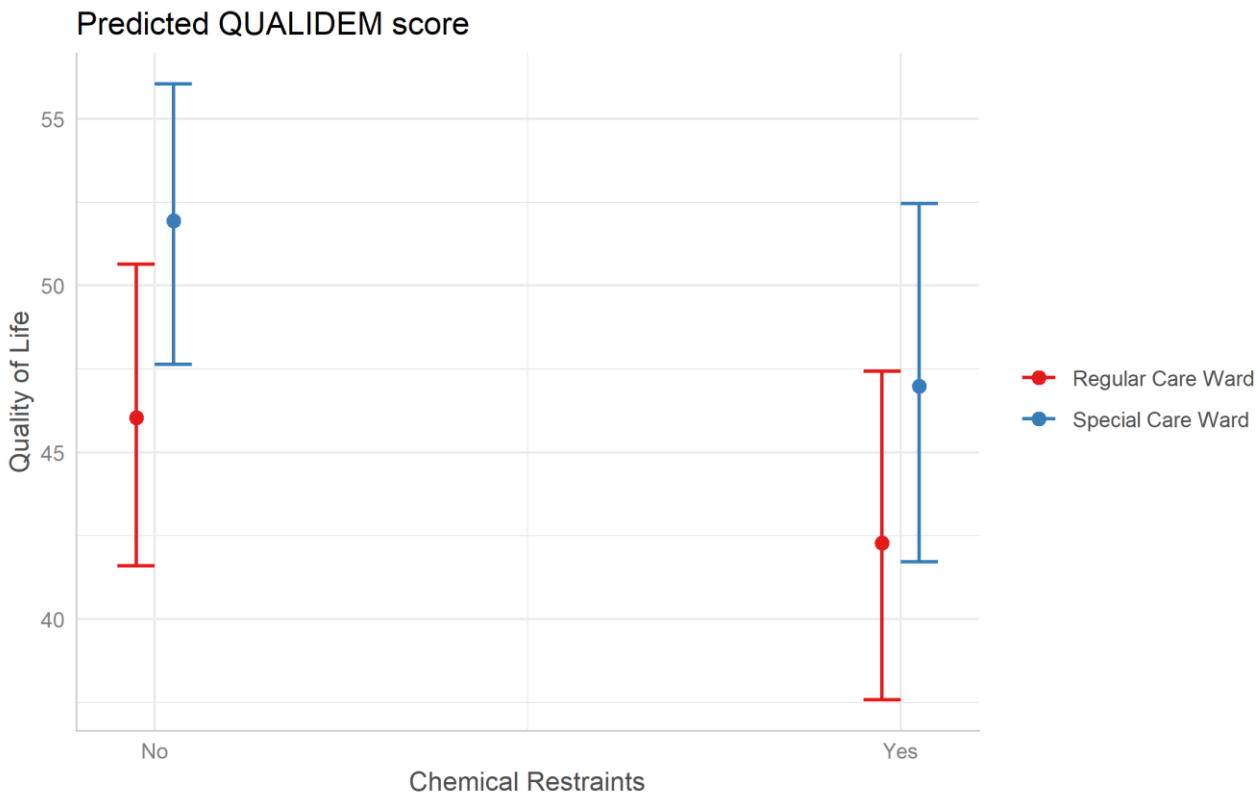


Table S3: *Coefficients for Models with Interaction Terms*

Terms	*Model: Interaction between Barthel and Hospital*		*Model: Interaction between Physical Restraints and Hospital*		*Model: Interaction between PAS and Hospital*		*Model: Interaction between Chemical Restraints and Hospital*	
	Estimates	*HDI (95%)*	*Estimates*	*HDI (95%)*	*Estimates*	*HDI (95%)*	*Estimates*	*HDI (95%)*
(Intercept)	36.18	26.19 – 45.67	36.31	27.00 – 46.10	36.57	27.11 – 45.71	36.16	26.45 – 45.69
Length of Stay	-0.07	-0.23 – 0.11	-0.06	-0.23 – 0.09	-0.07	-0.23 – 0.10	-0.06	-0.22 – 0.10
Age	0.12	0.01 – 0.23	0.12	0.02 – 0.23	0.12	0.01 – 0.22	0.12	0.02 – 0.23
Moderate Dementia	1.21	-2.79 – 5.13	1.19	-2.72 – 5.24	1.19	-2.83 – 4.72	1.25	-2.95 – 4.89
Severe Dementia	-0.34	-4.30 – 3.70	-0.44	-4.45 – 3.65	-0.74	-4.54 – 3.10	-0.30	-4.45 – 3.59
Female	0.16	-2.06 – 2.29	0.19	-1.98 – 2.29	0.52	-1.53 – 2.70	0.23	-1.76 – 2.53
Charlson's Comorbidity Index	-0.12	-0.81 – 0.59	-0.11	-0.77 – 0.64	-0.11	-0.77 – 0.59	-0.11	-0.82 – 0.59
Barthel Score	1.66	0.27 – 2.99	2.02	1.15 – 2.88	2.07	1.17 – 2.97	2.05	1.13 – 2.98
Physical Restraints (yes)	5.74	3.48 – 7.90	5.76	2.67 – 8.72	5.95	3.64 – 8.18	5.94	3.44 – 8.62
Special Care Ward (Intervention)	-5.02	-7.65 – -2.54	-4.92	-8.57 – -1.11	-4.94	-7.60 – -2.41	-4.96	-7.83 – -2.58
PAS-Score	-2.96	-3.30 – -2.62	-2.94	-3.30 – -2.61	-2.22	-2.77 – -1.66	-2.93	-3.27 – -2.59
Psychotropic Drug Use (yes, as-needed)	-4.43	-7.10 – -1.83	-4.40	-7.08 – -1.87	-4.30	-6.81 – -1.56	-3.76	-7.83 – -0.03
Barthel * Intervention	0.52	-1.00 – 2.06						
Phys. Restr. * Intervention			-0.09	-4.46 – 4.66				
PAS * Intervention					-1.09	-1.76 – -0.42		
Chem. Restr. * Intervention							-1.15	-6.63 – 4.27

Supplemental Material 4 – R Source Code

```
library(tidyverse)
library(ggridges)
library(sjmisc)
library(sjlabelled)
library(sjstats)
library(sjPlot)
library(insight)
library(bayestestR)
library(brms)

# Data available at https://doi.org/10.5281/zenodo.1479676

# load data ----

load("Dataset.RData")

# divide age by 10

d$age10 <- d$age / 10


# Labels for final model ----

labs <-
  c(
    stay_c = "Length of Stay",
    age = "Age",
    age10 = "Age",
    mmse2 = "Moderate Dementia",
    mmse3 = "Severe Dementia",
    sex2 = "Female Sex",
    barthel_code = "Barthel-Index",
    groupintervention = "Special Care Ward",
    physres1 = "Physical Restraints",
    pas_c = "PAS-Score",
    cci_c = "Charlson's Comorbidity Index",
    chemicalres1 = "Psychotropic Drug Use", # oder as-needed
    b_stay_c = "Length of Stay",
    b_age = "Age",
    b_age10 = "Age",
    b_mmse2 = "Moderate Dementia",
    b_mmse3 = "Severe Dementia",
    b_sex2 = "Female Sex",
    b_barthel_code = "Barthel-Index",
    b_groupintervention = "Special Care Ward",
    b_physres1 = "Physical Restraints",
    b_pas_c = "PAS-Score",
    b_cci_c = "Charlson's Comorbidity Index",
    b_chemicalres1 = "Psychotropic Drug Use" # oder as-needed
  )

# prior-definition in brms ----

# scale is 2.5 * sd(y) / sd(x)

bprior <-
  prior(normal(0, 6), class = "b", coef = "stay_c") +
  prior(normal(.1554, 40), class = "b", coef = "age10") +
  prior(normal(0, 42), class = "b", coef = "mmse2") +
  prior(normal(-.444, 42), class = "b", coef = "mmse3") +
  prior(normal(-3.219, 42), class = "b", coef = "sex2") +
  prior(normal(0, 29), class = "b", coef = "barthel_code") +
  prior(normal(-5, 42), class = "b", coef = "physres1") +
  prior(normal(0, 42), class = "b", coef = "groupintervention") +
  prior(normal(0, 13), class = "b", coef = "pas_c") +
  prior(normal(.2925, 42), class = "b", coef = "chemicalres1") +
  prior(normal(0, 26.77), class = "b", coef = "cci_c")

# see:
# Beerens HC, Sutcliffe C, Renom-Guiteras A, et al. Quality of Life and
# Quality of Care for People With Dementia Receiving Long Term Institutional
# Care or Professional Home Care: The European RightTimePlaceCare Study.
```

```
# Journal of the American Medical Directors Association. 2014;15(1):54-61.
# doi:10.1016/j.jamda.2013.09.010
#
# QoL-Scale ranges from 13-52 (40 points). Effects from those study are
# multiplied by 2.5 (rescaling 40-points to 100-point-scale of QUALIDEM)

# see:
# Barca, M. L., Engedal, K., Laks, J., & Selbæk, G. (2011). Quality of Life
# among Elderly Patients with Dementia in Institutions. Dementia and Geriatric
# Cognitive Disorders, 31(6), 435-442. https://doi.org/10.1159/000328969

# QoL-scale ranges from 11-55 (45 points). Effects from those study are
# multiplied by 2.2 (rescaling 45-points to 100-point-scale of QUALIDEM)


# model formula ----

mf <-
  formula(
    QoL ~ stay_c + age10 + mmse + sex + cci_c +
      barthel_code + physres + group + pas_c +
      chemicalres + (1 | maindiag)
  )


# brms-model ----

set.seed(1207)

m2a <- brm(
  formula = mf,
  data = d,
  prior = bprior,
  sample_prior = TRUE
)


# Figure 3 ----

theme_set(theme_sjplot2(base_size = 14, base_family = "serif"))

p <- plot_model(
  m2a,
  title = "",
  axis.labels = labs,
  sort.est = T,
  colors = c("grey30"),
  axis.title = "Change in QUALIDEM-Score",
  wrap.title = 100,
  wrap.labels = 20,
  width = .2,
  grid.breaks = 2,
  size.inner = .1
) +
  ylab("Change in QUALIDEM Total Score") +
  theme_sjplot2(base_size = 14, base_family = "serif")

p_pdf <- p +
  theme_sjplot2(base_size = 28, base_family = "serif") +
  theme(
    panel.grid.major = element_line(size = .1),
    panel.grid.minor = element_line(size = .05),
    axis.line.x      = element_line(size = .15),
    axis.line.y      = element_line(size = .15)
  )

ggsave(filename = "Fig3.tiff", plot = p, width = 170, height = 120, units = "mm", dpi = 300,
compression = "lzw")
ggsave(filename = "Fig3.pdf", scale = 2, plot = p_pdf, width = 170, height = 120, units = "mm",
dpi = 300)


# Appendix S1: Test for practical equivalence ----

rope(m2a, rope = c(-6, 6))
```

```
rope(m2a, rope = c(-7.5, 7.5))

equivalence_test(m2a, parameters = "^(?!prior)")
equivalence_test(m2a, parameters = "^(?!prior)") %>% plot()


# Appendix S1, Table Regression Coefficients ----

tab_df(tidy_stan(m2a, prob = c(.5, .89), digits = 1))


# Appendix S1, Prior Adjustement ----

insight::get_priors(m2a)


# Appendix S1, Figure distribution Posterior Samples ----

tmp <- m2a %>%
  as_tibble() %>%
  select(2:12) %>%
  gather(key = "predictor", value = "estimate") %>%
  to_factor(predictor)

tmp$predictor <- lvls_reorder(tmp$predictor, c(9, 4, 8, 3, 7, 11, 10, 6, 1, 2, 5))

p2 <- ggplot(tmp, aes(x = estimate, y = predictor)) +
  geom_vline(xintercept = 0, colour = "grey70", size = .8) +
  geom_density_ridges2(rel_min_height = 0.001, scale = 2, alpha = .5) +
  scale_x_continuous(breaks = seq(-8, 8, 2)) +
  scale_y_discrete(labels = labs) +
  labs(x = "Change in QUALIDEM-Score", y = NULL) +
  theme_sjplot2(base_size = 14, base_family = "serif")

ggsave(filename = "FigS1.tiff", plot = p2, width = 190, height = 120, units = "mm", dpi = 300)


# Appendix S1, test for practical equivalence ----

## Short version

equivalence_test(m2a, parameters = "^(?!prior)")
equivalence_test(m2a, parameters = "^(?!prior)") %>% plot()


## More beautiful tweaked version

tmp.hdi <- hdi(m2a, prob = .95) %>%
  slice(c(-c(1, 13:23)))

tmp2 <- m2a %>%
  as_tibble() %>%
  select(2:12) %>%
  map2_df(tmp.hdi$CI_low, function(x, y) {
    x[x < y] <- NA
    x
  }) %>%
  map2_df(tmp.hdi$CI_high, function(x, y) {
    x[x > y] <- NA
    x
  }) %>%
  gather(key = "predictor", value = "estimate") %>%
  to_factor(predictor)

tmp2$predictor <- lvls_reorder(tmp2$predictor, c(9, 4, 8, 3, 7, 11, 10, 6, 1, 2, 5))

tmp2$grp <- dplyr::case_when(
  tmp2$predictor %in% c("b_stay_c", "b_cci_c") ~ "reject",
  tmp2$predictor %in% c("b_age10", "b_mmse2", "b_mmse3", "b_sex2", "b_barthel_code") ~
"undecided",
  TRUE ~ "accept"
)

p3 <- ggplot(tmp2, aes(x = estimate, y = predictor, fill = grp)) +
  # rope based on "equi_test(model)".
```

```
  annotate("rect", xmin = -1.7, xmax = 1.7, ymin = 0, ymax = Inf, fill = sjplot_pal(palette =
"us")[1], alpha = 0.15) +
  geom_vline(xintercept = 0, colour = "grey70", size = .8) +
  geom_density_ridges2(rel_min_height = 0.01, scale = 2, alpha = .4) +
  scale_x_continuous(breaks = seq(-8, 8, 2)) +
  scale_y_discrete(labels = labs) +
  scale_fill_manual(values = sjplot_pal()[c(3, 1, 7)]) +
  labs(x = "Change in QUALIDEM-Score", y = NULL, fill = "Acceptance of Parameters") +
  theme_sjplot2(base_size = 14, base_family = "serif") +
  theme(
    legend.title = element_text(size = 13),
    legend.position = "bottom",
    axis.line.x = element_line(colour = "grey50"),
    axis.line.y = element_line(colour = "grey50")
  )

ggsave(filename = "FigS2.tiff", plot = p3, width = 190, height = 120, units = "mm", dpi = 300)


# Appendix S1, Posterior-Prior-Check ----

## Short version

plot_model(m2a, type = "diag", axis.lim = c(-20, 20))

## More beautiful tweaked version

pr_samp <- prior_samples(m2a) %>%
  select(starts_with("b_")) %>%
  gather(key = "Term", value = "Estimate") %>%
  mutate(Sample = "prior")

ps_samp <- posterior_samples(m2a) %>%
  select(starts_with("b_"), -b_Intercept) %>%
  gather(key = "Term", value = "Estimate") %>%
  mutate(Sample = "posterior")

m_pp_data <- bind_rows(pr_samp, ps_samp) %>% to_factor(Term)
m_pp_data$Term <- lvls_reorder(m_pp_data$Term, rev(c(8, 3, 10, 4, 7, 5, 11, 6, 1, 2, 9)))

p4 <- ggplot(m_pp_data, aes(x = Estimate, fill = Sample)) +
  geom_density(alpha = .4) +
  scale_x_continuous(limits = c(-20, 20)) +
  facet_wrap(
    ~ Term,
    scales = "free",
    labeller = labeller(Term = labs),
    nrow = 4
  ) +
  labs(x = NULL, y = NULL) +
  bayesplot::theme_default(base_size = 13) +
  theme(
    axis.line.x      = element_line(colour = "grey50"),
    axis.line.y      = element_line(colour = "grey50"),
    axis.text        = element_text(colour = "grey10"),
    axis.title       = element_text(colour = "black"),
    # strip.background = element_rect(colour = "grey50", fill = "grey90"),
    # strip.text       = element_text(colour = "grey20"),
    legend.title     = element_text(colour = "grey10"),
    legend.text      = element_text(colour = "grey20"),
    legend.position  = c(.5, .15),
    legend.justification = c(-2, 1)
  ) +
  scale_fill_manual(values = sjplot_pal("breakfast club")[c(1, 3)])

ggsave(filename = "FigS3.tiff", plot = p4, width = 190, height = 214, units = "mm", dpi = 300)


# Appendix S1, Traceplot ----

p <- rstan::traceplot(m2a$fit, pars = colnames(as.data.frame(m2a))[1:12], inc_warmup = F)
p$data$parameter <- as.character(p$data$parameter)
tmp <- p$data %>%
  filter(parameter != "b_Intercept")
```

```
for (i in 1:length(labs)) {
  if (names(labs)[i] %in% tmp$parameter) {
    r <- which(tmp$parameter == names(labs)[i])
    tmp$parameter[r] <- labs[i]
  }
}

p5 <- ggplot(tmp, aes(x = iteration, y = value, colour = chain)) +
  geom_line() +
  facet_wrap(~parameter, scales = "free_y", ncol = 3) +
  scale_color_manual(values = sjplot_pal("us", n = 4)) +
  labs(x = NULL, y = NULL) +
  bayesplot::theme_default(base_size = 13) +
  theme(
    axis.line.x      = element_line(colour = "grey50"),
    axis.line.y      = element_line(colour = "grey50"),
    axis.text        = element_text(colour = "grey10"),
    axis.title       = element_text(colour = "black"),
    # strip.background = element_rect(colour = "grey50", fill = "grey90"),
    # strip.text       = element_text(colour = "grey20"),
    legend.title     = element_text(colour = "grey10"),
    legend.text      = element_text(colour = "grey20"),
    legend.position  = c(.5, .15),
    legend.justification = c(-4.2, 0.7)
  )

ggsave(filename = "FigS4.tiff", compress = "lzw", plot = p5, width = 190, height = 214, units =
"mm", dpi = 300)
```

STROBE 2007 (v4) Statement—Checklist of items that should be included in reports of *case-control studies*

Section/Topic	Item #	Recommendation	Reported on page #
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	4-5
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	6-7
Participants	6	*(*a) Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls	6-7
		(*b*) For matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	8-10
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	8-10
Bias	9	Describe any efforts to address potential sources of bias	10
Study size	10	Explain how the study size was arrived at	7
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	10-11
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	10-11
		(*b*) Describe any methods used to examine subgroups and interactions	10-11
		(*c*) Explain how missing data were addressed	10
		(*d*) If applicable, explain how matching of cases and controls was addressed	N/A
		(*e*) Describe any sensitivity analyses	N/A
Results			

Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	12
		(b) Give reasons for non-participation at each stage	N/A
		(c) Consider use of a flow diagram	N/A
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	12-13
		(b) Indicate number of participants with missing data for each variable of interest	12-13
Outcome data	15*	Report numbers in each exposure category, or summary measures of exposure	12-13
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	13-14
		(*b*) Report category boundaries when continuous variables were categorized	N/A
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	N/A
Discussion			
Key results	18	Summarise key results with reference to study objectives	15-16
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	17-18
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	15-18
Generalisability	21	Discuss the generalisability (external validity) of the study results	18-19
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	20

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Quality of Life of Patients with Dementia in Acute Hospitals: A non-randomised case-control-study comparing a regular ward to a special care ward with dementia care concept.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-030743.R2
Article Type:	Research
Date Submitted by the Author:	26-Jul-2019
Complete List of Authors:	Lüdecke, Daniel; University Medical Center Hamburg-Eppendorf, Department of Medical Sociology Poppele, Georg; Evangelisches Krankenhaus Alsterdorf Klein, Jens; University Medical Center Hamburg-Eppendorf, Department of Medical Sociology Kofahl, Christopher; University Medical Center Hamburg-Eppendorf, Department of Medical Sociology
Primary Subject Heading:	Patient-centred medicine
Secondary Subject Heading:	Health services research, Nursing
Keywords:	Dementia < NEUROLOGY, Quality of Life, Acute Hospital, Quality of Care, INTERNAL MEDICINE

Title

Quality of Life of Patients with Dementia in Acute Hospitals: A non-randomised case-control-study comparing a regular ward to a special care ward with dementia care concept.

Authors

Daniel Lüdecke, Georg Poppele, Jens Klein, Christopher Kofahl

Corresponding author

Daniel Lüdecke
Department of Medical Sociology
University Medical Center Hamburg-Eppendorf
Martinistr. 52
20246 Hamburg
Germany
Email: d.luedecke@uke.de
ORCID iD: 0000-0002-8895-3206
Tel.: ++49 40 7410 52878

Co-authors

Georg Poppele, Evangelisches Krankenhaus Alsterdorf (Protestant Hospital Alsterdorf), Hamburg, Germany

Jens Klein, University Medical Center Hamburg-Eppendorf, Department of Medical Sociology, Hamburg, Germany

Christopher Kofahl, University Medical Center Hamburg-Eppendorf, Department of Medical Sociology, Hamburg, Germany

Keywords

Dementia, Quality of Life, Acute Hospital, Quality of Care, Internal Medicine

Word Count: 5.256 (excluding title page, abstract, strength and limitations, references, tables)

Abstract

Objectives: To identify factors that predict the quality of life (QoL) of patients with dementia in acute hospitals and to analyse if a special care concept can increase the patients' QoL.

Design: A non-randomised case-control-study, including two internal medicine wards from hospitals in Hamburg, Germany.

Setting and Participants: In all, 526 patients with dementia from two hospitals were included in the study (intervention: n=333; control: n=193). Inclusion criteria was an at least mild cognitive impairment or dementia. The intervention group was a hospital with a special care ward for internal medicine focussing on patients with dementia. The control group was from a hospital with a regular care ward without special dementia care concept.

Outcome Measures: Our main outcome was the QoL (range 0-100) from patients with dementia in two different hospitals. A Bayesian multilevel analysis was conducted to identify predictors like age, dementia, agitation, physical and chemical restraints or functional limitations that affect QoL.

Results: QoL differs significantly between the control (40.7) and intervention group (51.2), $p<0.001$. Regression analysis suggests that physical restraint (estimated effect: -4.9), psychotropic drugs use (-4.4) and agitation (-2.9) are negatively associated with QoL. After controlling for confounders, the positive effect of the special care concept remained (5.7).

Conclusions: A special care ward will improve the quality of care and has a positive impact on the QoL of patients with dementia. Health policies should consider the benefits of special care concepts and develop incentives for hospitals to improve the QoL and quality of care for these patients.

Article Summary

Strengths and Limitations

- This is one of the first studies to investigate quality of life of patients with dementia in acute internal medicine wards in Germany.
- The statistical method applied in this study explicitly incorporates and accounts for information and knowledge from previous research.
- There are no studies which have evaluated the reliability and validity for the use of the assessment instrument for our main outcome (quality of life) in hospitals settings.
- The structural differences between the hospitals from the control and intervention may limit the generalizability of the results regarding the benefit of special care concepts for regular internal medicine wards.

Introduction

Acute hospitals face the challenge of an increase in old-age patients, which particularly affects internal medicine wards [1]. The average age of patients in internal medicine wards is above 70 years and may even come close to 80 years [2,3], leading to an increasing prevalence of cognitive impairments [4]. Although data on the occurrence of dementia or cognitive impairments of patients in hospitals is inconsistent, the larger proportion of studies report prevalence rates of about 40% [5–7].

Many hospitals are insufficiently prepared for patients with cognitive impairments, especially in acute care units predominantly focussing on somatic diseases [8]. Patients with cognitive impairments or dementia do not fit into the typical routines and standardised workflows of hospitals as these patients need more resources for care and treatment [9,10]. These patients often become disoriented, anxious, agitated and challenge hospital staff with erratic behaviour when placed in regular care wards. This results in an increased likelihood of falls, complications during the hospital stay and post-operative complications [11,12]. Hence, patients with dementia are a vulnerable group with a higher risk for long-lasting functional impairments [13,14].

To control the challenging behaviour of patients with dementia, a common practice – at least in German acute hospitals, and especially for older patients – is to use physical restraints like side rails to keep patients in bed, or chemical restraints such as the on-demand-use of psychotropic medication or hypnotics and sedatives [15,16]. This practise is not limited to German hospitals; however, there seems to be large variations between countries [17]. Physical and chemical restraints as well as anxiety and challenging behaviour are associated with poorer outcomes in quality of life [QoL] and quality of care of patients with dementia [18,19]. Therefore, hospitals in general, and specifically internal medicine wards with their increasing proportion of patients with

dementia, need to address these issues in order to improve the quality of care for these patients.

At least in Germany, there were lately no care concepts that fully address the needs of patients with dementia in internal medicine [20]. The special care ward "DAVID" in the Protestant Hospital Alsterdorf in Hamburg was one of the first internal medicine wards in Germany that implemented a comprehensive care concept for patients with dementia, aiming to improve the patients' QoL during their hospital stay. QoL is an important indicator of quality of care and a major dimension when assessing patient reported outcomes, particularly in older people as global outcome measure for interventions [21,22]. The assumption of this care concept is that a special care ward for patients with dementia leads to better outcomes in QoL compared to regular internal medicine wards. A study ("DAVID 2") was conducted to investigate the impact of such a care concept. This paper shows the results of this study and addresses two research questions. First, which factors predict the QoL of patients with dementia in acute hospitals? Second, beyond these factors, can a special care concept for patients with dementia in acute hospitals increase the patients' QoL?

Methods

Study Design and Setting

The aim of this study was to compare the quality of care for patients with dementia within a specialised dementia care concept as opposed to regular care in acute hospitals. The present study was designed as a non-randomised case-control-study, including two internal medicine wards in two hospitals located in Hamburg, Germany. The intervention group was a hospital that implemented a special care ward for internal medicine focussing on patients with dementia. The control group was from a hospital

with a regular care ward for internal medicine, which had no special dementia care concept.

Intervention Group

The special care ward "DAVID" is an internal medicine ward in the Protestant Hospital Alsterdorf, a not-for-profit organization, and has 14 beds. In the year of data collection (2016), 349 patients were treated. The ward employed nine care workers as nursing staff.

Key components of the special care concept are a) a specific architectonical design, including a homelike lounge, a specific colouring of doors and walls, and a light concept with minimum 500 lux at eye level; b) doctors, nurses and service staff are trained in coping with challenging behaviour and other dementia related issues, like basal stimulation or validation therapy, but also included case conferences to discuss issues with current patients [23]; duration of training courses and case conferences was about one hour and were provided on a monthly basis by external instructors; additionally, twice per year, an internal training course was offered for employees, lasting for half a day; c) mobile devices for diagnostics, to perform as many treatments as possible in the different rooms of the special care ward; d) involvement of relatives into assessment, care and discharge planning; and e) regular therapeutic offers like occupational or speech therapy, and social offers like music, playing or spending more time than usual to care for the patients.

To fulfil these high standards of quality of care, the ward "DAVID" employs more care staff in relation to the number of patients as compared to other regular internal medicine wards in Germany. With respect to the total number of full-time equivalents [FTE] nurses, the staff-patient-ratio is one FTE nurse per 39 patients.

The Protestant Hospital Alsterdorf has a second ward for internal medicine, however, patients with dementia were usually immediately transferred to the special care ward after admission to hospital. Thus, as almost no patients with dementia were treated in the second internal medicine ward, the control group was taken from another hospital.

Control Group

The regular care ward is part of a larger private-company hospital with emergency hospitalisation. It has 80 beds and in the year of data collection, about 3.500 patients were treated in this internal medicine ward. Twenty-six employees worked as care staff in this ward. Trainees sometimes supported the care team. The staff-patient-ratio in the regular care ward is approximately one FTE nurse per 130 patients. However, since the internal medicine ward in this hospital also treats patients from the emergency ambulance, the staff-patient-ratio related to the number of patients who actually stayed longer in hospital (three days and more) is lower. Unfortunately, the hospital management was not willing to provide more detailed information beside the publicly available quality reports, so we cannot quantify the staff-patient-ratio exactly.
The regular care ward had no specific care concept for dementia patients. The care staff was not particularly trained in dementia topics.

Data collection and participants

An assessment questionnaire was developed to obtain data from patients with dementia. Study nurses were trained in using this assessment questionnaire and then conducted the data collection in both hospitals. Two study nurses were responsible for the special care ward and one for the regular care ward. A pre-test of two months was conducted to test and revise the questionnaire. As a result, some items were removed and instructions

for study nurses were defined more precisely. After the pre-test, data was collected over a period of about 12 months (from July 2015 to June 2016 in the special care ward and from August 2015 to September 2016 in the regular care ward). To detect small to medium effect sizes (Cohen's d ~ 0.1 to 0.2), a power analysis was performed prior to the data collection and yielded a sample size of at least 173 subjects per group. Patients were included when they showed at least mild cognitive impairments or memory problems. In the special care ward (intervention group) all patients were assessed because a diagnosed dementia was a requirement for admission to that hospital. Hence, the participation rate for the special care ward was about 94% and excluded only a few patients that were not responsive. For the regular care ward (control group), patients who already had a diagnosed dementia or cognitive impairments were included in the study. A short dementia screening was carried out by the study nurse to assess the severity of dementia of patients who had no clarified dementia diagnosis, and to identify further patients who qualify for the study [24]. The total sample size for the present analysis consists of N=526 patients (special care ward: n=333; regular care ward: n=193). For both the intervention and control group, patients were excluded from the study when they were completely confined to bed due to severe health-related dependency. As both care wards had no particular selection criteria for patients such as age, mobility, or the main diagnosis that lead to hospital admission, no further exclusion criteria for the study were defined.

Prior to the study, a study protocol was developed and submitted to the ethical committee of the medical association of Hamburg. The ethical committee approved the proposal and attested that the study conforms to ethical and legal requirements (approval code PV5102). Study participants were not able to give their informed consent due to their cognitive impairments. However, as data mostly derived from the

hospitals' regular documentation and was completely anonymous, the ethics committee waived the need of an informed consent.

Patient and Public involvement

Patients and the public were not involved in the development of the research question nor study design.

Measures

Outcome: QoL in patients with dementia was assessed using the QUALIDEM [25,26]. After observing patients for about one week (depending on the length of stay), the study nurses rated their QoL. QUALIDEM comprises 37 items reflecting nine different subdomains of QoL: "care relationship" (7 items, 0-21 points), "positive affect" (6 items, 0-18 points), "negative affect" (3 items, 0-9 points), "restless and tense behaviour" (3 items, 0-9 points), "positive self-image" (3 items, 0-9 points), "social relations" (6 items, 0-18 points), "social isolation" (3 items, 0-9 points), "feeling at home" (4 items, 0-12 points) and "have something to do" (2 items, 0-6 points). For patients with very severe dementia (Minimental State Examination Test [27] [MMSE] < 7), only six of the nine subscales apply, where the dimensions "positive self-image", "feeling at home" and "have something to do" were omitted. The recommendation is to report descriptive results of the QUALIDEM separately for each subscale. For regression analyses, a QoL index was calculated by summing up and normalizing the QUALIDEM subscales (six subscales for patients with very severe dementia, nine subscales for the remaining patients) to a range from 0 to 100 points. A higher score indicates better QoL. Due to normalization of the QUALIDEM total score for all severities of dementia, all patients' scores are consistent and comparable [28].

Independent Variables: Age, gender, main diagnosis for admission to hospital and length of stay were recorded. Details about the distribution of the main diagnoses among patients and by hospitals are shown in the Supplementary File 1. If a main diagnosis was mentioned no more than one time in both hospital wards, it was recoded into the category "other". The final variable "main diagnosis" comprised 20 different diagnoses. A modified version of the Charlson's Comorbidity Index [CCI], which included depression and hypertension as new items, was built based upon the assessment of comorbidities and chronical diseases [29,30]. If patients had no chronic illnesses, the CCI had a score of zero points. Else, higher scores indicated more serious comorbid disease. Shortly after admission to hospital, the study nurses measured functional limitations and cognitive status of patients. Functional limitations in daily living were assessed with the Barthel-Index [31]. This score ranges from 0 (completely dependent) to 100 points (no basic functional limitations) and was recoded according to the classification of the ICD-10 [32] (German adaption) into a score from 1 to 6 points. The Minimental State Examination Test [27] [MMSE] measures the cognitive impairments of patients, ranging from 0 (very strong cognitive impairments) to 30 (very mild or no cognitive impairments) points. This score was recoded into three categories, also based on ICD-10 classification: severe dementia (0-16), moderate dementia (17-23 points) and mild dementia (24-27 points). After about one week of hospital stay, the study nurses rated the patients' agitation and challenging behaviour and recorded psychotropic drug use (chemical restraint) and physical restraints. Agitation and challenging behaviour of patients was assessed using the Pittsburgh Agitation Scale [PAS] [33] ranging from 0 to 16 points (higher scores indicate stronger agitation).

Physical restraints were defined as the use of one the following measures: Side rails to keep a patient in bed, tying a patient to a bed, and use of "therapeutic" chairs that

prevent patients to stand up. The variable was dichotomised, indicating whether patients (in the course of the hospital stay) were mechanically restrained by at least one of these measures or not.

Psychotropic drug use was defined as on-demand-use ("as-needed") of medication for the nervous system by means of the Anatomical Therapeutical Chemical (ATC) classification [34] and comprises antipsychotics, anxiolytics, hypnotics/sedatives and antidepressants (N05A-C, N06A). Indicated use of psychotropic drugs was defined as medications that were prescribed for regular, not on-demand-use and not only given to patients in order to control their challenging behaviour. Such use of psychotropic drugs was excluded from the analysis. The on-demand-use variable was dichotomised and shows whether, during the complete hospital stay, chemical restraints were applied to patients or not.

While these variables already cover many different aspects that have an effect on the QoL, we decided to add a further predictor as proxy for the intervention to the model. Therefore, we included a binary variable with two categories ("control" as reference and "intervention") representing the two hospitals, to estimate the impact of the special care concept. This should reflect how much of the change in QoL is attributable to the special care concept.

Missing Data

In total, 11% of individual items across all scales were missing (at random), 6% of individual items when looking at the QUALIDEM only. The missing data pattern was analysed and missing data was imputed using the multivariate imputation by chained equations method [35], using 11 imputation steps corresponding to the proportion of missing data [36]. The method for imputing missing values depends on the variable's

nature. For continuous variables, predictive mean matching was applied, while logistic regressions were used for binary variables.

Statistical Methods

Descriptive results for the total sample and each hospital are reported. Statistically significant differences of $p<0.05$ between the two hospital wards were tested using t-tests, χ^2-tests or Mann-Whitney-U-tests, depending on the level of measurement and distribution of variables. Differences between the hospitals in the QUALIDEM subscales are presented as boxplots, showing the median value and upper and lower quartiles of the value distribution.

As multivariate analysis, a Bayesian linear mixed model was applied to analyse the associations between the independent variables and the outcome. Computations were based on Stan [37], a probabilistic programming language for specifying Bayesian models, using Markov Chain Monte Carlo sampling (in particular, Hamiltonian Monte Carlo) [38]. We assume that the patients' main diagnosis is associated with different degrees of physical impairments, which affect the QoL. Therefore, the variable 'main diagnosis' was used as level-2 unit (*random intercept*) in the multilevel model to control for the variation in the outcome. We used informative priors for the predictors age, female gender, severe dementia, psychotic drug use and physical restraints, based on information from former research [18,39,40]. Weakly informative priors were used for the remaining predictors. The prior and posterior distributions of the model are summarised in the supplemental material (see Supplementary File 2).

Continuous predictors were centred before entering the model. Age was divided by 10, so a 1-unit change in the predictor of age reflects a change of 10 years in patients. The median value of the posterior distribution is used as "Bayesian point estimate", which

minimises the difference of estimates from true values over posterior samples, but there are many other plausible values (the "posterior distribution") to describe the association between predictors and outcome. Hence, 50% and 89% highest density intervals [41] [HDI] are shown to indicate the range of most credible values and to reflect the (un-)certainty of the estimates. The intraclass correlation coefficient [42] was calculated to see how much of the proportion of the variance in the outcome can be explained by the grouping structure ('main diagnosis'). We developed post-hoc additional regression models with interaction terms for need predictors (Barthel-Index, physical and chemical restraints, PAS-Score) to check if the associations between the complexity of patients' needs and QoL differ between hospitals. We found no significant interaction terms and decided to present the most parsimonious model here and show further results in the appendix (see Supplementary File 3).

All analyses were conducted with the R statistical package [43], including the packages *mice* [35], *ggplot* [44], *brms* [45] and *sjPlot* [46]. The source code is available in the supplemental material (see Supplementary File 4). Data is available online [47].

Results

Sample Characteristics

Table 1 gives an overview of the sample characteristics. The proportion of female to male patients is similar in both groups. The mean age is 4 years higher in the control group. There are also significant group differences in the Barthel-Index indicating higher functional impairment in the control group, while the dementia severity was the same in both hospitals. Comorbid conditions are slightly higher in the control group. Patients stayed 9.4 days in hospital on average and nearly one day longer in the intervention group as compared to the control group. Large differences between the two hospitals

can be seen in the use of medical and physical restraints with significantly less use in the intervention group. Agitation- and QoL-scores also show strong group differences to the disadvantage of the patients in the control group.

In most cases, the distribution of main diagnoses of patients were comparable between the two hospital wards (see Supplemental File 1). Most frequent were pneumonia (13.5% in the intervention group and 11.9% in the control group), a worsening medical condition of patients (8.7% and 7.2%) or exsiccosis (4.8% and 6.7%). Noticeable differences between the two wards were found in urinary tract infections (UTI) (9.9% in the intervention group and 3.1% in the control group) or dyspnoea (1.2% and 7.8%).

Table 1: Sample Characteristics

Characteristic	**Control Group (Regular Care Ward, n=193)**	**Intervention Group (Special Care Ward, n=333)**	**Total (N=526)**	**p-value of difference**
Female, %	59.1	61.6	60.6	.637
Mean Age (SD)	83.1 (7.2)	79.0 (11.9)	80.5 (10.6)	**<.001**
Mean Barthel-Index (SD)	29.9 (27.9)	40.7 (30.4)	36.7 (29.9)	**<.001**
Mild Dementia, %	7.8	9.6	8.9	.580
Moderate Dementia, %	30.0	26.7	27.9	.473
Severe Dementia, %	62.2	63.7	63.2	.805
Mean Length of Stay, in Days (SD)	8.9 (7.5)	9.7 (5.5)	9.4 (6.3)	**.002**
Physical Restraints (yes), %	54.4	28.2	37.8	**<.001**
Psychotropic Drug Use (yes, as-needed), %	25.9	14.1	18.4	**.001**
Mean-Score Pittsburg Agitation Scale (SD)	3.9 (3.1)	3.0 (3.2)	3.3 (3.2)	**<.001**
Mean Charlson' Comorbitiy Index (SD)	3.2 (3.0)	2.5 (2.0)	2.8 (1.6)	**<.001**
Mean QUALIDEM Total Score (SD)	40.7 (14.5)	51.2 (17.2)	47.3 (17.0)	**<.001**

Barthel-Index: 0-100 (higher = better functioning); Dementia (MMSE score): mild: 24-27; moderate: 17-23; severe: <= 16; Pittsburg Agitation Scale: 0-16 (higher = stronger agitation and anxiety); Charlson' Comorbitiy Index: 0-9 (higher = more comorbidities); QUALIDEM: 0-100 (higher = better QoL)

Quality of life

Looking at the QoL for patients with severe to mild cognitive impairments (these are the ones, in which all subdomains of the QUALIDEM could be applied), there is a consistent pattern across all QUALIDEM-domains: Patients in the control group have a lower QoL compared to the intervention group. Except for the last subdomain ('having something to do'), all differences are statistically significant (Figure 1).

The same consistent pattern can be found for patients with very severe dementia symptoms (MMSE score < 7). Here, only the second of the six applied subdomains ('positive affect') does not differ significantly between intervention and control (Figure 2).

Predictors of quality of life

Figure 3 shows the results from the Bayesian mixed model. Three predictors are clearly negatively associated with QoL: physical restraint, psychotropic drug use and agitation (PAS-score). Physical restraint is associated with a 4.9-point decrease in QoL. With 50% probability, the QoL decreases by 4.1 to 5.8 points and with a chance of 89% by -7 to -2.8 points respectively. The application of psychotropic drugs as-needed shows similar results, with a posterior median of -4.4. The third clearly negative associated predictor is agitation, which shows a decrease in QoL of about 2.9 points for each additional point in the PAS-score.

Dementia and gender are not clearly associated with QoL. Neither are the length of hospital stay and the CCI.

The age of the patient correlates slightly positive with QoL, where an increase of 10 years means an increase of about 1.2 points in the QoL. The posterior median of the Barthel-Index is 2.0, so for a one-category change in functional impairments the QoL

changes by two points. This means that patients with severe functional impairments differ by about 10 points in QoL compared to patients with no functional impairments. Controlling for all other predictors, the intervention (special care ward) shows the strongest association with our outcome of interest, the patients' QoL. The posterior median is 5.7, and with an 89% probability, the credible values describing the effect of the intervention on QoL are within the range from 3.8 to 7.6.

The intraclass correlation coefficient of the model is rather low (0.01). This means, the 'main diagnosis' does not explain much of the variance in the patients' QoL and there is almost no regularization ("shrinkage") of estimated model parameters and no larger differences between hospitals according to the patients' needs, as indicated by their main diagnosis.

Discussion

The study reported in this paper sought to understand those factors that influence the QoL in patients with dementia and whether a special care concept for these patients performs better in this regard as opposed to regular care wards.

One of our main findings is that QoL differs significantly between the control and intervention group. We found substantial differences between the two hospitals in the patients' total QoL score in favour of the special care ward. Beyond the statistical significance, this finding also has a clinical impact. Studies suggest a change in 3 points for the Quality of Life – Alzheimer's Disease Scale [48], which has a range of 40 points, to be clinically relevant [49,50]. Transferred to the range of the QUALIDEM scale, a difference of about 7.5 points would be considered as an important improvement in QoL. Another indicator to evaluate the clinical relevance of a change in QoL is an increase of the score of half a standard deviation [51], which would be about 8.5 points

for our data. Taking these reference points as a basis, we found evidence for the clinical relevant improvement in QoL of patients in a special care ward.

A second key finding is the identification of those factors that are clearly associated with QoL. The use of physical and chemical restraints, both happening more frequently in the control group, are associated with lower outcomes in QoL. This finding is in line with other studies that suggest a negative association between physical and chemical restraints and QoL [18,40] and explains why the regular care ward performs less good in this regard than the special care ward. Agitation was also negatively associated with QoL. This is understandable as agitation is an expression of anxiety and indisposition of people with dementia and typically occurs after admission to hospital. Furthermore, agitation is often a reason for psychotropic drug use or physical restraint and, thus, also negatively affects QoL [52,53].

Independent from these factors, the special care ward itself shows the strongest impact on QoL, indicating that patients with dementia explicitly benefit from specialised care concepts. Other studies also report these benefits, both in a nursing home or hospital setting [54,55]. Since we controlled for patient characteristics like main diagnosis, age, functional limitations, chronic comorbidities, agitation, length of stay etc. in our model, we do not assume that the positive effect of the special care ward is completely a result of a biased sample between intervention and control group. Although the two compared hospitals differ in their structures and size, patients' characteristics are largely comparable between the samples in the control and intervention group. For instance, there is no substantial difference between the two hospitals regarding the relationship between functional impairments and physical restraints. Moreover, to see if the complexity of patients' need affects our findings, we calculated regression models with

interaction terms between need factors moderated by hospitals (see Supplementary File 3).

The association between complexity of needs and QoL is not significantly different between the intervention and control group. Based on our results we suggest that the special care concept mainly explains the differences in the QoL. Although it is certainly difficult to determine the exact effect of the special care concept on the patients' QoL, our findings seem plausible in the light of the key elements of this intervention. A higher ratio of care staff as to patients, smaller facilities or systematically trained employees can be considered essential for health care provision to patients with dementia and are much better conditions for less physical or chemical restraints, independent of the functional limitations of patients. The special care ward provides a more dementia-friendly interior design, including orientation and navigation aids and the use of light and colours, which are considered as important components to reduce agitation for patients with dementia [56].

These findings and conclusions are in line with other studies on hospital care that suggest that an increased staff ratio or the implementation of multiple components, which particularly address the needs of patients with dementia, lead to reduced use of physical restraints and psychotropic drug use and improve the quality of care [57,58]. Furthermore, dementia-specific educational programmes, as implemented in the special care ward, have positive effects on nurses regarding their interaction with patients with dementia. Trained nurses can improve their coping skills in handling challenging behaviour of these patients, and better attend to the patients' unmet physical and psychological needs [59].

Studies suggest that the use of both physical and chemical restraints is reduced for nurses who completed a dementia-specific training as opposed to nurses who did not

complete such an educational programme. Trained nurses had better skills in providing patient-centred care and thus improving the QoL for patients with dementia [59–61]. The special care ward benefits from a higher staff ratio, i.e. nurses have to care for fewer patients with dementia compared to the control group. While this is an intentional element of the concept, the downside is higher personnel costs. Only few studies investigated the follow-up costs for patients with dementia in home care settings after hospitalization. Costa et al. predicted additional monthly costs in home care of about 445 Euros due to increased agitation of patients with dementia [62]. Thus, if patients with dementia benefit from special care concepts and perceive better outcomes in quality of life and care, the increased costs for more care personnel may be compensated by reducing follow-up costs for the ambulatory care. However, further research is needed to give more exact projections of the increased costs and potential of saving money.

Another finding is that the severity of cognitive impairments, measured with the MMSE, is a rather improper indicator to represent the underlying problems of and with the dementia disease, as these factors were not consistently associated with QoL. Direct measures of the problems associated with dementia, as agitation or challenging behaviour, should be considered as well when it comes to investigate the QoL of patients with dementia.

Our study has several limitations. One concerns the structural differences between the two hospitals. The hospital with the special care ward is much smaller than the hospital that hosted the control group. A second control group or an intervention group in a hospital of a similar size as the hospital with the regular care ward may have permitted a more distinct comparison. We tried to keep the impact of the structural differences as

minimal as possible, for instance by accounting for many different patient characteristics including functional status, comorbidities and behavioural problems. Furthermore, the main diagnoses of patients were also considered in the analysis. We assume that we could at least partly adjust our analysis for a bias due to patient selection mechanisms. To validate our assumptions, we investigated to which extent the association between patient characteristics and QoL is affected by differences between the control and intervention group (details shown in Supplementary File 3). Results suggest that our data provides no strong evidence for noticeably differences between the intervention and control group regarding the association between complexity of patients' needs and QoL. However, although we adjusted our analysis for many patient characteristics, we cannot eliminate a potential bias due to different hospital structures. In particular, the higher mean age and stronger functional limitations in the control group may indicate a selection bias in our sample. We suggest that further studies should take a second control group or a more comparable intervention group into account to gain more insight into potential biases due to structural differences of the control and intervention group.

Another structural difference between the intervention and control group that certainly affects the results are the different staff-patient-ratios. In the special care ward, nurses have to care for fewer patients than in the regular care ward. Although we assume that this aspect probably has the highest impact on the outcomes in QoL, this is not a "selection bias" per se rather than a core component of the intervention. A higher staff-patient-ratio, dementia-specific training programmes, or a specific architectonical design are key elements of the special care concept, which, in their entirety, are reflected in the resulting differences between hospitals.

A further limitation is possibly the first and thus rather exploratory use of the QUALIDEM assessment in a hospital setting. Although studies show reliable results of the QUALIDEM in nursing homes even for a short observation period of about one week [63], there are no studies that evaluate the reliability and validity for use in hospitals. We have done checks of internal consistencies, which showed that most subdomains of the QUALIDEM perform well with our data and are comparable to results from other validation studies [64]. This indicates that the use of the QUALIDEM is feasible for hospital research. However, due to financial and logistic limitations, it was not possible to monitor the complete data collection and accurate completion of questionnaires. Hence, we cannot give evidence on the interrater reliability apart from the intense training of the study nurses.

Another debatable issue regarding the QUALIDEM concerns the computation concept of the total score for patients with very severe dementia. We followed the QUALIDEM authors' instruction to use only six of the nine subscales to calculate the total score for this group [65]. Technically, this is similar to mean value imputation for the missing scores of the three omitted subscales. This, however, may result in biased and/or underestimated measurement error variance for this group. Therefore, we also calculated a regression model with a QUALIDEM total score based on imputation for missing values for all nine subscales for patients with very severe dementia (see Supplementary File 5). In the results section, we have provided the analyses as suggested by the QUALIDEM authors for comparability reasons. In order to meet different views on the computation concept, we also provide the results of the alternative analysis in the Supplementary File 5. These are very similar to the first analysis and do not differ significantly.

Finally, due to the nature of the study design, it was not possible that study nurses in the intervention and control group were blinded. This might affect the results insofar as study nurses may have generated more generous responses for the assessment scales [66].

Conclusions

On the whole, we think that a special care ward will improve the quality of care and is effective regarding the positive impact on the QoL of patients with dementia. Our study showed that after controlling for different predictors, the intervention still has a perceptible effect concerning clinical important differences in our outcome of interest, the patients' quality of life. However, such improvements can only be achieved by implementing a concept with multiple components that address the explicit needs of patients with dementia. The implementation of a special care concept usually increases the costs for hospitals because it requires a higher staff-patient-ratio, regular training of employees or more therapeutic offers. On the other hand, costs that accumulate in informal care after hospital stay as a result of poorer quality of care in hospitals can be much higher than additional personnel costs and could probably be reduced [62,67]. Health policies should consider the benefits of special care concepts and develop incentives for hospitals to improve the QoL and quality of care for patients with dementia.

Acknowledgement

The authors are grateful to the clinical staff of both hospitals who supported this study by helping with the data collection. Furthermore, we are thankful to the Authority of Health and Consumer Protection and the Homann-foundation for funding this study.

Author Contributions

DL was involved in conception and design of the study, analysis and interpretation of data and drafting the article. GP was involved in conception and design of the study and interpretation of data. JK was involved in interpretation of data and drafting the article. CK was involved in conception and design of the study, interpretation of data and drafting the article.

Competing interests

The authors have declared that no competing interests exist.

Funding

The present study was funded by the Behörde für Gesundheit und Verbraucherschutz (Authority of Health and Consumer Protection) in Hamburg and the Homann-foundation (grant number Z/65454/2014/F500-02.10/10,002). No one of the funders was involved in the data analysis process and did not influence or hold back results of the study.

Ethics approval and consent to participate

Prior to the study, a study protocol was developed and submitted to the ethical committee of the medical association of Hamburg ("Ethik-Kommission der Ärztekammer Hamburg"). The IRB of the ethical committee approved the proposal and attested that the study conforms to ethical and legal requirements (approval code PV5102). Study participants were not able to give their informed consent due to their cognitive impairments. However, as data mostly derived from the hospitals' regular documentation and was completely anonymous, the IRB of the ethics committee waived the need of an informed consent.

Availability of data and material

All data generated or analysed during this study are licensed under CC BY-NC 4.0 and available in the Zenodo repository (DOI: 10.5281/zenodo.3351450) at https://doi.org/10.5281/zenodo.3351450.

References

1 Recine U, Scotti E, Bruzzese V, *et al.* The change of hospital internal medicine: a study on patients admitted in internal medicine wards of 8 hospitals of the Lazio area, Italy. *Ital J Med* 2015;**9**:252. doi:10.4081/itjm.2015.523

2 Raveh D, Gratch L, Yinnon AM, *et al.* Demographic and clinical characteristics of patients admitted to medical departments. *J Eval Clin Pract* 2005;**11**:33–44. doi:10.1111/j.1365-2753.2004.00492.x

3 Sonnenblick M, Raveh D, Gratch L, *et al.* Clinical and demographic characteristics of elderly patients hospitalised in an internal medicine department in Israel: Characteristics of elderly patients in medical wards. *Int J Clin Pract* 2007;**61**:247–54. doi:10.1111/j.1742-1241.2006.00925.x

4 Mukadam N, Sampson EL. A systematic review of the prevalence, associations and outcomes of dementia in older general hospital inpatients. *Int Psychogeriatr* 2011;**23**:344–55. doi:10.1017/S1041610210001717

5 Zekry D, Herrmann FR, Grandjean R, *et al.* Does dementia predict adverse hospitalization outcomes? A prospective study in aged inpatients. *Int J Geriatr Psychiatry* 2009;**24**:283–91. doi:10.1002/gps.2104

6 Lyketsos CG, Sheppard JM, Rabins PV. Dementia in elderly persons in a general hospital. *Am J Psychiatry* 2000;**157**:704–7. doi:10.1176/appi.ajp.157.5.704

7 Sampson EL, Blanchard MR, Jones L, *et al.* Dementia in the acute hospital: prospective cohort study of prevalence and mortality. *Br J Psychiatry J Ment Sci* 2009;**195**:61–6. doi:10.1192/bjp.bp.108.055335

8 Pinkert C, Holle B. Menschen mit Demenz im Akutkrankenhaus: Literaturübersicht zu Prävalenz und Einweisungsgründen. *Z Für Gerontol Geriatr* 2012;**45**:728–34. doi:10.1007/s00391-012-0319-1

9 Dutzi I, Schwenk M, Micol W, *et al.* Patienten mit Begleitdiagnose Demenz. *Z Für Gerontol Geriatr* 2013;**46**:208–13. doi:10.1007/s00391-013-0483-y

10 Timmons S, O'Shea E, O'Neill D, *et al.* Acute hospital dementia care: results from a national audit. *BMC Geriatr* 2016;**16**:113. doi:10.1186/s12877-016-0293-3

11 Pi H-Y, Gao Y, Wang J, *et al.* Risk Factors for In-Hospital Complications of Fall-Related Fractures among Older Chinese: A Retrospective Study. *BioMed Res Int* 2016;**2016**:1–11. doi:10.1155/2016/8612143

12 Hu C-J, Liao C-C, Chang C-C, *et al.* Postoperative Adverse Outcomes in Surgical Patients with Dementia: A Retrospective Cohort Study. *World J Surg* 2012;**36**:2051–8. doi:10.1007/s00268-012-1609-x

13 Motzek T, Junge M, Marquardt G. Einfluss der Demenz auf Verweildauer und Erlöse im Akutkrankenhaus. *Z Für Gerontol Geriatr* 2017;**50**:59–66. doi:10.1007/s00391-016-1040-2

14 Gross AL, Jones RN, Habtemariam DA, *et al.* Delirium and Long-term Cognitive Trajectory Among Persons With Dementia. *Arch Intern Med* 2012;**172**:1324. doi:10.1001/archinternmed.2012.3203

15 Nakanishi M, Okumura Y, Ogawa A. Physical restraint to patients with dementia in acute physical care settings: effect of the financial incentive to acute care hospitals. *Int Psychogeriatr* 2017;:1–10. doi:10.1017/S104161021700240X

16 Krüger C, Mayer H, Haastert B, *et al.* Use of physical restraints in acute hospitals in Germany: a multi-centre cross-sectional study. *Int J Nurs Stud* 2013;**50**:1599–606. doi:10.1016/j.ijnurstu.2013.05.005

17 Heinze C, Dassen T, Grittner U. Use of physical restraints in nursing homes and hospitals and related factors: a cross-sectional study: Physical restraints in nursing homes and hospitals. *J Clin Nurs* 2012;**21**:1033–40. doi:10.1111/j.1365-2702.2011.03931.x

18 Beerens HC, Sutcliffe C, Renom-Guiteras A, *et al.* Quality of Life and Quality of Care for People With Dementia Receiving Long Term Institutional Care or Professional Home Care: The European RightTimePlaceCare Study. *J Am Med Dir Assoc* 2014;**15**:54–61. doi:10.1016/j.jamda.2013.09.010

19 Beerens HC, Zwakhalen SMG, Verbeek H, *et al.* Factors associated with quality of life of people with dementia in long-term care facilities: A systematic review. *Int J Nurs Stud* 2013;**50**:1259–70. doi:10.1016/j.ijnurstu.2013.02.005

20 Kricheldorff C, Hofmann W. Demenz im Akutkrankenhaus. *Z Für Gerontol Geriatr* 2013;**46**:196–7. doi:10.1007/s00391-013-0484-x

21 Treurniet HF, Essink-Bot M-L, Mackenbach JP, *et al.* Health-related quality of life: an indicator of quality of care? *Qual Life Res* 1997;**6**:363–9. doi:10.1023/A:1018435427116

22 Valderas JM, Alonso J. Patient reported outcome measures: a model-based classification system for research and clinical practice. *Qual Life Res* 2008;**17**:1125–35. doi:10.1007/s11136-008-9396-4

23 Tondi L, Ribani L, Bottazzi M, *et al.* Validation therapy (VT) in nursing home: A case-control study. *Arch Gerontol Geriatr* 2007;**44**:407–11. doi:10.1016/j.archger.2007.01.057

24 Kaiser AK, Hitzl W, Iglseder B. Three-question dementia screening: Development of the Salzburg Dementia Test Prediction (SDTP). *Z Für Gerontol Geriatr* 2014;**47**:577–82. doi:10.1007/s00391-013-0568-7

25 Ettema TP, Dröes R-M, de Lange J, *et al.* QUALIDEM: development and evaluation of a dementia specific quality of life instrument. Scalability, reliability and internal structure. *Int J Geriatr Psychiatry* 2007;**22**:549–56. doi:10.1002/gps.1713

26 Dichter MM, Ettema TP, Schwab CGG, *et al.* Benutzerhandbuch für die deutschsprachige QUALIDEM Version 2.0. Witten: 2016.

27 Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;**12**:189–98.

28 Dichter MN, Meyer G. Quality of Life of People with Dementia in Nursing Homes. In: Schüssler S, Lohrmann C, eds. *Dementia in Nursing Homes*. Cham: : Springer 2017. 139–57. doi:10.1007/978-3-319-49832-4_11

29 Charlson ME, Pompei P, Ales KL, *et al.* A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;**40**:373–83.

30 Charlson ME, Charlson RE, Peterson JC, *et al.* The Charlson comorbidity index is adapted to predict costs of chronic disease in primary care patients. *J Clin Epidemiol* 2008;**61**:1234–40. doi:10.1016/j.jclinepi.2008.01.006

31 Mahoney FI, Barthel DW. Functional Evaluation: The Barthel Index. *Md State Med J* 1965;**14**:61–5.

32 Deutsches Institut f. medizinische Dokumentation u. Information (DIMDI). *ICD-10-GM 2017 Systematisches Verzeichnis*. Lich, Hess: : pictura Werbung 2016.

33 Rosen J, Burgio L, Kollar M, *et al.* The Pittsburg Agitation Scale: A user-friendly instrument for rating agitation in dementia patients. *Am J Geriatr Psychiatry Off J Am Assoc Geriatr Psychiatry* 1994;**2**:52–9.

34 WHO Collaborating Centre for Drug Statistics Methodology, editor. *Guidelines for ATC classification and DDD assignment 2013*. Oslo: 2012.

35 Buuren S van, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. *J Stat Softw* 2011;**45**. doi:10.18637/jss.v045.i03

36 White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med* 2011;**30**:377–99. doi:10.1002/sim.4067

37 Carpenter B, Gelman A, Hoffman MD, *et al.* Stan: A Probabilistic Programming Language. *J Stat Softw* 2017;**76**. doi:10.18637/jss.v076.i01

38 Gelman A, Carlin JB, Stern HS, *et al. Bayesian data analysis*. Third edition. Boca Raton: : CRC Press 2014.

39 Barca ML, Engedal K, Laks J, *et al.* Quality of Life among Elderly Patients with Dementia in Institutions. *Dement Geriatr Cogn Disord* 2011;**31**:435–42. doi:10.1159/000328969

40 Beer C, Flicker L, Horner B, *et al.* Factors Associated with Self and Informant Ratings of the Quality of Life of People with Dementia Living in Care Facilities: A Cross Sectional Study. *PLoS ONE* 2010;**5**:e15621. doi:10.1371/journal.pone.0015621

41 McElreath R. *Statistical rethinking: a Bayesian course with examples in R and Stan*. Boca Raton London New York: : CRC Press, Taylor & Francis Group 2016.

42 Hox JJ. *Multilevel analysis: techniques and applications*. 2nd ed. New York: : Routledge 2010.

43 R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: : R Foundation for Statistical Computing 2019. https://www.R-project.org/

44 Wickham H. *ggplot2: Elegant Graphics for Data Analysis.* 2nd ed. New York: : Springer 2016.

45 Bürkner P-C. brms: An R Package for Bayesian Multilevel Models Using Stan. *J Stat Softw* 2017;**80**. doi:10.18637/jss.v080.i01

46 Lüdecke D. *sjPlot: Data Visualization for Statistics in Social Science.* 2018. doi:10.5281/zenodo.1308157

47 Lüdecke D. Dataset: Quality Of Life Of Patients With Dementia In Acute Hospitals-Version 1.0. 2018. doi:10.5281/zenodo.1479677

48 Thorgrimsen L, Selwood A, Spector A, *et al.* Whose quality of life is it anyway? The validity and reliability of the Quality of Life-Alzheimer's Disease (QoL-AD) scale. *Alzheimer Dis Assoc Disord* 2003;**17**:201–8.

49 Selwood A, Thorgrimsen L, Orrell M. Quality of life in dementia--a one-year follow-up study. *Int J Geriatr Psychiatry* 2005;**20**:232–7. doi:10.1002/gps.1271

50 Hoe J, Hancock G, Livingston G, *et al.* Changes in the quality of life of people with dementia living in care homes. *Alzheimer Dis Assoc Disord* 2009;**23**:285–90. doi:10.1097/WAD.0b013e318194fc1e

51 Norman GR, Sloan JA, Wyrwich KW. Interpretation of Changes in Health-related Quality of Life: The Remarkable Universality of Half a Standard Deviation. *Med Care* 2003;**41**:582–92. doi:10.1097/01.MLR.0000062554.74615.4C

52 Lam K, Kwan JSK, Wai Kwan C, *et al.* Factors Associated With the Trend of Physical and Chemical Restraint Use Among Long-Term Care Facility Residents in Hong Kong: Data From an 11-Year Observational Study. *J Am Med Dir Assoc* 2017;**18**:1043–8. doi:10.1016/j.jamda.2017.06.018

53 Ó Flatharta T, Haugh J, Robinson SM, *et al.* Prevalence and predictors of bedrail use in an acute hospital. *Age Ageing* 2014;**43**:801–5. doi:10.1093/ageing/afu081

54 Kok JS, Berg IJ, Scherder EJA. Special Care Units and Traditional Care in Dementia: Relationship with Behavior, Cognition, Functional Status and Quality of Life - A Review. *Dement Geriatr Cogn Disord Extra* 2013;**3**:360–75. doi:10.1159/000353441

55 von Renteln-Kruse W, Neumann L, Klugmann B, *et al.* Geriatric Patients With Cognitive Impairment. *Dtsch Ärztebl Int* 2015;**112**:103–12. doi:10.3238/arztebl.2015.0103

56 Büter K, Motzek T, Dietz B, *et al.* Demenzsensible Krankenhausstationen: Expertenempfehlungen zu Planung und Gestaltung. *Z Für Gerontol Geriatr* 2017;**50**:67–72. doi:10.1007/s00391-016-1079-0

57 Blair A, Anderson K, Bateman C. The "Golden Angels": effects of trained volunteers on specialling and readmission rates for people with dementia and delirium in rural hospitals. *Int Psychogeriatr* 2018;:1–10. doi:10.1017/S1041610218000911

58 Sinvani L, Warner-Cohen J, Strunk A, *et al.* A Multicomponent Model to Improve Hospital Care of Older Adults with Cognitive Impairment: A Propensity Score-Matched

Analysis. *J Am Geriatr Soc* Published Online First: 11 August 2018. doi:10.1111/jgs.15452

59 Schindel Martin L, Gillies L, Coker E, *et al.* An Education Intervention to Enhance Staff Self-Efficacy to Provide Dementia Care in an Acute Care Hospital in Canada: A Nonrandomized Controlled Study. *Am J Alzheimers Dis Dementiasr* 2016;**31**:664–77. doi:10.1177/1533317516668574

60 Matsuda Y, Hashimoto R, Takemoto S, *et al.* Educational benefits for nurses and nursing students of the dementia supporter training program in Japan. *PLOS ONE* 2018;**13**:e0200586. doi:10.1371/journal.pone.0200586

61 Galvin JE, Kuntemeier B, Al-Hammadi N, *et al.* "Dementia-friendly Hospitals: Care not Crisis." *Alzheimer Dis Assoc Disord* 2010;**24**:372–9. doi:10.1097/WAD.0b013e3181e9f829

62 Costa N, Wübker A, De Mauléon A, *et al.* Costs of Care of Agitation Associated With Dementia in 8 European Countries: Results From the RightTimePlaceCare Study. *J Am Med Dir Assoc* 2018;**19**:95.e1-95.e10. doi:10.1016/j.jamda.2017.10.013

63 Dichter MN, Schwab CGG, Meyer G, *et al.* Item distribution, internal consistency and inter-rater reliability of the German version of the QUALIDEM for people with mild to severe and very severe dementia. *BMC Geriatr* 2016;**16**. doi:10.1186/s12877-016-0296-0

64 Dichter M, Dortmann O, Halek M, *et al.* Scalability and internal consistency of the German version of the dementia-specific quality of life instrument QUALIDEM in nursing homes – a secondary data analysis. *Health Qual Life Outcomes* 2013;**11**:91. doi:10.1186/1477-7525-11-91

65 Dichter MN, Quasdorf T, Schwab CGG, *et al.* Dementia care mapping: effects on residents' quality of life and challenging behavior in German nursing homes. A quasi-experimental trial. *Int Psychogeriatr* 2015;**27**:1875–92. doi:10.1017/S1041610215000927

66 Schulz KF, Grimes DA. Blinding in randomised trials: hiding who got what. *The Lancet* 2002;**359**:696–700. doi:10.1016/S0140-6736(02)07816-9

67 Leon J, Cheng CK, Neumann PJ. Alzheimer's disease care: costs and potential savings. *Health Aff Proj Hope* 1998;**17**:206–16.

Figure Titles and Legends

Figure 1: QUALIDEM subdomain scores by care ward in patients with mild to severe dementia (MMSE score from 7 to 27, n = 400)

Figure 2: QUALIDEM subdomain scores by care ward, patients with very severe dementia (MMSE score < 7, n = 126)

Figure 3: Predictors of Health Related Quality of Life, Regression Coefficients, Bayesian Linear Mixed Model, Posterior Median (+50% and 89% High Density Interval)

Supplementary Files

Supplementary File 1: Figure, Distribution of Patients' Main Diagnosis by Hospital

Supplementary File 2: Methodological comments

Supplementary File 3: Regression Models with Interaction Terms

Supplementary File 4: R Source Code (to use with R statistics, CC BY-NC 4.0 license)

Supplementary File 5: Regression Model with Alternative QUALIDEM-Score

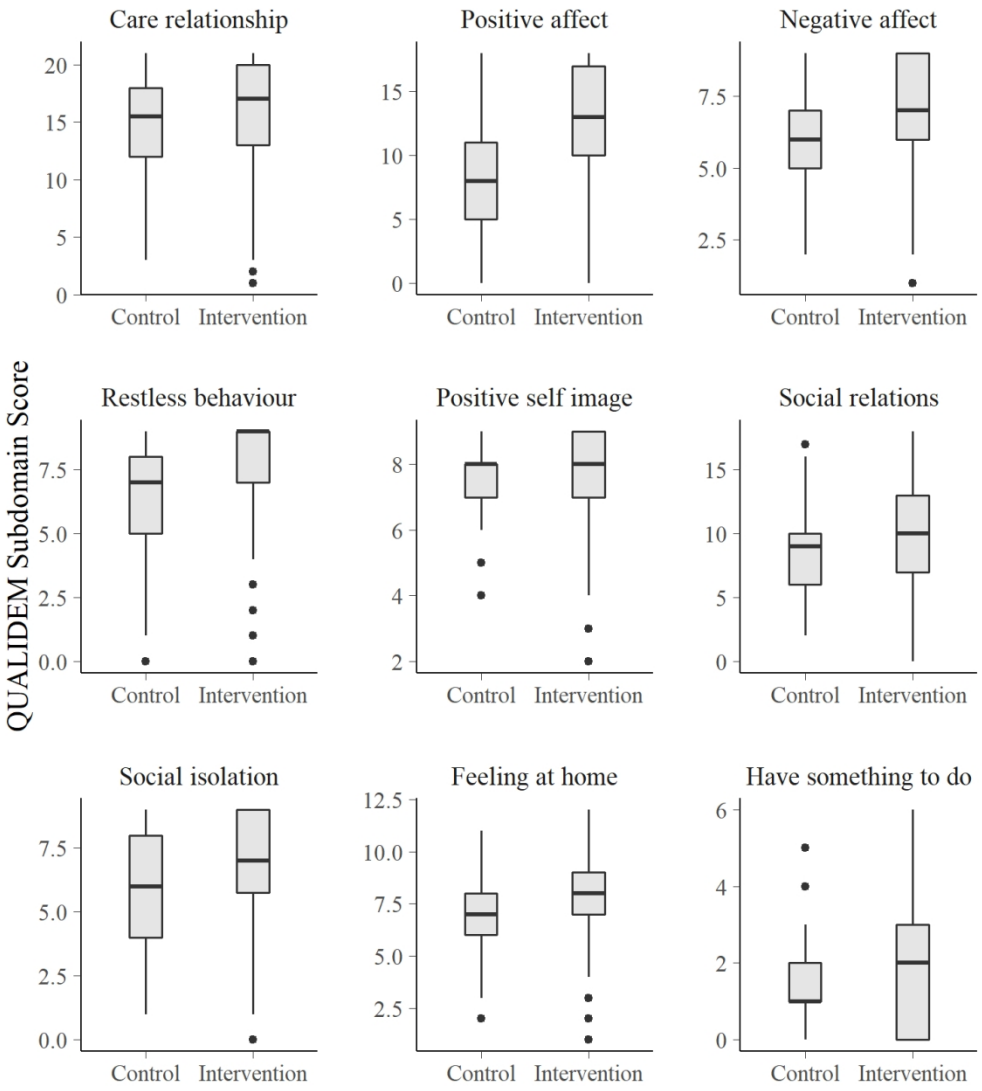


Figure 1: QUALIDEM subdomain scores by care ward in patients with mild to severe dementia (MMSE score from 7 to 27, n = 400)

169x189mm (220 x 220 DPI)

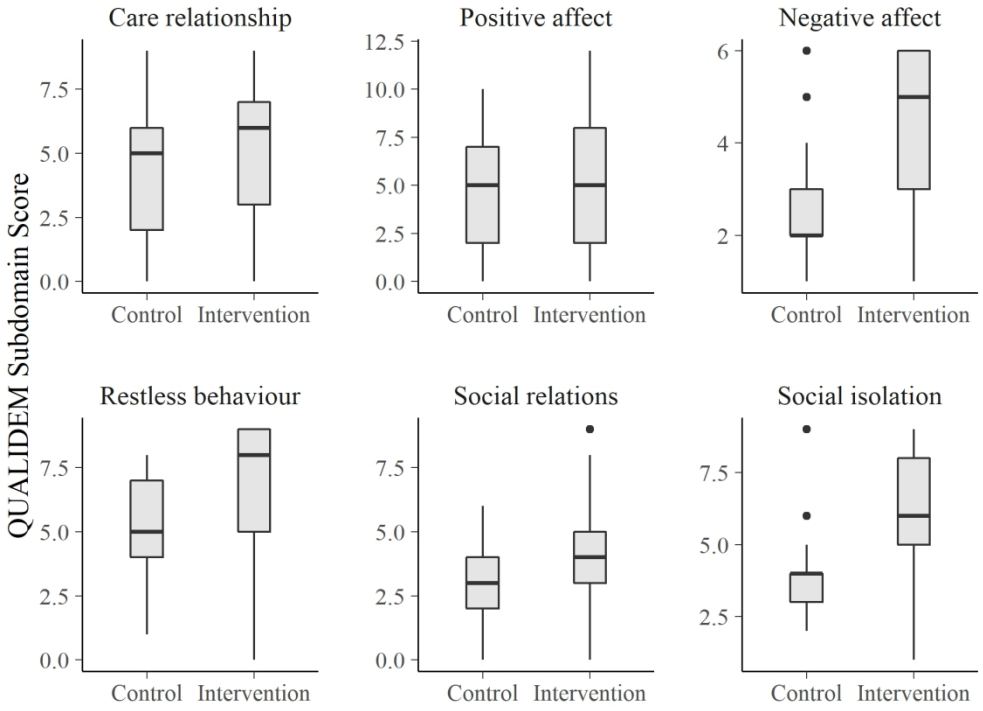


Figure 2: QUALIDEM subdomain scores by care ward, patients with very severe dementia (MMSE score < 7, n = 126)

169x125mm (300 x 300 DPI)

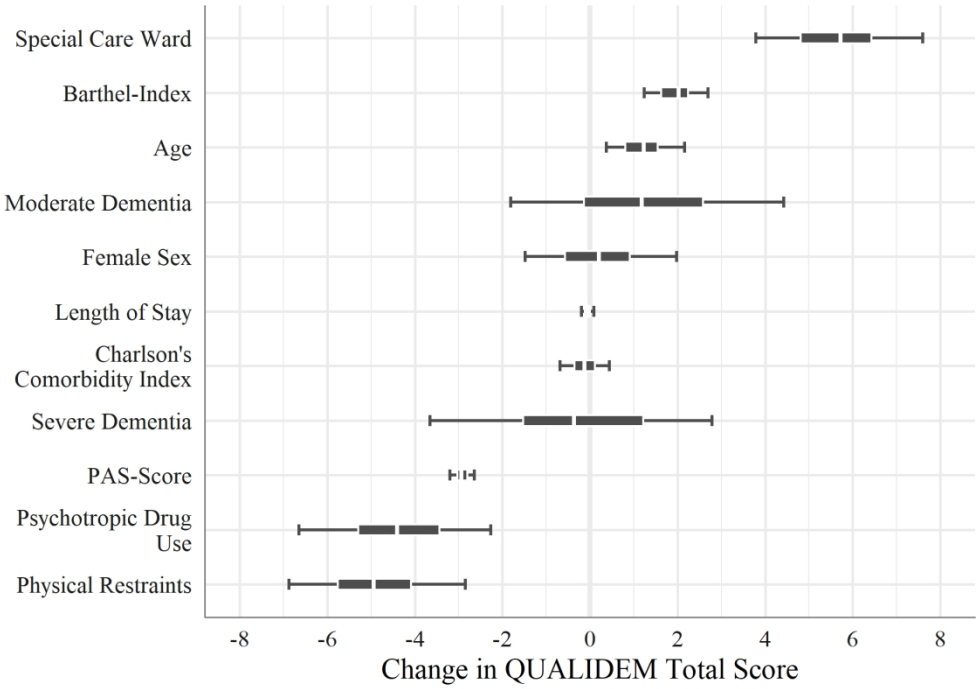
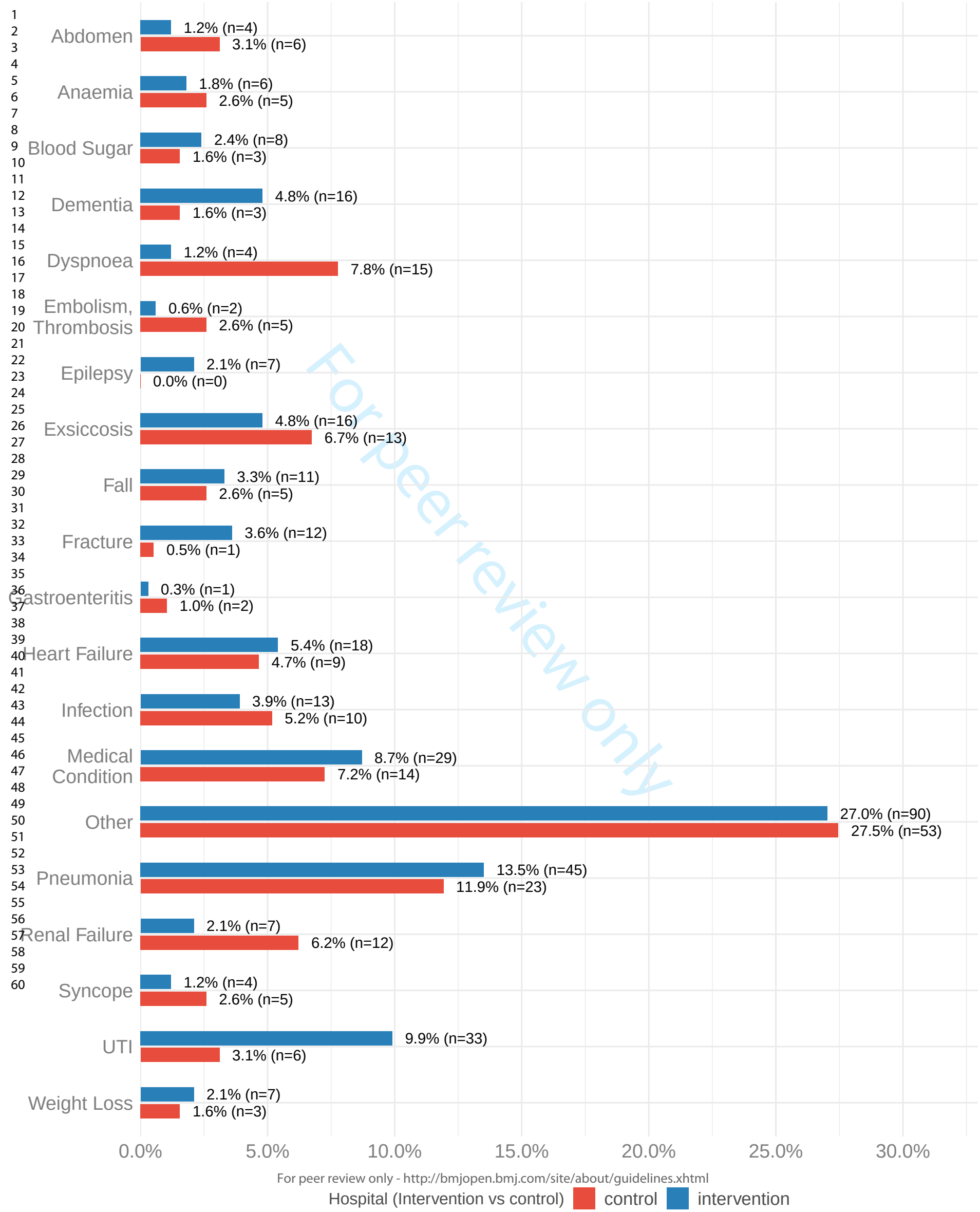


Figure 3: Predictors of Health Related Quality of Life, Regression Coefficients, Bayesian Linear Mixed Model, Posterior Median (+50% and 89% High Density Interval)

169x119mm (300 x 300 DPI)

Supplemental Material 1

Main Diagnosis by Hospital Ward



Supplemental Material 2

Quality of Life of Patients with Dementia in Acute Hospitals: Comparing a regular ward to a special care ward with dementia care concept. A Bayesian Multilevel Model Analysis.

Methodological comments

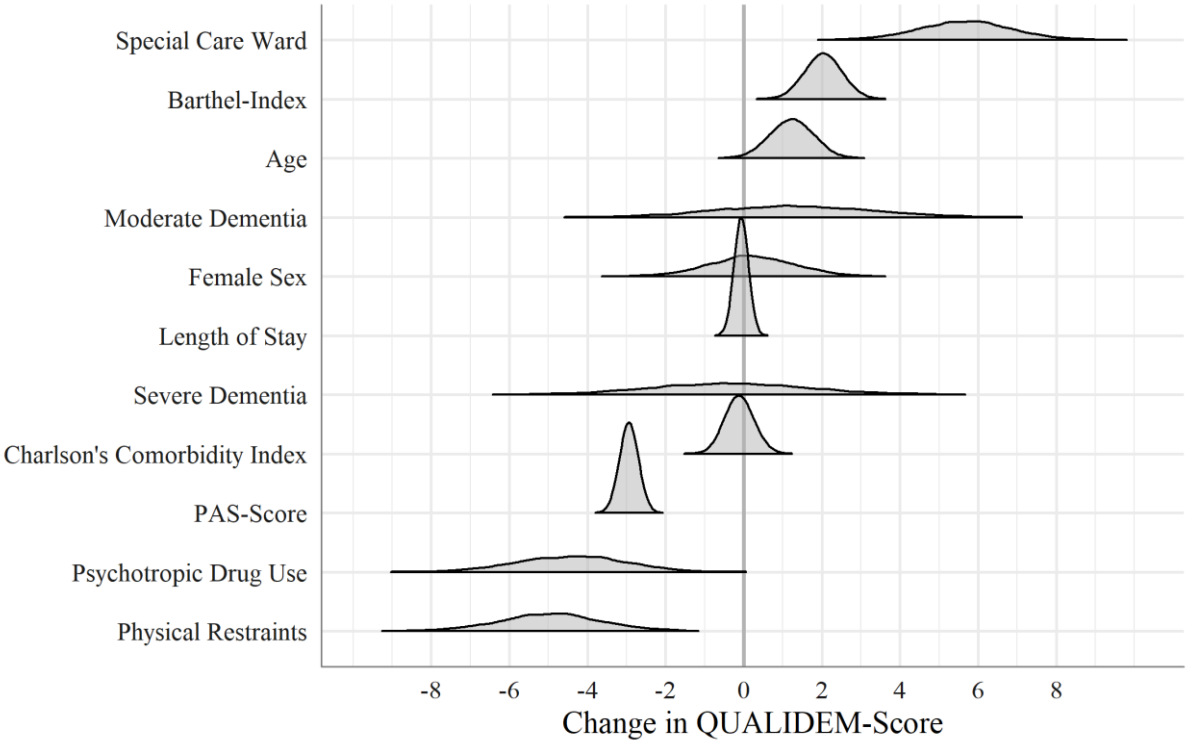
1. Short note on Bayesian Methods

Bayesian methods are probably not very common, and readers may ask "Why use Bayesian regression models?" Gelman et al.[1] give a well summarized answer to this question: "A pragmatic rationale for the use of Bayesian methods is the inherent flexibility introduced by their incorporation of multiple levels of randomness and the resultant ability to combine information from different sources, while incorporating all reasonable sources of uncertainty in inferential summaries. Such methods naturally lead to smoothed estimates in complicated data structures and consequently have the ability to obtain better real-world answers."[1]. Shortly speaking, advantages of Bayesian methods are, for instance, the ability to derive probability statements for every quantity of interest or explicitly incorporate prior knowledge about parameters into the model.

2. Distribution of Posterior Samples from the Regression Model

Contrary to a frequentist approach, there is no unique point estimate in Bayesian analysis. The posterior median minimises the expected absolute error (i.e. the difference of estimates from true values over posterior samples), but there are many other plausible values to describe the association between predictors and outcome, which is called the *distribution of posterior samples*. Credible or Highest Density Intervals are based on quantiles of this distribution and indicate the probability of an estimate being inside this interval. In terms of inference, this allows us to say that with, e.g. 50% probability, the true effect of a *Special Care Ward* on our outcome lies between 5.1 and 6.6. Figure S1 shows the distribution of posterior samples from our regression model.

Figure S1: *Distribution of Posterior Samples from Regression Model*



3. Tabular Summary from the Regression Model

Table S1 provides a summary of the regression model, including the data for two different Highest Density Intervals as well as the ratio of effective number of samples. The ratio should be 1.0 or close to 1.0 and indicates whether the samples were efficient, which means that the results for the related coefficient are reliable. Furthermore, all Rhat values of the models were approximately 1. The Rhat statistic[2] measures the ratio of the average variance of the draws within each chain to the variance of the pooled draws across chains. If the chains have not converged to a common distribution, the Rhat statistic will tend to be greater than one, indicating that more iterations may be required for reliable results.

Table S1: *Predictors of Quality of Life, Regression Coefficients, Bayesian Linear Mixed Model, Posterior Median (+50% and 89% Highest Density Intervals)*

Term	*Estimate*	*SE*	*50% HDI*	*89% HDI*	*Ratio of effective Number of Samples*
(Intercept)	46.2	2.2	45.2 – 48.2	42.7 – 49.8	1.00
Length of Stay	-0.1	0.1	-0.1 – 0.0	-0.2 – 0.1	1.00
Age	1.2	0.5	0.8 – 1.5	0.4 – 2.1	1.00
Moderate Dementia	1.2	2.0	-0.1 – 2.7	-1.8 – 4.6	1.00
Severe Dementia	-0.4	2.0	-1.6 – 1.1	-3.6 – 2.7	1.00
Female	0.2	1.1	-0.5 – 1.0	-1.6 – 1.9	1.00
Barthel Score	2.0	0.5	1.8 – 2.4	1.3 – 2.8	1.00
Physical Restraints (yes)	-4.9	1.2	-5.8 – -4.1	-7.0 – -2.8	1.00
Special Care Ward (Intervention)	5.7	1.2	4.9 – 6.5	3.8 – 7.6	1.00
PAS-Score	-2.9	0.2	-3.1 – -2.8	-3.2 – -2.7	1.00
Charlson's Comorbidity Index	-0.1	0.3	-0.4 – 0.1	-0.6 – 0.5	1.00
Psychotropic Drug Use (yes, as-needed)	-4.4	1.4	-5.3 – -3.5	-6.5 – -2.1	1.00
sigma	11.9	0.4	11.5 – 12.0	11.3 – 12.5	1.00

All Rhat values ~ 1, all mcse < 0.05. Results based on 4 chains each 1000 iterations for each chain. LOO-adjusted R^2: 0.500

4. Test for Practical Equivalence of Parameters

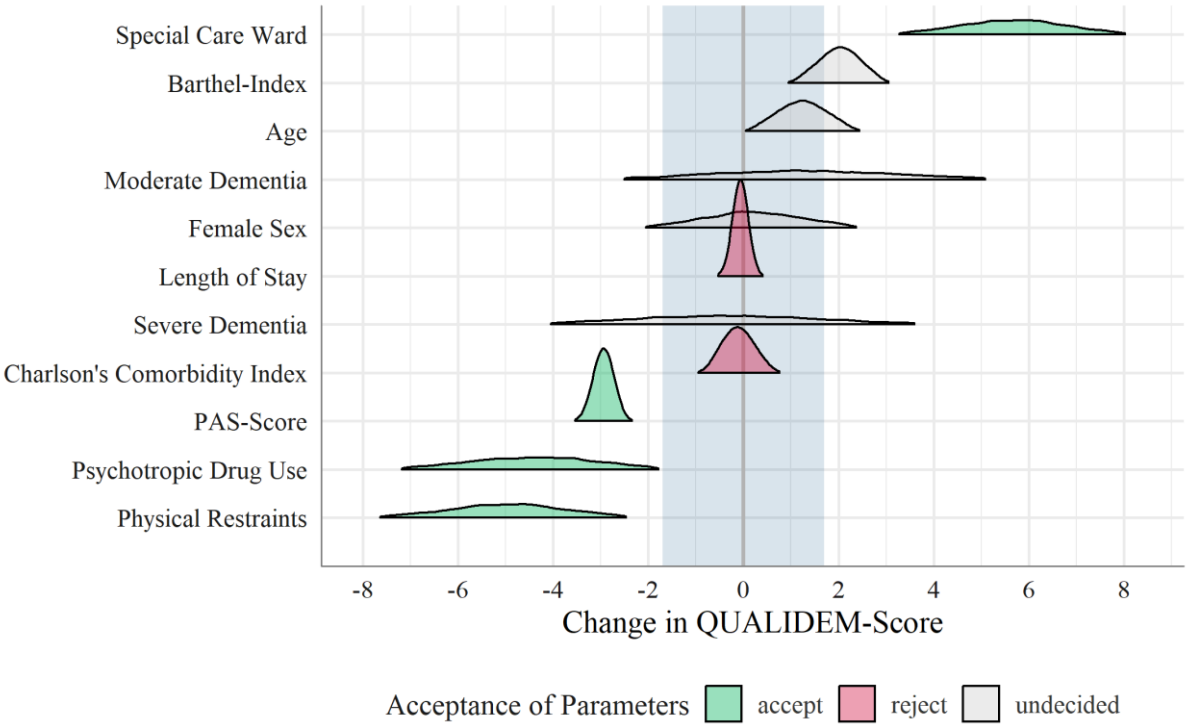
Bayesian methods do not perform classical "null hypothesis significance tests". Rather, to account for the uncertainty in estimation and the magnitude of parameter values, Kruschke[3] suggests checking whether parameter values lie inside a certain range that is considered as "practically no effect".

Figure S2 shows the distribution of the 95% Highest Density Intervals of posterior samples and indicates whether these lie completely outside, completely inside or partially inside the region of practical equivalence (ROPE). If the HDI is completely outside the ROPE, the "null hypothesis" for this parameter is "rejected". If the ROPE completely covers the HDI, i.e. all most credible values of a parameter are inside the region of practical equivalence, the null hypothesis is accepted. Else, it is undecided whether to accept or reject the null hypothesis.

The ROPE limits are specified following Kruschke's suggestion, which is defined as *0 +/- SD(dependent variable) * 0.1* for linear models.

Physical restraint, psychotropic drug use, agitation (PAS-score) and the intervention (special care ward) are those predictors that can be considered as relevant for our outcome.

Figure S2: *95%-Range of the Distribution of Posterior Samples from Regression Model; Region of Practical Equivalence emphasized in light-blue.*



5. Summary of the Distribution of Prior and Posterior Samples

Bayesian regression models can incorporate prior information about the parameters. Informative or weakly informative priors have the advantage of regularization of parameters to rule out large effects. This leads to less biased results and reductions of outliers in situations where the sample size in general, or for certain groups in the data, is small. Where we found information about our parameters, we included it by specifying *informative priors*. For instance, studies suggest that physical restraints are, on average, associated with an about 5-point reduction in quality of life, so the *location* of the prior distribution of our parameter physical restraint is at -5.0 (see Table S2). *Weakly informative priors* simply have a location of zero.

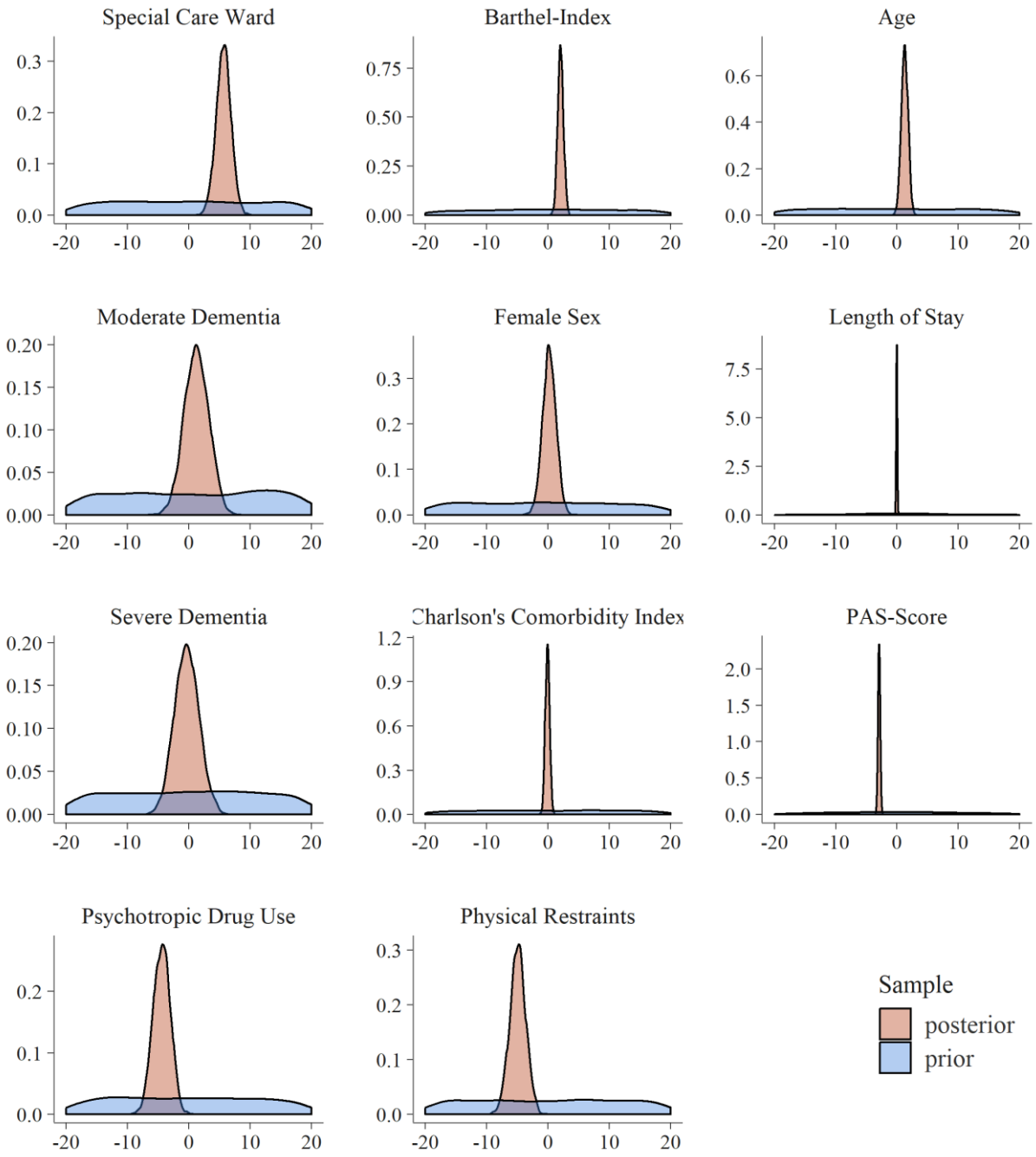
Table S2: *Prior Summary*

Predictor	Location	Scale
Length of Stay	0.00	6
Age	0.16	40
Moderate dementia	0.00	42
Severe dementia	-0.44	42
Sex, female	-3.22	42
Barthel-Score	0.00	29
Physical restraints	-5.00	42
Special Care Ward	0.00	42
PAS-Score	0.00	13
Charlson's Comorbidity Index	0.00	27
Psychotic Drug Use	0.29	42

To prevent a too strong shrinkage of uncertainty, the *scale* of the prior distribution should be large enough. If no assumptions about the scale of the effect are made, a recommended way (for normally distributed priors) is to use a scale of 2.5 and adjust it by dividing the scale by the standard deviation of the related predictor and multiply it with the standard deviation of the outcome (*2.5 * SD(y) / SD(x)*, see https://cran.r-project.org/web/packages/rstanarm/vignettes/priors.html).

When the evidence in the data is *strong*, the impact of the prior information is rather low. On the other hand, if the evidence of the data is *weak*, for instance due to small sample size, the impact of the prior information increases. Figure S3 shows the summary of the prior and posterior sample distributions and suggests that the evidence in the data was rather strong and results were not biased due to unrealistic prior information.

Figure S3: *Posterior versus Prior Summary*

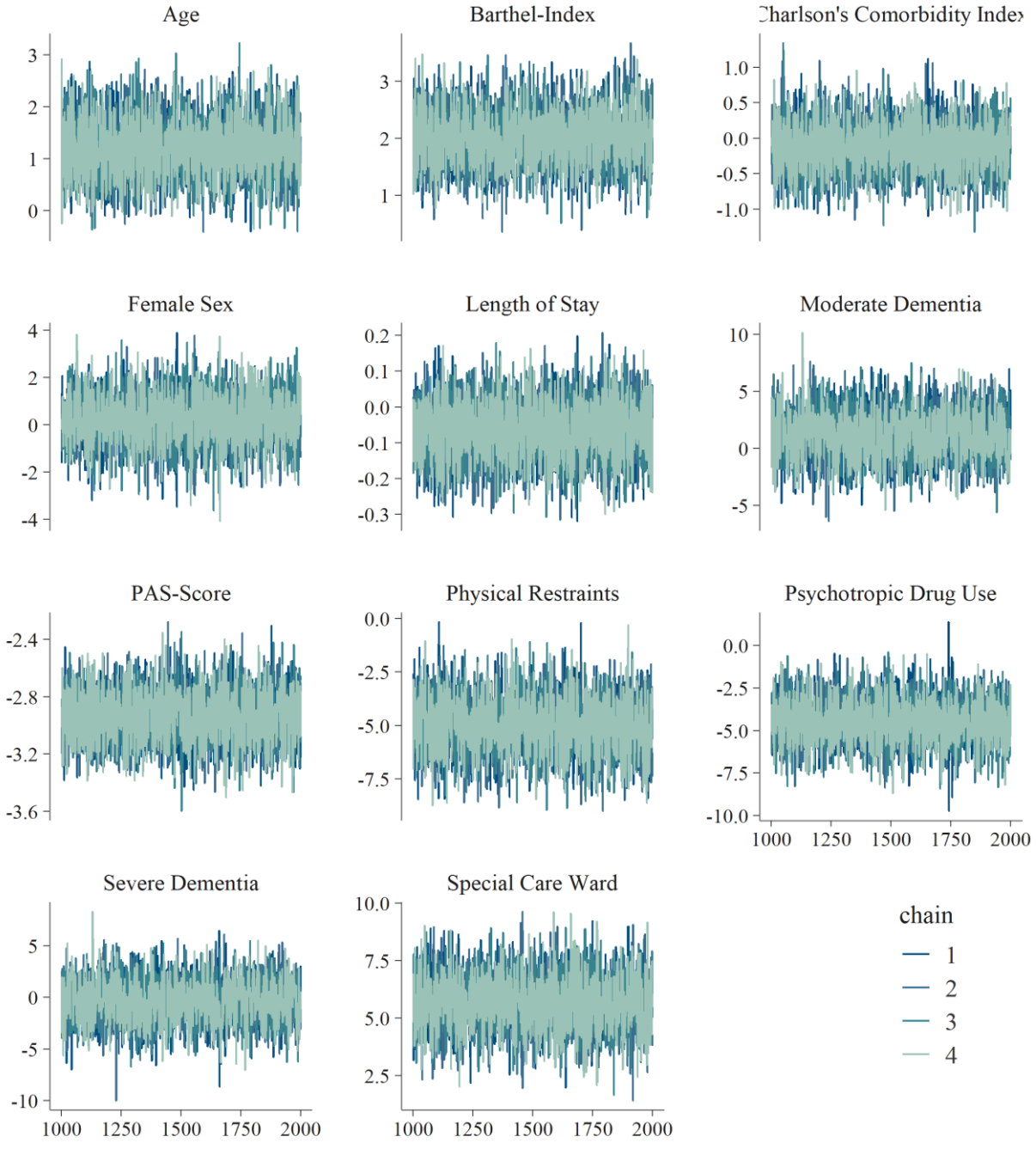


6. Trace plots

A trace plot provides a visual way to inspect sampling the behaviour and can be used, besides the Rhat-statistics, to assess convergence. The desired pattern is a "fat, hairy caterpillar", which shows no suspicious bends4, 5. If the chains converged to the posterior distribution, we can be confident in our model-based inferences.

The trace plot in Figure S4 indicates that all chains have converged well.

Figure S4: *Markov Chain trace plot, 4 chains with 1000 iterations each (excluding warmup-samples)*



7. References

1. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. Bayesian Data Analysis. Third edition. Boca Raton: CRC Press; 2014: p.24

2. Gelman A, Rubin DB. Inference from Iterative Simulation Using Multiple Sequences. Statistical Science. 1992;7(4):457-472. doi:10.1214/ss/1177011136

3. Kruschke JK. Rejecting or Accepting Parameter Values in Bayesian Estimation. Advances in Methods and Practices in Psychological Science. 2018; doi: 10.1177/2515245918771304

4. Sorensen T, Hohenstein S, Vasishth S. Bayesian linear mixed models using Stan: A tutorial for psychologists, linguists, and cognitive scientists. The Quantitative Methods for Psychology. 2016;12:175–200.

5. Lunn D, Jackson C, Best N, Thomas A, Spiegelhalter D. The BUGS book: a practical introduction to Bayesian analysis. Boca Raton, FL: CRC Press, Taylor & Francis Group; 2013.

Supplemental Material 3

To see if the complexity of patients' needs affect our findings, we calculated regression models with interaction terms between need factors moderated by hospitals. The figures S5 to S8 show the results of the interaction terms. Table S3 shows the coefficients in details. We found no "significant" associations for the interaction terms, concluding that the association between complexity of patients' needs and quality of life is not differing noticeably between the intervention and control group.

Plots were creates using the *ggeffects* package in R (Lüdecke D. ggeffects: Tidy Data Frames of Marginal Effects from Regression Models. Journal of Open Source Software. 2018;3: 772. doi:10.21105/joss.00772).

Figure S5: *Interaction between Barthel-Index and Intervention/Control-Group*

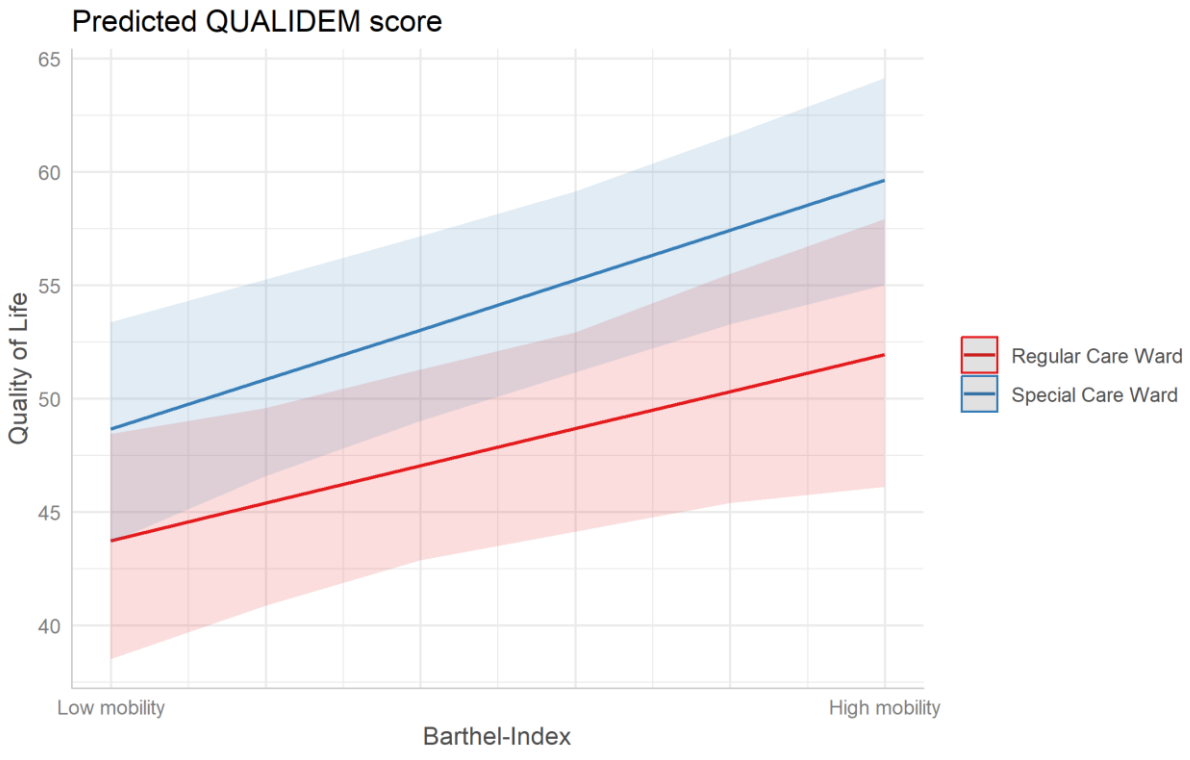


Figure S6: *Interaction between Physical Restraints and Intervention/Control-Group*

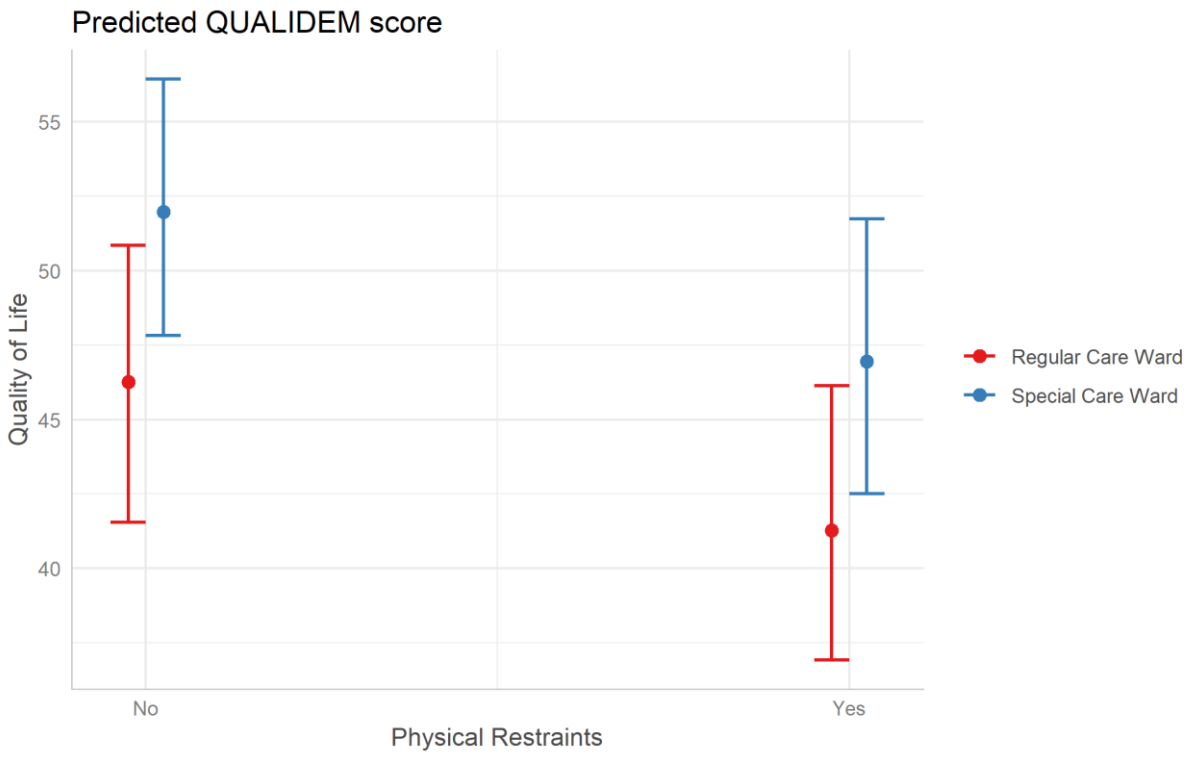


Figure S7: *Interaction between PAS and Intervention/Control-Group*

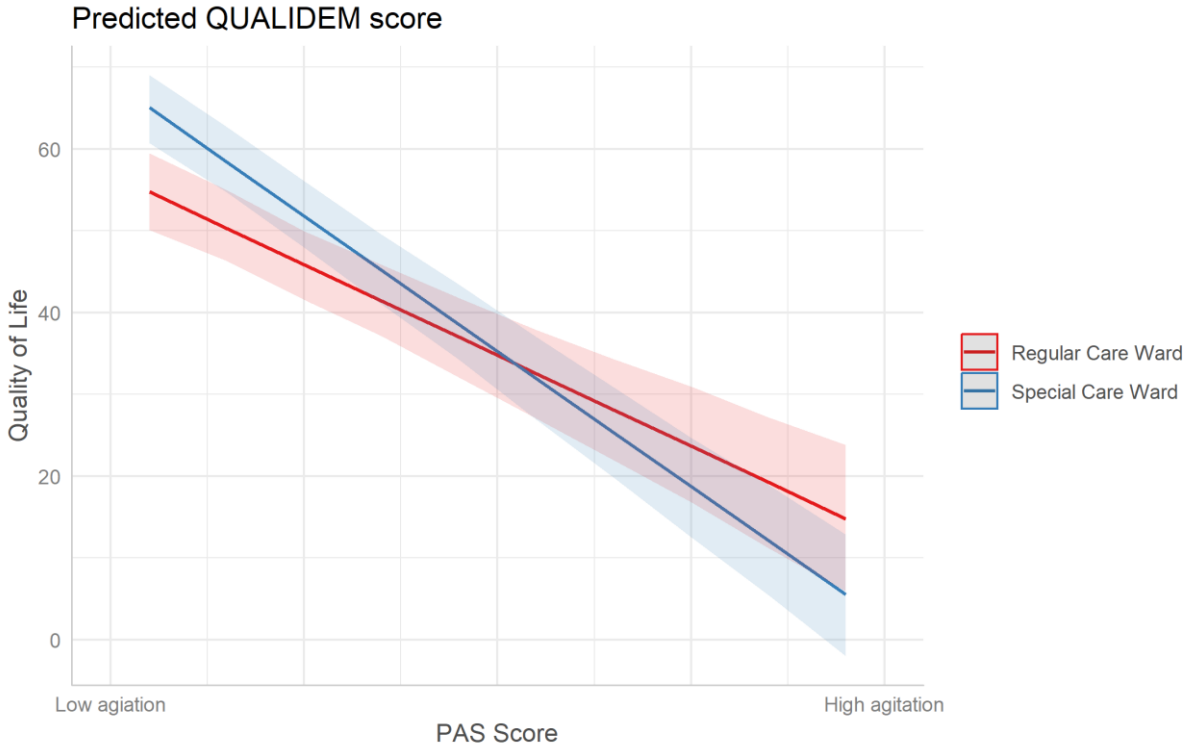


Figure S8: *Interaction between Chemical Restraints and Intervention/Control-Group*

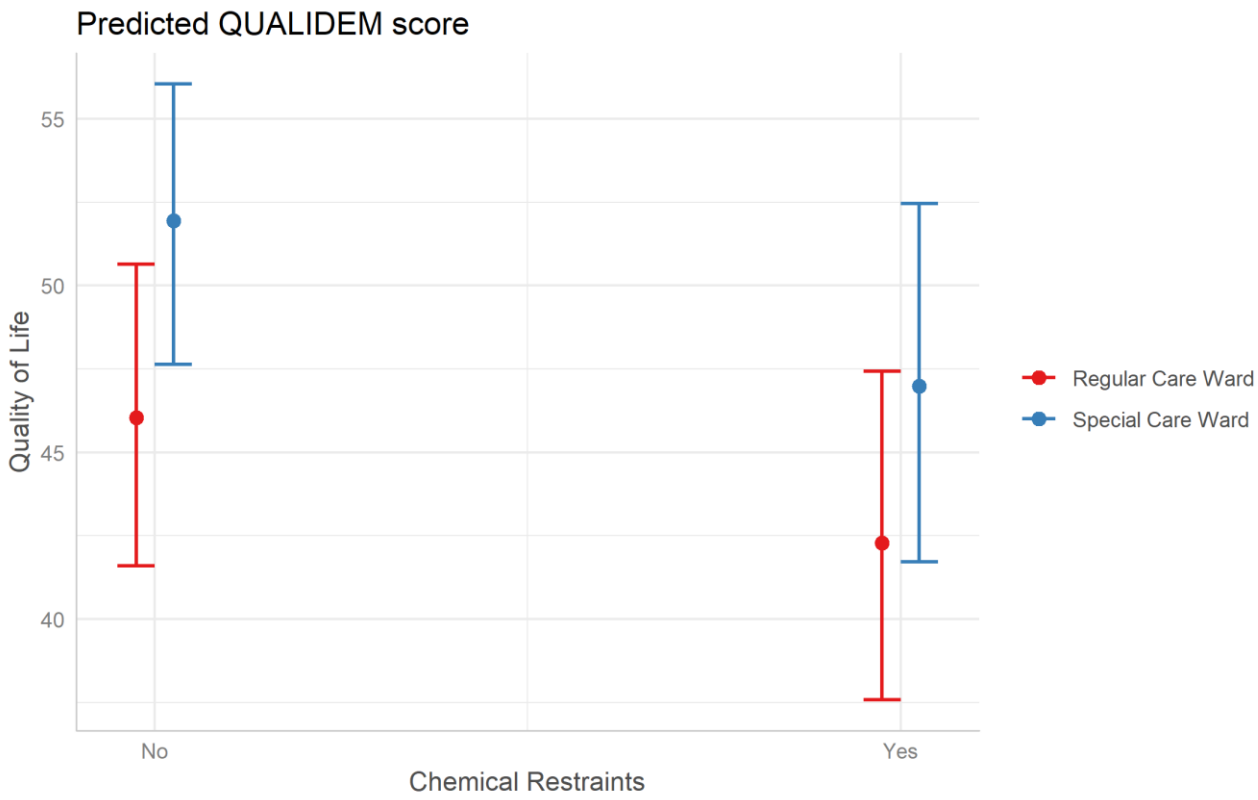


Table S3: *Coefficients for Models with Interaction Terms*

Terms	*Model: Interaction between Barthel and Hospital*		*Model: Interaction between Physical Restraints and Hospital*		*Model: Interaction between PAS and Hospital*		*Model: Interaction between Chemical Restraints and Hospital*	
	Estimates	*HDI (95%)*	*Estimates*	*HDI (95%)*	*Estimates*	*HDI (95%)*	*Estimates*	*HDI (95%)*
(Intercept)	36.18	26.19 – 45.67	36.31	27.00 – 46.10	36.57	27.11 – 45.71	36.16	26.45 – 45.69
Length of Stay	-0.07	-0.23 – 0.11	-0.06	-0.23 – 0.09	-0.07	-0.23 – 0.10	-0.06	-0.22 – 0.10
Age	0.12	0.01 – 0.23	0.12	0.02 – 0.23	0.12	0.01 – 0.22	0.12	0.02 – 0.23
Moderate Dementia	1.21	-2.79 – 5.13	1.19	-2.72 – 5.24	1.19	-2.83 – 4.72	1.25	-2.95 – 4.89
Severe Dementia	-0.34	-4.30 – 3.70	-0.44	-4.45 – 3.65	-0.74	-4.54 – 3.10	-0.30	-4.45 – 3.59
Female	0.16	-2.06 – 2.29	0.19	-1.98 – 2.29	0.52	-1.53 – 2.70	0.23	-1.76 – 2.53
Charlson's Comorbidity Index	-0.12	-0.81 – 0.59	-0.11	-0.77 – 0.64	-0.11	-0.77 – 0.59	-0.11	-0.82 – 0.59
Barthel Score	1.66	0.27 – 2.99	2.02	1.15 – 2.88	2.07	1.17 – 2.97	2.05	1.13 – 2.98
Physical Restraints (yes)	5.74	3.48 – 7.90	5.76	2.67 – 8.72	5.95	3.64 – 8.18	5.94	3.44 – 8.62
Special Care Ward (Intervention)	-5.02	-7.65 – -2.54	-4.92	-8.57 – -1.11	-4.94	-7.60 – -2.41	-4.96	-7.83 – -2.58
PAS-Score	-2.96	-3.30 – -2.62	-2.94	-3.30 – -2.61	-2.22	-2.77 – -1.66	-2.93	-3.27 – -2.59
Psychotropic Drug Use (yes, as-needed)	-4.43	-7.10 – -1.83	-4.40	-7.08 – -1.87	-4.30	-6.81 – -1.56	-3.76	-7.83 – -0.03
Barthel * Intervention	0.52	-1.00 – 2.06						
Phys. Restr. * Intervention			-0.09	-4.46 – 4.66				
PAS * Intervention					-1.09	-1.76 – -0.42		
Chem. Restr. * Intervention							-1.15	-6.63 – 4.27

Supplemental Material 4 – R Source Code

```
library(tidyverse)
library(ggridges)
library(sjmisc)
library(sjlabelled)
library(sjstats)
library(sjPlot)
library(insight)
library(bayestestR)
library(brms)

# Data available at https://doi.org/10.5281/zenodo.1479676

# load data ----

load("Dataset.RData")

# divide age by 10

d$age10 <- d$age / 10


# Labels for final model ----

labs <-
  c(
    stay_c = "Length of Stay",
    age = "Age",
    age10 = "Age",
    mmse2 = "Moderate Dementia",
    mmse3 = "Severe Dementia",
    sex2 = "Female Sex",
    barthel_code = "Barthel-Index",
    groupintervention = "Special Care Ward",
    physres1 = "Physical Restraints",
    pas_c = "PAS-Score",
    cci_c = "Charlson's Comorbidity Index",
    chemicalres1 = "Psychotropic Drug Use", # oder as-needed
    b_stay_c = "Length of Stay",
    b_age = "Age",
    b_age10 = "Age",
    b_mmse2 = "Moderate Dementia",
    b_mmse3 = "Severe Dementia",
    b_sex2 = "Female Sex",
    b_barthel_code = "Barthel-Index",
    b_groupintervention = "Special Care Ward",
    b_physres1 = "Physical Restraints",
    b_pas_c = "PAS-Score",
    b_cci_c = "Charlson's Comorbidity Index",
    b_chemicalres1 = "Psychotropic Drug Use" # oder as-needed
  )

# prior-definition in brms ----

# scale is 2.5 * sd(y) / sd(x)

bprior <-
  prior(normal(0, 6), class = "b", coef = "stay_c") +
  prior(normal(.1554, 40), class = "b", coef = "age10") +
  prior(normal(0, 42), class = "b", coef = "mmse2") +
  prior(normal(-.444, 42), class = "b", coef = "mmse3") +
  prior(normal(-3.219, 42), class = "b", coef = "sex2") +
  prior(normal(0, 29), class = "b", coef = "barthel_code") +
  prior(normal(-5, 42), class = "b", coef = "physres1") +
  prior(normal(0, 42), class = "b", coef = "groupintervention") +
  prior(normal(0, 13), class = "b", coef = "pas_c") +
  prior(normal(.2925, 42), class = "b", coef = "chemicalres1") +
  prior(normal(0, 26.77), class = "b", coef = "cci_c")

# see:
# Beerens HC, Sutcliffe C, Renom-Guiteras A, et al. Quality of Life and
# Quality of Care for People With Dementia Receiving Long Term Institutional
# Care or Professional Home Care: The European RightTimePlaceCare Study.
```

```
# Journal of the American Medical Directors Association. 2014;15(1):54-61.
# doi:10.1016/j.jamda.2013.09.010
#
# QoL-Scale ranges from 13-52 (40 points). Effects from those study are
# multiplied by 2.5 (rescaling 40-points to 100-point-scale of QUALIDEM)

# see:
# Barca, M. L., Engedal, K., Laks, J., & Selbæk, G. (2011). Quality of Life
# among Elderly Patients with Dementia in Institutions. Dementia and Geriatric
# Cognitive Disorders, 31(6), 435-442. https://doi.org/10.1159/000328969

# QoL-scale ranges from 11-55 (45 points). Effects from those study are
# multiplied by 2.2 (rescaling 45-points to 100-point-scale of QUALIDEM)


# model formula ----

mf <-
  formula(
    QoL ~ stay_c + age10 + mmse + sex + cci_c +
      barthel_code + physres + group + pas_c +
      chemicalres + (1 | maindiag)
  )


# brms-model ----

set.seed(1207)

m2a <- brm(
  formula = mf,
  data = d,
  prior = bprior,
  sample_prior = TRUE
)


# Figure 3 ----

theme_set(theme_sjplot2(base_size = 14, base_family = "serif"))

p <- plot_model(
  m2a,
  title = "",
  axis.labels = labs,
  sort.est = T,
  colors = c("grey30"),
  axis.title = "Change in QUALIDEM-Score",
  wrap.title = 100,
  wrap.labels = 20,
  width = .2,
  grid.breaks = 2,
  size.inner = .1
) +
  ylab("Change in QUALIDEM Total Score") +
  theme_sjplot2(base_size = 14, base_family = "serif")

p_pdf <- p +
  theme_sjplot2(base_size = 28, base_family = "serif") +
  theme(
    panel.grid.major = element_line(size = .1),
    panel.grid.minor = element_line(size = .05),
    axis.line.x      = element_line(size = .15),
    axis.line.y      = element_line(size = .15)
  )

ggsave(filename = "Fig3.tiff", plot = p, width = 170, height = 120, units = "mm", dpi = 300,
compression = "lzw")
ggsave(filename = "Fig3.pdf", scale = 2, plot = p_pdf, width = 170, height = 120, units = "mm",
dpi = 300)


# Appendix S1: Test for practical equivalence ----

rope(m2a, rope = c(-6, 6))
```

```
rope(m2a, rope = c(-7.5, 7.5))

equivalence_test(m2a, parameters = "^(?!prior)")
equivalence_test(m2a, parameters = "^(?!prior)") %>% plot()


# Appendix S1, Table Regression Coefficients ----

tab_df(tidy_stan(m2a, prob = c(.5, .89), digits = 1))


# Appendix S1, Prior Adjustement ----

insight::get_priors(m2a)


# Appendix S1, Figure distribution Posterior Samples ----

tmp <- m2a %>%
  as_tibble() %>%
  select(2:12) %>%
  gather(key = "predictor", value = "estimate") %>%
  to_factor(predictor)

tmp$predictor <- lvls_reorder(tmp$predictor, c(9, 4, 8, 3, 7, 11, 10, 6, 1, 2, 5))

p2 <- ggplot(tmp, aes(x = estimate, y = predictor)) +
  geom_vline(xintercept = 0, colour = "grey70", size = .8) +
  geom_density_ridges2(rel_min_height = 0.001, scale = 2, alpha = .5) +
  scale_x_continuous(breaks = seq(-8, 8, 2)) +
  scale_y_discrete(labels = labs) +
  labs(x = "Change in QUALIDEM-Score", y = NULL) +
  theme_sjplot2(base_size = 14, base_family = "serif")

ggsave(filename = "FigS1.tiff", plot = p2, width = 190, height = 120, units = "mm", dpi = 300)


# Appendix S1, test for practical equivalence ----

## Short version

equivalence_test(m2a, parameters = "^(?!prior)")
equivalence_test(m2a, parameters = "^(?!prior)") %>% plot()


## More beautiful tweaked version

tmp.hdi <- hdi(m2a, prob = .95) %>%
  slice(c(-c(1, 13:23)))

tmp2 <- m2a %>%
  as_tibble() %>%
  select(2:12) %>%
  map2_df(tmp.hdi$CI_low, function(x, y) {
    x[x < y] <- NA
    x
  }) %>%
  map2_df(tmp.hdi$CI_high, function(x, y) {
    x[x > y] <- NA
    x
  }) %>%
  gather(key = "predictor", value = "estimate") %>%
  to_factor(predictor)

tmp2$predictor <- lvls_reorder(tmp2$predictor, c(9, 4, 8, 3, 7, 11, 10, 6, 1, 2, 5))

tmp2$grp <- dplyr::case_when(
  tmp2$predictor %in% c("b_stay_c", "b_cci_c") ~ "reject",
  tmp2$predictor %in% c("b_age10", "b_mmse2", "b_mmse3", "b_sex2", "b_barthel_code") ~
"undecided",
  TRUE ~ "accept"
)

p3 <- ggplot(tmp2, aes(x = estimate, y = predictor, fill = grp)) +
  # rope based on "equi_test(model)".
```

```
  annotate("rect", xmin = -1.7, xmax = 1.7, ymin = 0, ymax = Inf, fill = sjplot_pal(palette =
"us")[1], alpha = 0.15) +
  geom_vline(xintercept = 0, colour = "grey70", size = .8) +
  geom_density_ridges2(rel_min_height = 0.01, scale = 2, alpha = .4) +
  scale_x_continuous(breaks = seq(-8, 8, 2)) +
  scale_y_discrete(labels = labs) +
  scale_fill_manual(values = sjplot_pal()[c(3, 1, 7)]) +
  labs(x = "Change in QUALIDEM-Score", y = NULL, fill = "Acceptance of Parameters") +
  theme_sjplot2(base_size = 14, base_family = "serif") +
  theme(
    legend.title = element_text(size = 13),
    legend.position = "bottom",
    axis.line.x = element_line(colour = "grey50"),
    axis.line.y = element_line(colour = "grey50")
  )

ggsave(filename = "FigS2.tiff", plot = p3, width = 190, height = 120, units = "mm", dpi = 300)


# Appendix S1, Posterior-Prior-Check ----

## Short version

plot_model(m2a, type = "diag", axis.lim = c(-20, 20))

## More beautiful tweaked version

pr_samp <- prior_samples(m2a) %>%
  select(starts_with("b_")) %>%
  gather(key = "Term", value = "Estimate") %>%
  mutate(Sample = "prior")

ps_samp <- posterior_samples(m2a) %>%
  select(starts_with("b_"), -b_Intercept) %>%
  gather(key = "Term", value = "Estimate") %>%
  mutate(Sample = "posterior")

m_pp_data <- bind_rows(pr_samp, ps_samp) %>% to_factor(Term)
m_pp_data$Term <- lvls_reorder(m_pp_data$Term, rev(c(8, 3, 10, 4, 7, 5, 11, 6, 1, 2, 9)))

p4 <- ggplot(m_pp_data, aes(x = Estimate, fill = Sample)) +
  geom_density(alpha = .4) +
  scale_x_continuous(limits = c(-20, 20)) +
  facet_wrap(
    ~ Term,
    scales = "free",
    labeller = labeller(Term = labs),
    nrow = 4
  ) +
  labs(x = NULL, y = NULL) +
  bayesplot::theme_default(base_size = 13) +
  theme(
    axis.line.x      = element_line(colour = "grey50"),
    axis.line.y      = element_line(colour = "grey50"),
    axis.text        = element_text(colour = "grey10"),
    axis.title       = element_text(colour = "black"),
    # strip.background = element_rect(colour = "grey50", fill = "grey90"),
    # strip.text       = element_text(colour = "grey20"),
    legend.title     = element_text(colour = "grey10"),
    legend.text      = element_text(colour = "grey20"),
    legend.position  = c(.5, .15),
    legend.justification = c(-2, 1)
  ) +
  scale_fill_manual(values = sjplot_pal("breakfast club")[c(1, 3)])

ggsave(filename = "FigS3.tiff", plot = p4, width = 190, height = 214, units = "mm", dpi = 300)


# Appendix S1, Traceplot ----

p <- rstan::traceplot(m2a$fit, pars = colnames(as.data.frame(m2a))[1:12], inc_warmup = F)
p$data$parameter <- as.character(p$data$parameter)
tmp <- p$data %>%
  filter(parameter != "b_Intercept")
```

```
for (i in 1:length(labs)) {
  if (names(labs)[i] %in% tmp$parameter) {
    r <- which(tmp$parameter == names(labs)[i])
    tmp$parameter[r] <- labs[i]
  }
}

p5 <- ggplot(tmp, aes(x = iteration, y = value, colour = chain)) +
  geom_line() +
  facet_wrap(~parameter, scales = "free_y", ncol = 3) +
  scale_color_manual(values = sjplot_pal("us", n = 4)) +
  labs(x = NULL, y = NULL) +
  bayesplot::theme_default(base_size = 13) +
  theme(
    axis.line.x      = element_line(colour = "grey50"),
    axis.line.y      = element_line(colour = "grey50"),
    axis.text        = element_text(colour = "grey10"),
    axis.title       = element_text(colour = "black"),
    # strip.background = element_rect(colour = "grey50", fill = "grey90"),
    # strip.text       = element_text(colour = "grey20"),
    legend.title     = element_text(colour = "grey10"),
    legend.text      = element_text(colour = "grey20"),
    legend.position  = c(.5, .15),
    legend.justification = c(-4.2, 0.7)
  )

ggsave(filename = "FigS4.tiff", compress = "lzw", plot = p5, width = 190, height = 214, units =
"mm", dpi = 300)
```

Supplemental Material 5

Regression Model with Alternative QUALIDEM-Score

A limitation of the QUALIDEM total score is how it is calculated for patients with very severe dementia. As recommended by the authors of QUALIDEM, only six of the nine subscales are used to calculate the total score [1]. Technically speaking, this is similar to simple mean value imputation of the missing scores for the three omitted subscales, resulting in biased and underestimated measurement error variance for this group. In order to see to which extent an alternative computation of the QUALIDEM total score might be more appropriate than the recommended approach, we calculated another regression model (called "New Model", see *Table S4*). This model used a QUALIDEM total score based on full imputation of missing values for *all nine* subscales for patients with very severe dementia (instead of only using six subscales).

Table S4 shows the result of the new model in comparison to the main model presented in the manuscript. Both results for the estimates as well as standard error and HDI are very similar. This suggests that in our particular case we found no improvement in the estimation accuracy after imputing the missing values for the originally omitted three subscales for people with very severe dementia. Still it might be statistically more sound to either have the same amount of subscales for all groups of dementia severity or to consider appropriate missing data imputation algorithms for the group with very severe dementia. Thus, future research in QoL for patients with dementia should also focus on the reliability and validity of the QUALIDEM.

Table S4: *Comparison of Models for two different QoL-Scores*

	New Model with full imputed QoL-Score for very severe dementia			*Main Model*		
Term	*Estimate*	*SE*	*89% HDI*	*Estimate*	*SE*	*89% HDI*
Length of Stay	-0.0	0.1	-0.2 – 0.1	-0.1	0.1	-0.2 – 0.1
Age	0.8	0.5	-0.0 – 1.6	1.2	0.5	0.4 – 2.1
Moderate Dementia	1.1	1.9	-1.8 – 4.1	1.2	2.0	-1.8 – 4.6
Severe Dementia	-0.3	1.9	-3.2 – 2.7	-0.4	2.0	-3.6 – 2.7
Female	0.6	1.0	-1.0 – 2.2	0.2	1.1	-1.6 – 1.9
Barthel Score	2.0	0.4	1.3 – 2.7	2.0	0.5	1.3 – 2.8
Physical Restraints (yes)	-4.1	1.3	-6.1 – -2.1	-4.9	1.2	-7.0 – -2.8
Special Care Ward (Intervention)	5.2	1.1	3.4 – 6.9	5.7	1.2	3.8 – 7.6
PAS-Score	-2.7	0.2	-3.0 – -2.4	-2.9	0.2	-3.2 – -2.7
Charlson's Comorbidity Index	0.0	0.3	-0.5 – 0.5	-0.1	0.3	-0.6 – 0.5
Psychotropic Drug Use (yes, as-needed)	-4.3	1.3	-6.3 – -2.3	-4.4	1.4	-6.5 – -2.1

All Rhat values ~ 1, all mcse < 0.05. Results based on 4 chains each 1000 iterations for each chain.

1 Dichter MN, Quasdorf T, Schwab CGG, et al. Dementia care mapping: effects on residents' quality of life and challenging behavior in German nursing homes. A quasi-experimental trial. International Psychogeriatrics 2015;27:1875–92. doi:10.1017/S1041610215000927

STROBE 2007 (v4) Statement—Checklist of items that should be included in reports of *case-control studies*

Section/Topic	Item #	Recommendation	Reported on page #
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	4-5
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	6-7
Participants	6	*(*a) Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls	6-7
		(*b*) For matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	8-10
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	8-10
Bias	9	Describe any efforts to address potential sources of bias	10
Study size	10	Explain how the study size was arrived at	7
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	10-11
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	10-11
		(*b*) Describe any methods used to examine subgroups and interactions	10-11
		(*c*) Explain how missing data were addressed	10
		(*d*) If applicable, explain how matching of cases and controls was addressed	N/A
		(*e*) Describe any sensitivity analyses	N/A
Results			

Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	12
		(b) Give reasons for non-participation at each stage	N/A
		(c) Consider use of a flow diagram	N/A
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	12-13
		(b) Indicate number of participants with missing data for each variable of interest	12-13
Outcome data	15*	Report numbers in each exposure category, or summary measures of exposure	12-13
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	13-14
		(*b*) Report category boundaries when continuous variables were categorized	N/A
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	N/A
Discussion			
Key results	18	Summarise key results with reference to study objectives	15-16
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	17-18
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	15-18
Generalisability	21	Discuss the generalisability (external validity) of the study results	18-19
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	20

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.