BMJ Open Quality

Albuminuria measurement in diabetic care: a multilevel analysis measuring the influence of accreditation on institutional performance

Nermin Ghith,[1,2] Juan Merlo,[2] Anne Frølich[1]

Check for updates

[1]Research Unit of Chronic Conditions, Center for Clinical Research and Prevention, Bispebjerg and Frederiksberg University Hospital, Frederiksberg, Denmark
[2]Unit for Social Epidemiology, Lunds Universitet, Lund, Sweden

Correspondence to
Dr Nermin Ghith;
Nermin.Ghith@regionh.dk

ABSTRACT

Background Studies assessing institutional performance regarding quality of care are frequently performed using single-level statistical analyses investigating differences between provider averages of various quality indicators. However, such analyses are insufficient as they do not consider patients' heterogeneity around those averages. Hence, we apply a multilevel analysis of individual-patient heterogeneity that distinguishes between 'general' ('latent quality' or measures of variance) and 'specific' (measures of association) contextual effects. We assess general contextual effects of the hospital departments and the specific contextual effect of a national accreditation programme on adherence to the standard benchmark for albuminuria measurement in Danish patients with diabetes.

Methods From the Danish Adult Diabetes Database, we extracted data on 137 893 patient cases admitted to hospitals between 2010 and 2013. Applying multilevel logistic and probit regression models for every year, we quantified general contextual effects of hospital department by the intraclass correlation coefficient (ICC) and the area under the receiver operating characteristic curve (AUC) values. We evaluated the specific effect of hospital accreditation using the ORs and the change in the department variance.

Results In 2010, the department context had considerable influence on adherence with albuminuria measurement (ICC=21.8%, AUC=0.770), but the general effect attenuated along with the implementation of the national accreditation programme. The ICC value was 16.5% in 2013 and the rate of compliance with albuminuria measurement increased from 91.6% in 2010 to 96% in 2013.

Conclusions Parallel to implementation of the national accreditation programme, departments' compliance with the standard benchmark for albuminuria measurement increased and the ICC values decreased, but remained high. While those results indicate an overall quality improvement, further intervention focusing on departments with the lowest compliance could be considered.

INTRODUCTION

In 2000, the Danish state developed the National Indicator Project for monitoring and improving quality of care for a number of chronic conditions.[1] By 2010, the National Indicator Project was expanded to include all clinical quality databases under one Secretariat, the Regions' Clinical Quality Improvement Program (RKKP).[1] Parallel to that, hospitals implemented three types of accreditation in different areas of the country between 2002 and 2015.[2] All hospitals were accredited by the Danish quality model (DDKM) in 2010–2013 and reaccredited in 2013–2015.[2]

Literature indicates that diabetes is responsible for, or associated with, a wide range of chronic conditions including cardiovascular diseases.[3] According to the Danish Diabetes Association, around 320 545 Danes (5.7% of the population in 2011) were diagnosed with diabetes, which is almost double the rate reported 10 years earlier. In 2004, Danish health authorities established the Adult Diabetes Database or *Dansk Voksen Diabetes Database* (DVDD) to register, monitor and support the quality of care in patients with diabetes.[4] All hospital departments are audited annually against a set of standard benchmarks.[4] Data on quality indicators are collected periodically for each hospital and aggregated for further analysis and release every year to the public and healthcare professionals. Hospital performance is compared with the national overall average for each indicator.

One of the mandatory high-volume processes is the albuminuria measurement that accounted for 12.8% of the total amount of diabetes care delivered in Denmark in 2010–2013. Albuminuria is a strong predictor of kidney impairment and cardiovascular disease in patients with both type 1 and type 2 diabetes.[5 6] Thus, periodic monitoring of albuminuria is crucial in those patients.[7] The DVDD audit commission recommended that hospital departments should measure albuminuria in at least 95% of patients with diabetes. Danish hospitals were audited against that benchmark.[4]

Several quality oversight schemes, including the National Indicator Project, use conventional single-level analyses for auditing healthcare providers' performance. This approach does not properly consider the potential correlation of patient observations within the healthcare unit (ie, multilevel structure of data).[8 9] Alternatively, multilevel regression analysis accounts for clustering of data while it allows a less biased estimation of uncertainty and leads to more appropriate ranking of healthcare providers.[9] Further, in single-level analyses, the size of the contextual (eg, healthcare unit) variance is difficult to interpret as it is quantified in isolation from the total patient variation. In the multilevel regression analysis of individual-patient heterogeneity, the contextual variance is just a component of the total patient variance. The relevance of the healthcare unit variance is, therefore, expressed by the share of the total patient variance that is at the unit level.[10 11] This variance partition coefficient or intraclass correlation coefficient (ICC) is fundamental in multilevel regression analysis,[12] and it allows us to evaluate the *general contextual effect* of the healthcare unit[10 11] that expresses the existence of both observable and not directly observable influences which conditions adherence (compliance) with clinical standards. The idea of *general contextual effect* is analogous to the concept of *latent healthcare unit quality variable*.[13 14] An analogous approach was embraced by the committee that set the statistical methods for profiling Medicaid and Medicare services in the USA.[15]

Besides the formal analysis of components of variance, relevant information on the *general contextual effect* can also be obtained by using measures of discriminatory accuracy such as the area under the receiver operating characteristic curve (AUC) for the estimation of random (healthcare unit) effects in multilevel regression analysis.[11 16 17] This AUC informs to what extent knowledge on a patient's healthcare provider or department offers accurate information to discriminate between patients who will and those who will not get the desired process or outcome of care.

Processes of care are influenced by the hospital's general context where the patients are treated. The concept of general contextual effect based on the variance partition coefficient or ICC can be traced to the work by Rodriguez and Goldman in 1995[18] (p. 74), among others, and it has been implicitly or explicitly applied over the last two decades in the evaluation of healthcare performance,[10 11 19–25] but it did not reach a wider application in literature on hospital accreditation. Therefore, we apply multilevel regression analysis of individual-patient heterogeneity, which combines multilevel analyses of variance components with the AUC measure[11] to quantify (1) the hospital departments' general contextual effects and (2) the specific contextual effect of accreditation on adherence to the mandatory standard benchmark for measuring albuminuria in patients with diabetes.

We assumed that the accreditation process should support hospital departments to provide more homogeneous, equitable, high-quality care to all patients with diabetes regardless of gender or age. We hypothesised that accreditation would increase the overall rates of compliance with the albuminuria measurement standard and reduce the general contextual effects (ie, between-department variance) over time.

POPULATION AND METHODS
Danish Adult Diabetes Database (DVDD), population and patient variables
The DVDD audit commission collects information on 16 quality indicators for patients with diabetes.[4] See online supplementary appendix SA1 for database inclusion criteria and completeness.

Study data were extracted from the DVDD on 137 893 patient cases admitted to 57 specialised outpatient departments in 32 hospitals between 2010 and 2013. Patients were 18 years or older. Data on albuminuria measurement are coded either as '1 for measured' or '0 for not measured'. Confounding factors are not usually a concern in analyses of process indicators that must be recorded for *all* eligible patients.[26] However, to investigate potential inequity in measuring albuminuria, we included gender as a dummy variable and age (in years) entered as a continuous variable in the statistical analyses.

Hospital accreditation variable
Accreditation is a dummy variable with four groups of departments (coded 1–4) categorised by duration and type of accreditation implemented by each hospital in Denmark. All hospitals in Copenhagen were accredited by the Joint Commission International (JCI) Accreditation programme during two different periods (groups 1 and 2). Group 1 included departments at five hospitals with the longest continuous experience with accreditation, as they were accredited by the JCI between 2002 and 2014, then by the mandatory Danish Quality Model (DDKM) between 2012 and 2015. We assigned group 1 as the reference group in comparisons. Group 2 included departments at seven hospitals that were accredited by the JCI between 2011 and 2014, then by DDKM in 2012–2015. Group 3 included departments at four small acute facilities in southern Denmark that were accredited by the British Health Quality Service model (HQS) in 2003/2004. These facilities merged into one larger hospital and were reaccredited by HQS in 2007, followed by DDKM between 2010 and 2015. Hence, groups 1–3 had DDKM as an extra dose of accreditation.[2] Group 4 included departments at 19 hospitals accredited for the first time by DDKM in 2010/2011 and in 2015. See online supplementary appendix SA2 for hospital accreditation timeline.

Statistical analyses
We applied multilevel logistic and probit regression analyses of variance[25 27] with patient cases with diabetes nested within departments in order to model albuminuria measurement. See online supplementary appendix SA3 for

equations, functions, models' estimation and fit, and statistical notes.

We developed three consecutive models (see below). For every model, we obtained two different measures of general contextual effects: (1) the ICC and (2) the AUC value. The ICC expresses the share of the total patient variation in the underlying propensity of being measured for albuminuria that is at the department level. The AUC is a measure of discriminatory accuracy[11 16 25 28] that measures the ability of the model to correctly discriminate between patient cases with or without a measurement for albuminuria.[29]

As quality profiling is carried out annually, we made separate analyses for every year (2010–2013).

Model 1: quantifying the general contextual effect (ie, hospital department)

Using hospital department ID codes, we fit model 1 as an empty two-level logistic regression model. Model 1 includes no explanatory variables (fixed effects). In this model, patient observations (albuminuria measurements) at the first level of the model are nested within hospital departments at the second level of the model. The model includes only a random intercept for the department level.

Model 2: quantifying specific contextual effect (ie, accreditation)

Model 2 provides a way to understand the mechanism underlying the observed department general effect in model 1. Therefore, to quantify the specific contextual effect, we added the department variable on accreditation as a fixed effect. In model 2, besides obtaining the ORs for the type of accreditation, we calculated the proportional change in variance as the percentage of department variance in the empty model 1 that was explained by adding accreditation in model 2.

Model 3: investigating potential inequities

In model 3, we wanted to investigate whether, aside from accreditation, the observed department variation in measuring albuminuria could be due to potential inequities (disparities)[30] in care induced by patient age or gender. We therefore extended model 2 by adding patient gender and age as fixed effects. To calculate the proportional change in variance between models 1 and 3, the variance of model 3 was rescaled to get an accurate value (see Hox et al,[31] p. 136).

We also aimed to determine the value-added of the information on gender and age to accurately classify the patients who will or will not have measured albuminuria over and above the department effects. Thus, we calculated the difference between the AUC values of models 3 and 1.

We performed the analyses using MLwiN V.2.22.[32]

Ranking of departments

For each year and model, we created a plot ranking departments according to their averages and their 95% CIs in the logarithmic OR scale. Further, we used the

Table 1 Descriptive statistics for patient cases with diabetes and hospital departments included in the study (2010–2013) (hospital departments are grouped by type of implemented accreditation)

	Group 1*	Group 2*	Group 3*	Group 4*
National compliance rates for measuring albuminuria (%)				
2010 (total %=91.6%)	79.3%	89.9%	95.8%	95.5%
2011 (total %=94.3%)	93.0%	95.3%	96.6%	94.3%
2012 (total %=96.5%)	96.8%	97.0%	96.2%	96.4%
2013 (total %=96%)	94.5%	95.7%	93.1%	96.5%
Patient cases (% of the total)				
2010 (total N=28 073)	20.54	10.26	4.96	64.24
2011 (total N=38 224)	13.30	7.95	4.74	74.02
2012 (total N=35 920)	13.40	9.98	5.28	71.33
2013 (total N=35 676)	13.83	5.90	5.52	74.76
Departments (N)				
2010 (N=40)	5	7	4	24
2011 (N=42)	5	7	4	26
2012 (N=42)	5	7	4	26
2013 (N=57)	7	7	8	35
Gender—females (%)				
2010	41.2	39.7	40.8	42.3
2011	40.8	40.6	41.2	41.5
2012	40.8	39.3	40.2	42.0
2013	40.1	37.6	40.4	41.7
Age—median				
2010	60.00	61.61	61.00	58.72
2011	59.00	62.69	61.00	59.00
2012	60.00	63.40	61.00	58.09
2013	61.00	63.04	62.00	58.20

*Group 1: hospitals accredited by the American model (Joint Commission International (JCI)) in 2002–2014, then by the Danish Quality Model (DDKM) in 2012–2015. Group 2: hospitals accredited by JCI in 2011–2014, then by DDKM in 2012–2015. Group 3: small hospitals accredited by the British model (Health Quality Service) in 2003–2010, then by DDKM in 2010–2015. Group 4: hospitals that were accredited for the first time by DDKM between 2010–2011 and 2015.

information provided by the ICC and the AUC values to interpret the league tables.[11 25 33 34]

RESULTS
Descriptive statistics

Table 1 shows that, overall, the compliance rate of albuminuria measurement increased between 2010 and 2012,

Table 2 Measures of association (fixed effects) obtained by two-level (patient and hospital departments) multilevel logistic regression (models 2 and 3) modelling measurement of albuminuria upon contact with hospital departments by patients with diabetes (values are ORs, 95% confidence intervals)

	Model 2*	Model 3*	Model 2*	Model 3*
	2010		2011	
Hospital accreditation (group 1)†	Reference	Reference	Reference	Reference
Group 2	2.69 (1.26 to 4.92)	2.91 (1.50 to 8.95)	2.03 (0.74 to 5.09)	2.67 (1.47 to 6.40)
Group 3	6.71 (2.52 to 17.73)	8.32 (1.95 to 23.62)	3.29 (1.07 to 10.39)	4.27 (1.49 to 11.28)
Group 4	5.99 (4.14 to 10.57)	7.75 (3.71 to 13.60)	2.14 (1.43 to 4.08)	2.94 (1.83 to 4.70)
Gender (male as reference)		0.91 (0.83 to 0.99)		1.02 (0.93 to 1.12)
Age (1 year increasing)		1.01 (1.00 to 1.01)		1.01 (1.01 to 1.01)
	2012		2013	
Hospital accreditation (group 1)†	Reference	Reference	Reference	Reference
Group 2	2.21 (1.00 to 5.46)	2.56 (1.03 to 7.70)	1.92 (0.65 to 5.28)	2.07 (0.84 to 5.28)
Group 3	0.78 (0.37 to 2.17)	1.13 (0.36 to 3.84)	0.97 (0.39 to 2.08)	0.96 (0.47 to 2.23)
Group 4	1.19 (0.82 to 2.55)	2.05 (0.83 to 3.49)	2.51 (1.46 to 4.65)	2.47 (1.57 to 4.78)
Gender (male as reference)		1.13 (1.00 to 1.27)		1.07 (0.96 to 1.18)
Age (1 year increasing)		1.02 (1.02 to 1.03)		1.02 (1.02 to 1.02)

*Model 1 is an empty model with no fixed effects, thus it is absent in this table. Model 2: a two-level model with accreditation. Model 3: a two-level model with accreditation, gender and age.
†Group 1: hospitals accredited by the American model (Joint Commission International (JCI)) in 2002–2014, then by the Danish Quality Model (DDKM) in 2012–2015. Group 2: hospitals accredited by JCI in 2011–2014, then by DDKM in 2012–2015. Group 3: small hospitals accredited by the British model (Health Quality Service) in 2003–2010, then by DDKM in 2010–2015. Group 4: hospitals that were accredited for the first time by DDKM between 2010–2011 and 2015.

to 96% in 2013. However, comparing 2013 with 2010, this increase was primarily due to the improved compliance in groups 1 and 2 with departments in hospitals accredited by JCI and DDKM.

Group 4 (accredited only by DDKM) maintained a high compliance rate throughout the whole study period. Group 3 (accredited by HQS and DDKM) also had a high compliance rate, but it slightly decreased in 2013.

Most of the patients were from departments that were accredited only by DDKM (group 4). During the study periods, most of the patients were men and the median age was similar in the four groups.

Measures of association

Table 2 shows that there were no substantial gender differences in measuring albuminuria. However, older patients had a slightly higher chance for having an albuminuria test than the younger ones.

We observed a clear association between department accreditation status and patient albuminuria measurement (table 2). However, the ORs for department accreditation were dissimilar. In 2010–2011, patients admitted to departments in hospitals accredited by HQS and by DDKM for the first time in 2010 (group 3) had, on average, better chances for having an albuminuria test than patients admitted to departments in hospitals that experienced JCI accreditation since 2002 (reference group 1). On the other hand, in 2012, in comparison with group 1, patients treated at departments in hospitals accredited for the first time by JCI in 2011 and DDKM in

2012 (group 2) had higher odds for albuminuria tests. Finally, in 2013, patients treated at departments in hospitals accredited only by DDKM (group 4) with no previous experience with international accreditation showed a higher OR than patients admitted at departments in group 1. We note that this latter group of departments (group 4) was in hospitals that were either preparing for the second cycle of DDKM accreditation in 2014 or were already reaccredited for the second time by DDKM in 2013.

Measures of variance and discrimination

Table 3 shows that, as expected, the ICC_{logit} values obtained from the multilevel logistic regression models were slightly higher than the values obtained from the probit models (ICC_{probit}).

Using the conservative measures of variance obtained from the multilevel probit regression models, we found that the component of the total variance at the department level (ie, the *general contextual effect*) decreased over time. We observed the largest department general effect in 2010 (ICC 21.8%), and the lowest in 2012 (ICC 12.3%), while there was a slight increase in the department general effect in 2013 (ICC 16.5%).

In 2010, accreditation explained more than one-fourth of between-department variance (proportional change in variance=27.39%). However, between 2011 and 2013, accreditation explained practically none or only a very minor part of the between-department variance (see table 3). In comparison with model 2, which includes

Table 3 Measures of variance (random effects) obtained by two-level (patient and hospital departments) multilevel logistic regression (models 1, 2 and 3) modelling measurement of albuminuria upon contact with hospital departments by patients with diabetes (variance measures are obtained separately using the logit and probit link functions and expressed as median values and 95% credible intervals)

	Model 1*	Model 2*	Model 3*	Model 1*	Model 2*	Model-3*
	2010			2011		
$Variance_{Logit}$	1.225 (0.789 to 2.041)	0.854 (0.523 to 1.465)	0.886 (0.543 to 1.537)	0.945 (0.573 to 1.642)	0.917 (0.545 to 1.64)	0.929 (0.564 to 1.618)
ICC_{Logit} (%)	27.25 (19.44 to 38.43)	20.71 (13.79 to 30.94)	21.32 (14.24 to 31.97)	22.42 (14.91 to 33.43)	21.9 (14.29 to 33.4)	22.12 (14.71 to 33.1)
$Variance_{Probit}$	0.278 (0.18 to 0.468)	0.202 (0.123 to 0.345)	0.198 (0.124 to 0.331)	0.179 (0.111 to 0.308)	0.171 (0.105 to 0.298)	0.178 (0.109 to 0.309)
$Variance_{Rescaled}$			0.191			0.174
ICC_{Probit} (%)	21.75 (15.25 to 31.88)	16.81 (10.95 to 25.65)	16.53 (11.03 to 24.87)	15.18 (9.99 to 23.55)	14.6 (9.5 to 22.96)	15.11 (9.83 to 23.61)
$ICC_{Probit\ Rescaled}$ (%)			16			14.8
PCV_{Probit} (%)	Reference	27.39	28.78	Reference	4.47	0.56
$PCV_{Probit\ Rescaled}$ (%)			31.16			2.61
	2012			2013		
$Variance_{Logit}$	0.752 (0.449 to 1.308)	0.748 (0.441 to 1.341)	0.776 (0.460 to 1.348)	1.026 (0.661 to 1.67)	0.953 (0.607 to 1.541)	0.908 (0.581 to 1.526)
ICC_{Logit} (%)	18.7 (12.07 to 28.57)	18.62 (11.88 to 29.08)	19.18 (12.33 to 29.19)	23.88 (16.82 to 33.81)	22.57 (15.66 to 32.03)	21.73 (15.09 to 31.82)
$Variance_{Probit}$	0.135 (0.084 to 0.239)	0.138 (0.083 to 0.241)	0.134 (0.08 to 0.232)	0.198 (0.129 to 0.324)	0.185 (0.119 to 0.302)	0.177 (0.115 to 0.284)
$Variance_{Rescaled}$			0.131			0.173
ICC_{Probit} (%)	11.9 (7.74 to 19.29)	12.13 (7.66 to 19.42)	11.82 (7.41 to 18.83)	16.53 (11.43 to 24.47)	15.61 (10.63 to 23.19)	15.04 (10.31 to 22.12)
$ICC_{Probit\ Rescaled}$ (%)			11.6			14.7
PCV_{Probit} (%)	Reference	−2.22	0.74	Reference	6.57	10.61
$PCV_{Probit\ Rescaled}$ (%)			2.54			12.72

*Model 1: an empty model with no fixed effects. Model 2: a two-level model with accreditation. Model 3: a two-level model with accreditation, gender and age.
ICC, intraclass correlation coefficient; PCV, proportional change in variance.

only the contextual variable on accreditation, adding patient characteristics (gender and age) in model 3 did not explain much of between-department variance for all years. Rescaling model 3 variance did not substantially change these findings.

Table 4 indicates that the highest values for AUC were observed for 2010 and 2013. As expected, adding accreditation as a contextual variable in model 2 did not improve AUC values in comparison with the empty model 1 which includes only the random intercepts for the department level. These values were almost unchanged when adding information on gender and age in model 3.

League tables

Figure 1 shows department differences in compliance with albuminuria measurement (2010–2013). The figure also indicates the values of the intraclass (department) correlation for measuring albuminuria (see models 1 and 2 in table 3). We present league tables for models 1 and 2 only, as adding gender and age in model 3 did not explain much of between-department variation.

Figure 1 shows that between 2010 and 2013 the 95% CI around a number of departments' averages did not overlap the overall mean of patient cases (study population's mean). However, in 2013, we observed an increased degree of uncertainty of the ranking which expresses itself as large overlapping of the 95% CIs.

DISCUSSION

We quantified both the general contextual effect of the hospitals and the specific contextual effect of accreditation on compliance with guidelines for albuminuria measurement (as per the standard benchmark) between 2010 and 2013. We also quantified the discriminatory accuracy of knowing where the patient was treated (ie, departments) for classifying diabetic patients who were or were not measured for albuminuria. As hypothesised, when national accreditation was progressively implemented, the overall rate of compliance with albuminuria measurement increased and general contextual effect of the department decreased, as indicated by the reduction in ICC and AUC values. These results suggest an improved performance along with an increased homogeneity in quality of care in these departments. However, department general effect remained high at the end of the observation period (ICC=16.5%), suggesting a potential need for further improvement in departments with the lowest compliance rates.

In practice, considering albuminuria measurement, empty multilevel models with no explanatory variables are appropriate for profiling hospital departments as processes of care are under full control of providers. In this case, there is no need to account for patient case-mix as confounding is not a major problem by definition.[35 36] The empty model, which includes only patient observations within departments, informs on the ceiling value of the department variance and its related *general contextual*

Table 4 Measures of the area under the receiver operating characteristic curve (AUC) along with their corresponding 95% credible intervals

	Model 1*	Model 2*	Model 3*
2010			
AUC values	0.770 (0.760 to 0.780)	0.770 (0.760 to 0.780)	0.772 (0.763 to 0.780)
Differences in AUC	Reference	0	0.002
2011			
AUC values	0.691 (0.680 to 0.701)	0.691 (0.680 to 0.701)	0.696 (0.686 to 0.707)
Differences in AUC	Reference	0	0.005
2012			
AUC values	0.698 (0.684 to 0.712)	0.698 (0.684 to 0.712)	0.716 (0.701 to 0.730)
Difference in AUC	Reference	0	0.018
2013			
AUC values	0.735 (0.722 to 0.748)	0.735 (0.722 to 0.748)	0.750 (0.737 to 0.763)
Difference in AUC	Reference	0	0.015

*Model 1: an empty model with no fixed effects. Model 2: a two-level model with accreditation. Model 3: a two-level model with accreditation, gender and age.
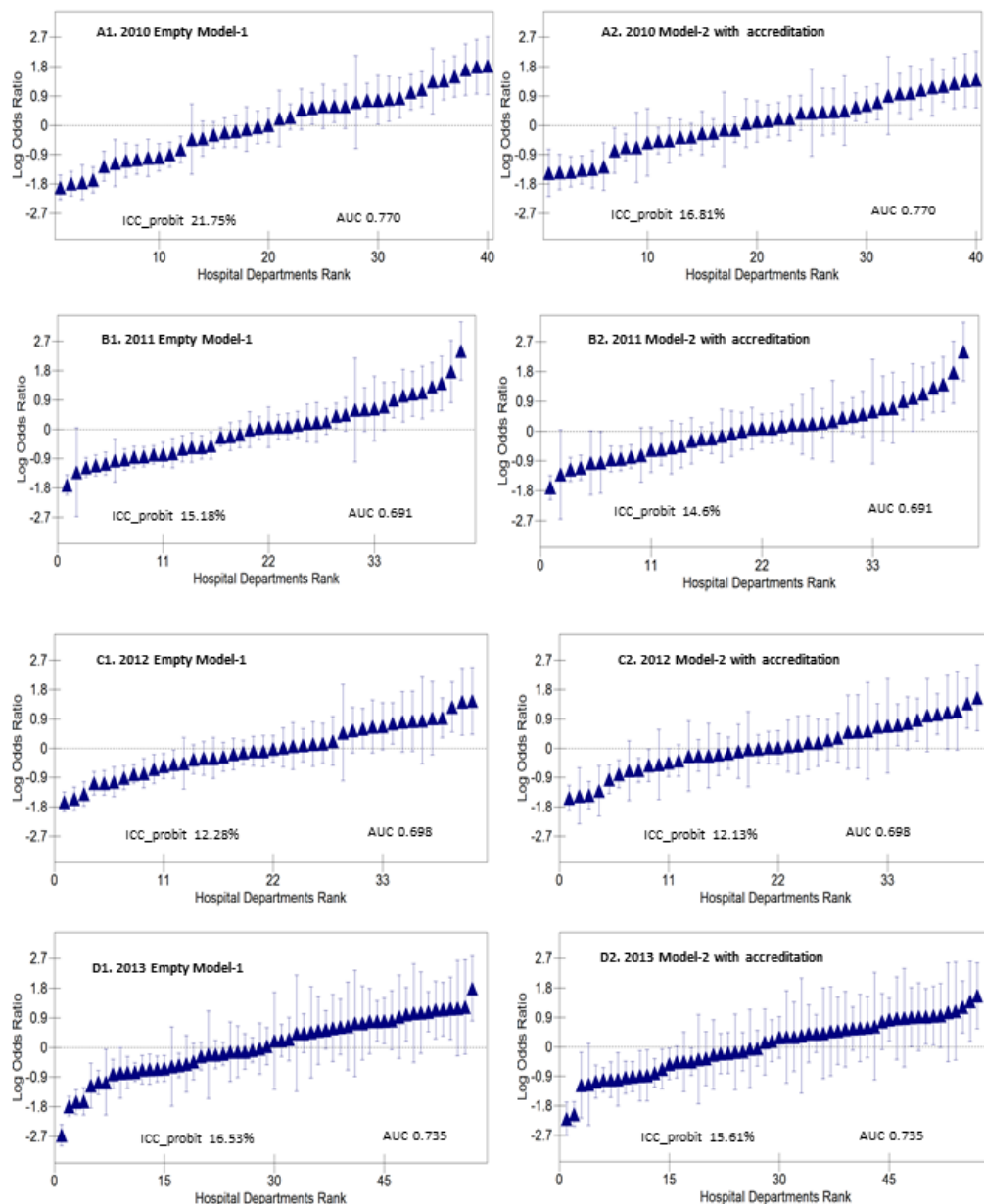
Figure 1 League tables of the hospital departments during the study period (A–D). Hospital departments were ranked by logarithm ORs (ie, shrunken residuals) with 95% confidence intervals (vertical lines) of measuring albuminuria using the overall average (study population's mean) as reference. AUC, area under the receiver operating characteristic curve; ICC, intraclass correlation coefficient.

effect.[25 37] If differences in the patient measurements of albuminuria are conditioned by the general context of the level of care (ie, indicated by meaningful ICC and AUC values), then the department care is heterogeneous. This indicates that the patient measurements of albuminuria were dependent on the department where the patient was treated.

In our study, the empty models had a rather good discriminatory ability, in reference to AUC values. Interpreting the ICC and AUC values driven from the empty models jointly with department league tables indicates a substantive department variance. This warrants ranking of departments, so departments with low performance

can be identified and targeted for further improvement, such that focused quality interventions might support below-average departments to increase their compliance rates. Hence, the *general contextual effect* (ie, differences between departments) could be further reduced, making the department performance more homogeneous. This premise is supported by the literature where a low provider variance would indicate a homogeneous clinical practice for a given level of care.[10 25 38]

The ICC values were highest in 2010, where accreditation explained some amount of between-department variance (proportional change in variance=27.39%). Yet, in 2011–2013, with the decrease in the ICC values, it seems

that accreditation could no longer explain much of the department variance.

In 2010, hospital departments in group 3 that experienced two different models of accreditation (HQS and DDKM) showed better compliance rates (OR values) than other departments. This could indicate that the 'extra dose' of accreditation contributed to an improved compliance rate by this group of departments. We could speculate that this group of departments could have benefited from their hospitals' experience with international accreditation (HQS).

In 2010, departments accredited only by DDKM (group 4) showed a better compliance rate (OR values) than departments accredited by JCI and DDKM (groups 1 and 2). It seems that DDKM accreditation activities improved department performance in this group.

Paradoxically, departments at hospitals that experienced accreditation by JCI since 2002 (group 1) did not show any superior performance (OR values) in comparison with other departments. Nevertheless, by 2010, departments in group 1 had already experienced the third accreditation survey by JCI in 2008 and were preparing for the fourth JCI survey in 2011 and the first DDKM survey in 2012. Thus, concerning the albuminuria indicator, these departments might have approached a ceiling level of performance which could not be further improved. Still, the required benchmark (95%), itself, is a ceiling level of compliance. Similar findings reported in literature, a study found that after the third/fourth accreditation cycle, hospital performance tends to stagnate as accreditation activities become less challenging.[39]

In 2010–2013, the department was a relevant level for understanding differences in a patient's measurement of albuminuria. However, considering department performance, it seems that accreditation by itself did not play a major (quantifiable) role during 2011–2013 as the proportional change in variance was relatively low. Interpreting only the ORs for the association between the accreditation type and albuminuria measurement for 2011–2013 would give misleading information on the role of accreditation. Yet, there might be a complementary influence for accreditation that could be identified with qualitative evaluations.[40 41]

Healthcare professionals tend to selectively implement and report elements of care that they perceive as important, but which are not necessarily part of the regulatory or accreditation mandates.[41–43] Disparities in care could therefore emerge, which could, in turn, be reflected in hospital performance. Some studies reported that elderly and female patients with diabetes might experience less than optimal care.[44–47] However, our results did not reveal any relevant inequities in care. In reference to albuminuria measurement, we can imply that quality improvement activities might have had a collective positive influence that shaped the Danish hospital departments as an equitable system.

Our results agree with previous findings in conventional literature. Several studies reported that accreditation

along with audit and feedback have the potential to influence and improve the delivery of clinical activities[30 48 49] such as measurement of albuminuria.[30] Hinchcliff *et al*[41] reported that accreditation can improve compliance with evidence-based processes of care. Further, an expert consensus report estimated that parallel to implementation of public reporting, feedback and pay for performance, among other strategies,[30] most commercial acute facilities in the USA increased compliance rates with a number of mandatory diabetes quality standards (including microalbuminuria) approaching 90%.[30]

We cannot rule out other influences beyond accreditation which could provide an alternative or complementary explanation for our results. Nevertheless, processes of care, such as the measurement of albuminuria, are under hospital control and hence are assumed to be sensitive to quality improvement interventions that target the hospital system in a given country.[49–53] Based on the literature,[22 25 41 54 55] it is reasonable to believe that the trend of increased compliance rates parallel to the implementation of the national accreditation (the Danish quality model) might actually reflect a behavioural change in diabetes care (ordering tests for albuminuria) by healthcare professionals.

In essence, the ultimate goal for performance monitoring and accreditation schemes is to encourage and guide providers to comply with evidence-based processes and standards of care to improve quality of clinical services.[55] Processes of diabetic care provide needed disease control activities against a set of risk factors[30] (eg, high levels of albuminuria) which in turn improve clinical outcomes (eg, prevention of kidney impairment).[5 7] Process-based indicators are considered to be direct sensitive performance measures for hospitals.[56] Such indicators measure clinical activities delivered to all eligible patients.[56] Process measures have direct interpretation, usually do not require risk adjustment and are actionable (ie, used for quality improvement).[56] Yet, process indicators could be manipulated and some measures may not be easily linked to patient outcomes.[56] However, any potential risk of gaming should be detected through the quality validation and controls routinely carried out by the health authorities to audit hospitals in Denmark.[1]

It is imperative to recognise that quality systems might be linked to diverse forms of organisational learning curves, non-linear butterfly effects[57] or even be confined to a 'hibernation phase' before showing any potential effects on processes and outcomes of care.[58 59] The hospital (department) variance could decrease, increase or stay the same over different observation periods and for different quality indicators.[58 59] Regardless of circumstance, careful assessment of the accreditation effect on hospital performance should be carried out for different conditions, indicators and at different phases of implementation.

A main limitation of our study is the short observation period, as data completeness was deficient before 2010. Nevertheless, we included all departments that

handle patients with diabetes and completeness of data on albuminuria measurements is approaching 100% in the database. Yet, our results concern Danish hospitals, so these cannot be directly generalised to other healthcare systems.

Several studies have employed multilevel analyses for evaluating hospital accreditation both in Denmark[60–62] and in other countries.[63 64] However, those studies consider variance and clustering of patient outcomes within hospitals as a sort of noise that should only be accounted for in the statistical analysis to obtain correct estimations of the SE around the regression coefficients.[25] In contrast, we consider variance and clustering as very relevant for assessing institutional performance and the influence of accreditation programmes on patient care. Traditional estimates of specific contextual effects based on average measures of association (eg, ORs for accreditation groups) provide insufficient information if they are not accompanied by measures of general contextual effects (ie, ICC).[65] Besides, compared with traditional single-level analyses, our study also provides more accurate department ranking. Our investigation, thus, provides improved information on the effect of accreditation in Denmark than that presented in previous publications.[62–64]

By applying our analytical approach to real-life evaluations of hospital performance, clinicians and policy-makers can assess whether or not outcomes of care are in compliance with clinical standards. If not, they can use quantitative information on the *general contextual effect* to plan for either a focused quality strategy targeting inconsistent performance or a compressive quality strategy targeting homogeneous poor performance. Such knowledge cannot be directly attained using traditional profiling models that do not report measures of provider contextual effects.[10 25 66 67]

CONCLUSION

Our multilevel approach represents an improved method for evaluating healthcare provider performance and the comparative effectiveness of large-scale interventions. Applying this approach, we analysed compliance with the mandatory standard of measuring albuminuria in patients with diabetes and the comparative effectiveness of hospital accreditation in Denmark on this process indicator. Accreditation appeared to be correlated with an overall improvement in quality expressed by an increased compliance rate across time and increased department performance homogeneity. The *general contextual effect* at the end of the observation period remained high, which suggests focused interventions targeting departments with the lowest compliance rates.

REFERENCES

1. Mainz J, Kristensen S, Bartels P. Quality improvement and accountability in the Danish health care system. *Int J Qual Health Care* 2015;27:523–7.
2. Triantafillou P. Against all odds? Understanding the emergence of accreditation of the Danish hospitals. *Soc Sci Med* 2014;101:78–85.
3. Marwick TH. Diabetic heart disease. *Postgrad Med J* 2008;84:188–92.
4. Jørgensen ME, Kristensen JK, Reventlov Husted G, *et al*. The danish adult diabetes registry. *Clin Epidemiol* 2016;8:429–34.
5. Heerspink HJ, Holtkamp FA, de Zeeuw D, *et al*. Monitoring kidney function and albuminuria in patients with diabetes. *Diabetes Care* 2011;34 Suppl 2:S325–S329.
6. Satchell SC, Tooke JE. What is the mechanism of microalbuminuria in diabetes: a role for the glomerular endothelium? *Diabetologia* 2008;51:714–25.
7. Gross JL, de Azevedo MJ, Silveiro SP, *et al*. Diabetic nephropathy: diagnosis, prevention, and treatment. *Diabetes Care* 2005;28:164–76.
8. Merlo J. Changing analytical approaches in European epidemiology—a short comment on a recent article. *Eur J Epidemiol* 2005;20:737–37.
9. Goldstein H, Spiegelhalter DJ. League tables and their limitations: statistical issues in comparisons of institutional performance. *J R Stat Soc Ser A Stat Soc* 1996;159:385–443.
10. Merlo J, *et al*. Survival after initial hospitalisation for heart failure: a multilevel analysis of patients in Swedish acute care hospitals. *Journal of Epidemiology & Community Health* 2001;55:323–9.
11. Merlo J, Wagner P, Ghith N, *et al*. An original stepwise multilevel logistic regression analysis of discriminatory accuracy: the case of neighbourhoods and health. *PLoS One* 2016;11:e0153778.
12. Goldstein H, Browne W, Rasbash J. Partitioning variation in multilevel models. *Understanding Statistics* 2002;1:223–31.
13. Normand SL. Registry studies for improving the quality of cardiovascular care: the role of variance components. *Circ Cardiovasc Qual Outcomes* 2012;5:e42–e43.
14. Krumholz HM, Lin Z, Normand SL. Measuring hospital clinical outcomes. *BMJ* 2013;346:f620.
15. Ash AS, Fienberg SE, Louis TA, *et al*. Statistical issues in assessing hospital performance commissioned by the Committee of presidents of statistical societies. 2011 https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/Downloads/Statistical-Issues-in-Assessing-Hospital-Performance.pdf.
16. Wagner P, Merlo J. Discriminatory accuracy of a random effect in multilevel logistic regression. *International Journal of Epidemiology* 2015;44:i49–i50.
17. Wagner P, Merlo J. Measures of discriminatory accuracy in multilevel analysis. *Eur J Epidemiol* 2013;28:135.
18. Rodriguez G, Goldman N. An assessment of estimation procedures for multilevel models with binary responses. *J R Stat Soc Ser A Stat Soc* 1995;158:73–8.

19. Davis P, Gribben B. Rational prescribing and interpractitioner variation. A multilevel approach. *Int J Technol Assess Health Care* 1995;11:428–42.
20. Merlo J, Gerdtham UG, Eckerlund I, *et al*. Hospital level of care and neonatal mortality in low- and high-risk deliveries: reassessing the question in Sweden by multilevel analysis. *Med Care* 2005;43:1092–100.
21. Ohlsson H, Librero J, Sundquist J, *et al*. Performance evaluations and league tables: do they capture variation between organizational units? An analysis of 5 Swedish pharmacological performance indicators. *Med Care* 2011;49:327–31.
22. Hjerpe P, Ohlsson H, Lindblad U, *et al*. Understanding adherence to therapeutic guidelines: a multilevel analysis of statin prescription in the Skaraborg Primary Care Database. *Eur J Clin Pharmacol* 2011;67:415–23.
23. Taljaard M, Donner A, Villar J, *et al*. Intracluster correlation coefficients from the 2005 WHO Global Survey on Maternal and Perinatal Health: implications for implementation research. *Paediatr Perinat Epidemiol* 2008;22:117–25.
24. Fink JC, Gardner JF, Armistead NC, *et al*. Within-center correlation in dialysis adequacy. *J Clin Epidemiol* 2000;53:79–85.
25. Ghith N, Wagner P, Frølich A, *et al*. Short term survival after admission for heart failure in Sweden: applying multilevel analyses of discriminatory accuracy to evaluate institutional performance. *PLoS One* 2016;11:e0148187.
26. Rubin HR, Pronovost P, Diette GB. The advantages and disadvantages of process-based measures of health care quality. *Int J Qual Health Care* 2001;13:469–74.
27. Austin PC, Merlo J. Intermediate and advanced topics in multilevel logistic regression analysis. *Stat Med* 2017;36:3257–77.
28. Lagerlund M, Merlo J, Vicente RP, *et al*. Does the neighborhood area of residence influence non-attendance in an urban mammography screening program? A multilevel study in a Swedish city. *PLoS One* 2015;10:e0140244.
29. Pepe MS, Janes H, Longton G, *et al*. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *Am J Epidemiol* 2004;159:882–90.
30. O'Connor PJ, Bodkin NL, Fradkin J, *et al*. Diabetes performance measures: current status and future directions. *Diabetes Care* 2011;34:1651–9.
31. Hox JJ, Moerbeek M, van de Schoot R. *Multilevel analysis: techniques and applications*. New York: Routledge, 2010.
32. Rasbash J, Charlton C Browne WJ *et al*. MLwiN Version 2.10. Centre for Multilevel Modelling: University of Bristol, 2009.
33. Merlo J, Ostergren PO, Broms K, *et al*. Survival after initial hospitalisation for heart failure: a multilevel analysis of patients in Swedish acute care hospitals. *J Epidemiol Community Health* 2001;55:323–9.
34. Merlo J, Viciana-Fernández FJ, Ramiro-Fariñas D, *et al*. Bringing the individual back to small-area variation studies: a multilevel analysis of all-cause mortality in Andalusia, Spain. *Soc Sci Med* 2012;75:1477–87.
35. Hermann RC, Rollins CK, Chan JA. Risk-adjusting outcomes of mental health and substance-related care: a review of the literature. *Harv Rev Psychiatry* 2007;15:52–69.
36. Mainz J. Developing evidence-based clinical indicators: a state of the art methods primer. *Int J Qual Health Care* 2003;15:5i–11.
37. Ding VY, Hubbard RA, Rutter CM, *et al*. Assessing the accuracy of profiling methods for identifying top providers: performance of mental health care providers. *Health Serv Outcomes Res Methodol* 2013;13:1–17.
38. Chin DL, Bang H, Manickam RN, *et al*. Rethinking thirty-day hospital readmissions: shorter intervals might be better indicators of quality of care. *Health Aff* 2016;35:1867–75.
39. Pomey MP, Lemieux-Charles L, Champagne F, *et al*. Does accreditation stimulate change? A study of the impact of the accreditation process on Canadian healthcare organizations. *Implement Sci* 2010;5:31.
40. Braithwaite J, Westbrook J, Pawsey M, *et al*. A prospective, multi-method, multi-disciplinary, multi-level, collaborative, social-organisational design for researching health sector accreditation [LP0560737]. *BMC Health Serv Res* 2006;6:113.
41. Hinchcliff R, Greenfield D, Moldovan M, *et al*. Narrative synthesis of health service accreditation literature. *BMJ Qual Saf* 2012;21:979–91.
42. Neuhauser D, Provost L, Bergman B. The meaning of variation to healthcare managers, clinical and health-services researchers, and individual patients. *BMJ Qual Saf* 2011;20:i36–i40.
43. Healthcare Research and Quality. *Patient safety and quality: an evidence-based handbook for nurses*. Rockville MD: Agency for Healthcare Research and Quality, 2008.
44. Bird CE, Fremont A, Hanson M. *Mapping gender differences in cardiovascular disease and diabetes care: a pilot assessment of LDL cholesterol testing rates in a California health plan*. Santa Monica. CA: RAND Corporation, 2014.
45. Sen G, Ostlin P. Gender inequity in health: why it exists and how we can change it. *Glob Public Health* 2008:3:1–12.
46. Siddiqui MA, Khan MF, Carline TE. Gender differences in living with diabetes mellitus. *Mater Sociomed* 2013;25:140–2.
47. Kleinsinger F. Working with the noncompliant patient. *Perm J* 2010;14:54–60.
48. Alkhenizan A, Shaw C. Impact of accreditation on the quality of healthcare services: a systematic review of the literature. *Ann Saudi Med* 2011;31:407.
49. Nicklin W. The value and impact of health care accreditation: a literature review Accreditation Canada. 2015 https://aventa.org/pdfs/valueimpactaccreditation.pdf
50. Greenfield D, Kellner A, Townsend K, *et al*. Health service accreditation reinforces a mindset of high-performance human resource management: lessons from an Australian study. *Int J Qual Health Care* 2014;26:372–7.
51. Ng GK, Leung GK, Johnston JM, *et al*. Factors affecting implementation of accreditation programmes and the impact of the accreditation process on quality improvement in hospitals: a SWOT analysis. *Hong Kong Med J* 2013;19:434–46.
52. Almoajel AM. Relationship between accreditation and quality indicators in hospital care: a review of the literature. *World Applied Sciences Journal* 2012;17:598–606.
53. Ng K. *A systematic review of the impact of accreditation on quality improvement in hospitals [master thesis]*: University of Hong Kong, 2010.
54. Groene O, Skau JK, Frølich A. An international review of projects on hospital performance assessment. *Int J Qual Health Care* 2008;20:162–71.
55. Ghith N, Frølich A, Merlo J. The role of the clinical departments for understanding patient heterogeneity in one-year mortality after a diagnosis of heart failure: a multilevel analysis of individual heterogeneity for profiling provider outcomes. *PLoS One* 2017;12:e0189050.
56. Gliklich R, Dreyer N, Leavy M. *Registries for evaluating patient outcomes: a user's guide*. Rockville, MD: Agency for Healthcare Research and Quality, 2014.
57. Brown JA. *healthcare quality handbook: a professional resource and study guide*. 28 edn. Sierra Madre, California: JB Quality Solutions, 2015.
58. Selby JV, Schmittdiel JA, Lee J, *et al*. Meaningful variation in performance: what does variation in quality tell us about improving quality? *Med Care* 2010;48:133–9.
59. Fung V, Schmittdiel JA, Fireman B, *et al*. Meaningful variation in performance: a systematic literature review. *Med Care* 2010;48:140–8.
60. Falstie-Jensen AM, Larsson H, Hollnagel E, *et al*. Compliance with hospital accreditation and patient mortality: a Danish nationwide population-based study. *Int J Qual Health Care* 2015;27:165–74.
61. Falstie-Jensen AM, Nørgaard M, Hollnagel E, *et al*. Is compliance with hospital accreditation associated with length of stay and acute readmission? A Danish nationwide population-based study. *Int J Qual Health Care* 2015;27:451–8.
62. Bogh SB, Falstie-Jensen AM, Bartels P, *et al*. Accreditation and improvement in process quality of care: a nationwide study. *Int J Qual Health Care* 2015;27:336–43.
63. Mumford V, Greenfield D, Hogden A, *et al*. Disentangling quality and safety indicator data: a longitudinal, comparative study of hand hygiene compliance and accreditation outcomes in 96 Australian hospitals. *BMJ Open* 2014;4:e005284.
64. Shaw CD, Groene O, Botje D, *et al*. The effect of certification and accreditation on quality management in 4 clinical services in 73 European hospitals. *Int J Qual Health Care* 2014;26:100–7.
65. Merlo J, Viciana-Fernández FJ, Ramiro-Fariñas D; Bringing the individual back to small-area variation studies: a multilevel analysis of all-cause mortality in Andalusia, Spain. *Soc Sci Med* 2012;75:1477–87.
66. Merlo J, Chaix B, Ohlsson H, *et al*. A brief conceptual tutorial of multilevel analysis in social epidemiology: using measures of clustering in multilevel logistic regression to investigate contextual phenomena. *J Epidemiol Community Health* 2006;60:290–7.
67. Merlo J. Contextual influences on the individual life course: building a research framework for social epidemiology. *Psychosocial Intervention* 2011;20:109–18.