International Journal for Quality in Health Care, 2015, 1–11
doi: 10.1093/intqhc/mzv092
Article

Article

# Community-acquired pneumonia (CAP) hospitalizations and deaths: is there a role for quality improvement through inter-hospital comparisons?

**W. AELVOET[1,2], N. TERRYN[1], A. BLOMMAERT[3], G. MOLENBERGHS[4], N. HENS[3,5], F. DE SMET[6,7], M. CALLENS[6], and P. BEUTELS[3,8]**

[1]Federal Public Service (FPS) Health, Food Chain Safety and Environment, Eurostation Bloc II—First Floor—01D327, Place Victor Horta 40 bte 10, B-1060 Brussels, Belgium, [2]Vrije Universiteit Brussel, Faculteit Geneeskunde en Farmacie, Brussels, Belgium, [3]Centre for Health Economics Research and Modeling Infectious Diseases (CHERMID), Vaccine and Infectious Disease Institute (WHO Collaborating Centre), University of Antwerp, Antwerp, Belgium, [4]Interuniversity Institute for Biostatistics and Statistical Bioinformatics (I-BioStat), Universiteit Hasselt and KU Leuven, Belgium, [5]Interuniversity Institute for Biostatistics and statistical Bioinformatics (I-BIOSTAT), Hasselt University, [6]National Alliance of Christian Mutualities, Brussels, Belgium, [7]Department of Public Health and Primary Care, Occupational, Environmental and Insurance Medicine, KU Leuven, Louvain, Belgium, and [8]School of Public Health and Community Medicine, The University of New South Wales, Sydney, Australia

Address reprints requests to: W. Aelvoet, Tel: +32-2-524-86-42; Fax: +32-2-524-86-98; E-mail: willem.aelvoet@health.fgov.be

Accepted 12 October 2015

## Abstract

**Objective:** To assess between-hospital variations in standardized in-hospital mortality ratios of community-acquired pneumonia (CAP), and identify possible leads for quality improvement.
**Design:** We used an administrative database to estimate standardized in-hospital mortality ratios for 111 Belgian hospitals, by carrying out a set of hierarchical logistic regression models, intended to disentangle therapeutic attitudes and biases. To facilitate the detection of false-negative/positive results, we added an inconclusive zone to the funnel plots, derived from the results of the study. Data quality was validated by comparison with (i) alternative data from the largest Belgian Sickness Fund, (ii) published German hospital data and (iii) the results of an on-site audit.
**Setting:** All Belgian hospital discharge records from 2004 to 2007.
**Study participants:** A total of 111 776 adult patients were admitted for CAP.
**Main outcome measure:** Risk-adjusted standardized in-hospital mortality ratios.
**Results:** Out of the 111 hospitals, we identified five and six outlying hospitals, with standardized mortality ratios of CAP consistently on the extremes of the distribution, as providing possibly better or worse care, respectively, and 18 other hospitals as having possible quality weaknesses/strengths. At the individuals' level of the analysis, adjusted odds ratios showed the paramount importance of old age, comorbidity and mechanical ventilation. The data compared well with the different validation sources.
**Conclusions:** Despite the limitations inherent to administrative data, it seemed possible to establish inter-hospital differences in standardized in-hospital mortality ratios of CAP and to identify leads for quality improvement. Monitoring is needed to assess progress in quality.

**Key words:** benchmarking, statistical methods, standardized mortality ratio, funnel plot, health services research

## Introduction

Inter-hospital comparisons of community-acquired pneumonia (CAP) standardized in-hospital mortality ratios (CAP-SMRs) may lead to an improved understanding of contextual influences on CAP, one of the leading causes of hospital admission, social and economic costs, and death throughout the world [1]. However, this type of comparison requires sufficiently reliable data, which can be challenging if these data serve multiple purposes (e.g. both reimbursement and quality assurance). Inadequate risk-adjustment and creep pose threats to data reliability such that widely used proprietary risk-adjustment may yield erroneous conclusions. Unfortunately, recommended severity scores [2, 3], laboratory data and physiologic information [4] are often not recorded in administrative databases, such as the ones we used here. Despite these imperfections, comparative information derived from administrative data is frequently put forward as a basis for quality improvement [5].

In an effort to encourage the hospital system to assume responsibility, the Belgian Ministry of Public Health decided to foster initiatives of quality improvement. To this end, a limited set of indicators was selected from the AHRQ Inpatient Quality Indicators, including the CAP-SMR [6]. We aimed, by establishing the existence of inter-hospital differences in CAP-SMR, (i) to evaluate to which extent Belgian discharge records allow the assessment of quality of care in the field of CAP, and (ii) to identify starting points for improvement.

## Methods

### Data source

Belgian hospitals are required to register discharge data, stored in the so-called the Minimal Clinical Data (MCD) database. It includes an unbounded number of ICD-9-CM coded diagnoses and procedures for each admission, which allows computing the Charlson's comorbidity index (CCI) (see Appendix 1: Charlson's comorbidity index, D'Hoore implementation) [7]. However, results of laboratory investigations, technical examinations such as X-rays, or patients' socio-economic status (SES) are not included. Moreover, the ICD-9-CM classification provides only limited information about severity of illness. The notion of 'intensity of care', based on the registration of invasive mechanical ventilation (IMV), non-invasive mechanical ventilation (NMV) and, otherwise, basic care (see Appendix 2: ICD-9 codes), allowed us to a certain extent to fill this gap [2, 8].

### Validity of the MCD data

A complementary Belgian data source, the Carenet hospitalization database (see Appendix 3: The Carenet hospitalization database), operating independently from the MCD database, was used to investigate the validity of the MCD data. Apart from a patient's age, gender and survival, it provides hospitalization data including primary and secondary diagnoses, patient and hospital identifiers and time and date of hospitalization. Carnet enabled us to compare between both registries the in-hospital 30-day mortality rates of hospitalized CAP globally and by age classes.

In addition, we compared the MCD's age distribution, age-specific incidence and the proportion of patients, admitted to hospital for CAP and who died during the follow-up period, with previously published German hospital data [2]. These German data had been collected according to a predefined quality report sheet as part of a nationwide mandatory performance measurement program.

Finally, we compared the MCD data with the results of audits, carried out by public service physicians, who compared the registered diagnostic codes with the original medical files, applying reference coding rules.

### Definition of the study population

In the MCD database, all admissions ($n = 146\,857$) having CAP [6] as principal diagnosis in the years 2004–07 (see Appendix1: ICD-9-CM codes) were selected. Records without information regarding vital status at discharge ($n = 77$), or concerning ages <18 years ($n = 37\,044$) or pregnant women ($N = 127$) or transfers to another hospital ($n = 2102$) were excluded. Thus, we retained 107 507 CAP patients. Striving for completeness rather than strictly applying the coding principles, we also included records with acute respiratory failure (ARF) as principal and CAP as secondary diagnosis ($n = 4269$), ending up with a potential study population of 111 776 observations, across 128 hospitals.

In the absence of personal identifiers, incidences were estimated excluding stays of patients transferred to another hospital ($n = 769$).

Concentrating on an inter-hospital comparison and in order to obtain statistical stability, we further excluded 17 hospitals registering fewer than 80 observations during the period 2004–07 ($n = 141$ stays) or fatalities in patients with an LOS < 3 days ($n = 2665$). The latter are highly dependent on the clinical status of the patient at presentation, whereas late mortality seems to be associated more closely with clinical management factors [9]. This way we obtained an *inter-hospital* study population of 108 213 cases admitted to 111 hospitals.

### Statistical methods

Since data of neither out-of-hospital cases nor out-of-hospital fatalities were available, our outcome of interest was the CAP-SMR. This standardized mortality ratio, defined as 100 times the ratio of the *observed* deaths ($O$) to the *expected* deaths ($E$), was constructed to identify both high- and low-performance quality outliers [10]. The expected deaths are the counterfactual, unobservable mortality experience, estimated from a hierarchical model, commonly applied in the field of hospital performance [10–12].

Hospitals with an SMR < 100 and a confidence interval not including 100 are considered high-performance outliers. Conversely, hospitals with an SMR > 100 and a confidence interval not including 100 are considered low-performance outliers. We constructed a hierarchical model, in this case, a mixed-effects multiple logistic regression model, with hospitals as random intercepts, accounting for within-hospital correlations [13]:

$$Y_{ij} \sim \text{Bernoulli}(p_{ij}),$$
$$\text{logit}(p_{ij}) = X_{ij}^T\beta + b_i,$$
$$b_i \sim N(0, \sigma^2),$$

with $p_{ij}$ the probability that patient $j$ within hospital $i$ dies, $\beta$ the vector of regression coefficients for $X_{ij}$, the matrix of risk-adjustment variables for the $j$th patient at the $i$th hospital. The model intercept is given by $\alpha$ and $b_i$ is the hospital-specific random intercepts$_i$, usually taken to be normally distributed with mean zero and standard deviation to be estimated.

For each hospital $i$, we calculated both the observed ($O_i$) and expected ($E_i$) number of fatalities:

$$O_i = \sum_{j=1}^{n_i} Y_{ij} \quad \text{and} \quad E_i = \sum_{j=1}^{n_i} \hat{p}_{ij},$$

with $\hat{p}_{ij} = \text{logit}^{-1}(X_{ij}^T\hat{\beta})$, the estimate of $\text{logit}^{-1}(X_{ij}^T\beta) = E(Y_{ij}|X_{ij}^T, b_i = 0)$, $\hat{\beta}$ is the vector of fitted regression coefficients and $n_i$ the number of CAP hospitalizations in hospital $i$. Since the random intercept component of the hierarchical model accounts for between-hospital variability, only the fixed-effects coefficients $\hat{\beta}$ were used to calculate the expected deaths, thus removing the impact of individual hospital quality on the expected mortality. In other words, the probability of death for a patient treated at an 'average quality' hospital (with random intercept $b_i = 0$) [13] is estimated.

From the observed and expected number of fatalities, we calculated the standardized mortality ratio ($SMR_i$) for hospital $i$ as $SMR_i = O_i/E_i \times 100$, that should be interpreted as a percentage deviation from the hypothetical average hospital. As we are modeling the ratio of the number of fatalities over the number of cases and since the criteria for approaching this binomial distribution by a Poisson distribution were not met, a hierarchical logistic regression was chosen. The SMRs are graphically represented using funnel plots with control limits based on the 99.8- and 95-percentiles of the exact binomial distribution as described in Spiegelhalter [14].

In our mixed-effects multiple logistic regression model, a generalized linear mixed model (GLMM), parameter estimation was carried out using an integral approximation method (the Gauss–Hermite quadrature with specification of 50 quadrature points), that numerically evaluates the marginal log-likelihood of the model. The advantage of this method is that it manipulates the likelihood and all of its derived quantities with high precision. By choosing the number of quadrature points sufficiently highly, arbitrary precision can be reached. As a consequence, trustworthy point estimates, standard errors, confidence intervals and likelihood ratio test statistics result. Thus, we were able to reject the hypothesis test of no random effects ($P < 0.0001$) implying significant inter-hospital differences.

To assess the need to include interactions, we fitted a series of models starting from a main effects model (M1) and successively introducing interaction variables. Although statistically significant interaction terms were present, we retained the main effects model as the 'initial model', starting point of our sensitivity analysis. Our choice was guided by the ease of its interpretation and by the modest improvement by adding interaction terms in the modeling (see Supplementary file: Modeling CAP mortality). The scaled Pearson statistic for the conditional distribution (0.95) did not suggest a problem of over- or under-dispersion.

Recognizing the limitations of administrative data regarding selection bias, inadequate risk-adjustment and other biases indirectly arising from differences in medical practices and in attitudes, including (i) whether or not providing IMV/NMV, (ii) early discharging patients (especially of terminal patients), (iii) artificially increasing of the case-mix and (iv) withholding optimal care in the elderly, whether or not by request of patient or family, we tried to take these biases into account by carrying out a sensitivity analysis. Therefore, we constructed two models, excluding from the analysis, respectively, patients discharged during the first week (as a proxy for early discharge) and patients aged over 79 years. Subsequently, we fitted two additional models wherein no adjustment was made, respectively, for intensity of care and comorbidities.

Moreover, random intercepts are believed to remove some of the biases typical of hospital-based registries as a result of differences including case-definition, case-ascertainment, coding and SES [15]. Also, the choice of analyzing a cause-specific SMR is considered more-reliable than that of a hospital-wide one [16].

Alternative statistical approaches in the domain of CAP-related mortality, e.g. risk prediction models [3] and data driven rules to predict mortality [4], could not be applied due to the absence of recommended severity scores, laboratory data and physiologic information in our data.

### Funnel plots

We generated scatter plots of the hospitals' SMR, against the number of admitted patients (the 'volume'). The vertical and horizontal axes of the plot represent the values of the SMR and of the volume, respectively, with 99.8 and 95% control limits. Taking the possible confounding effects of unmeasured or mismeasured variables into account [17], we delimited an *inconclusive zone* (shaded on the graphs), including all SMRs finding themselves between 33% above and 25% below the reference SMR of 100. It was intended to facilitate the detection of false-negative results in small hospitals and false-positive results in large hospitals. The fitted model and the event proportion of the 'average Belgian hospital' are displayed in footnotes.

Given these limits, five performance categories are usually [14] defined, ranging from 'action' (above or equal to the upper bound of the 99.8% limits); over 'alarm' (above or equal to the upper bound of the 95% limits, but lower than the upper bound of the 99.8% limits), 'normal', 'good' (below or equal to the lower bound of the 95% limits, but higher than the lower bound of the 99.8% limits); to 'excellent' (below or equal to the lower bound of the 99.8% limits).

However, to interpret the results of the sensitivity analysis, we defined a hospital's performance as 'To be assessed' when the performance category changed by more than one contiguous category. Otherwise, if the performance category equaled 'Excellent' or 'Action' in one of the analyses, we labeled the hospital 'Possibly better' or 'Possibly worse' performing, respectively. Hospitals belonging to the categories 'Action' or 'Excellent' are numbered in the figures. A further, 'To be assessed' category, consisted of hospitals finding themselves in the sensitivity analysis at least once outside the inconclusive zone. A final, 'Normally performing', category encompassed the remaining hospitals. All analyses were carried out in SAS 9.2. The program code used to create the funnel plots is freely available from the authors.

The study being (i) of a retrospective, non-interventional type and (ii) anonymous with respect to patients, no approval by an ethics committee is required under the Belgian law.

## Results

### Patient and hospital characteristics

The proportion of patients who died during the follow-up period in the MCD *inter-hospital* study population amounted to 12.13% (95% CI: 11.93–12.32) overall, 12.88% (95% CI: 12.62–13.15) in males and 11.15% (95% CI: 10.87–11.44) in females. In the case of ARF, this proportion amounted to 37.62% (95% CI: 35.67–39.61) in males and 34.63% (95% CI: 32.32–37.01) in females.

In both sexes, we observed the highest admission numbers (more than 50%) in the age window of 70–89 years and increasing mortality ratios with increasing age (Table 1). Conversely, although higher in deceased patients of both sexes, IMV markedly decreased with increasing age: from ~40% in age-class 40–49 years to 20% in age-class 80–89 years.

The volume of patients admitted varied hugely between hospitals.

Adjusted odds ratios (Table 2) showed the paramount importance of old age, multiple comorbidities and IMV. Small volume, admission from another hospital or from a rest and nursing home, and, to a

**Table 1** Distribution of patient, stay and hospital characteristics for the *inter-hospital* study population

| | *Inter-hospital* study population ($n$ = 108 213) | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Males | | | | | | Females | | | | | |
| | Dec. | Cases | OR | 95% CI | | Col% | Dec. | Cases | OR | 95% CI | | Col% |
| Intensity of care | | | | | | | | | | | | |
| IMV | 2357 | 4151 | 12.69 | 11.85 | 13.59 | 7 | 1314 | 2524 | 11.52 | 10.57 | 12.55 | 5 |
| NMV | 871 | 75 30 | 1.26 | 1.17 | 1.36 | 12 | 561 | 5302 | 1.26 | 1.14 | 1.38 | 11 |
| Basic care | 4614 | 49 193 | 1 | | | 81 | 3404 | 39 513 | 1 | | | 83 |
| ARF | | | | | | | | | | | | |
| Yes | 895 | 2379 | 4.47 | 4.10 | 4.88 | 4 | 562 | 1623 | 4.47 | 4.10 | 4.88 | 3 |
| No | 6947 | 58 495 | 1 | | | 96 | 4717 | 45 716 | 1 | | | 97 |
| CCI | | | | | | | | | | | | |
| 0 | 913 | 16 932 | 0.15 | 0.13 | 0.16 | 28 | 927 | 17 010 | 0.17 | 0.14 | 0.19 | 36 |
| 1 | 1195 | 13 298 | 0.25 | 0.23 | 0.28 | 22 | 852 | 10 092 | 0.27 | 0.23 | 0.31 | 21 |
| 2 | 1232 | 9462 | 0.38 | 0.35 | 0.42 | 16 | 978 | 7527 | 0.43 | 0.38 | 0.50 | 16 |
| 3 | 1154 | 7201 | 0.49 | 0.44 | 0.54 | 12 | 787 | 4821 | 0.57 | 0.49 | 0.65 | 10 |
| 4 | 855 | 4222 | 0.65 | 0.58 | 0.72 | 7 | 486 | 2823 | 0.60 | 0.51 | 0.71 | 6 |
| 5 | 533 | 2439 | 0.71 | 0.63 | 0.81 | 4 | 348 | 1531 | 0.85 | 0.72 | 1.01 | 3 |
| 6 | 435 | 1628 | 0.93 | 0.81 | 1.06 | 3 | 227 | 881 | 1.01 | 0.83 | 1.22 | 2 |
| 7 | 250 | 876 | 1.02 | 0.86 | 1.20 | 1 | 134 | 488 | 1.10 | 0.87 | 1.39 | 1 |
| 8 | 242 | 1062 | 0.75 | 0.64 | 0.89 | 2 | 143 | 569 | 0.97 | 0.78 | 1.22 | 1 |
| 9 | 192 | 771 | 0.84 | 0.70 | 1.01 | 1 | 66 | 307 | 0.79 | 0.59 | 1.07 | 1 |
| 10+ | 841 | 2983 | 1 | | | 5 | 331 | 1290 | 1 | | | 3 |
| Age class | | | | | | | | | | | | |
| <40 years | 49 | 4967 | 1 | | | 8 | 40 | 4767 | 1 | | | 10 |
| 40–49 years | 112 | 3912 | 2.96 | 2.11 | 4.15 | 6 | 70 | 3080 | 2.75 | 1.86 | 4.06 | 7 |
| 50–59 years | 357 | 6144 | 6.19 | 4.58 | 8.37 | 10 | 174 | 4293 | 4.99 | 3.53 | 7.06 | 9 |
| 60–69 years | 933 | 9798 | 10.56 | 7.91 | 14.11 | 16 | 331 | 5417 | 7.69 | 5.53 | 10.70 | 11 |
| 70–79 years | 2472 | 17 358 | 16.67 | 12.54 | 22.15 | 29 | 1158 | 10 302 | 14.97 | 10.90 | 20.55 | 22 |
| 80–89 years | 3079 | 15 430 | 25.02 | 18.83 | 33.24 | 25 | 2311 | 14 093 | 23.18 | 16.93 | 31.74 | 30 |
| 90–99 years | 824 | 3214 | 34.61 | 25.83 | 46.35 | 5 | 1147 | 5231 | 33.19 | 24.15 | 45.62 | 11 |
| 100 years+ | 16 | 51 | 45.88 | 23.83 | 88.33 | | 48 | 156 | 52.52 | 33.13 | 83.28 | |
| Admitted from | | | | | | | | | | | | |
| Hospital | 345 | 1531 | 2.25 | 1.99 | 2.55 | 3 | 224 | 1002 | 3.02 | 2.59 | 3.52 | 2 |
| Rest and nursing home | 1410 | 6059 | 2.35 | 2.20 | 2.51 | 10 | 1745 | 8442 | 2.73 | 2.56 | 2.91 | 18 |
| Other | 154 | 1360 | 0.99 | 0.84 | 1.17 | 2 | 74 | 746 | 1.15 | 0.91 | 1.47 | 2 |
| Home | 5933 | 51 924 | 1 | | | 85 | 3236 | 37 149 | 1 | | | 78 |
| Weekend admission | | | | | | | | | | | | |
| Yes | 1708 | 13 126 | 1.01 | 0.96 | 1.07 | 22 | 1213 | 10 151 | 1.11 | 1.03 | 1.18 | 21 |
| No | 6134 | 47 748 | 1 | | | 78 | 4066 | 37 188 | 1 | | | 79 |
| LOS (in days) | | | | | | | | | | | | |
| 0 | 0 | 941 | n/a | n/a | n/a | 5 | 0 | 666 | n/a | n/a | n/a | 5 |
| 1 | 0 | 762 | n/a | n/a | n/a | 4 | 0 | 543 | n/a | n/a | n/a | 4 |
| 2 | 0 | 1089 | n/a | n/a | n/a | 5 | 0 | 731 | n/a | n/a | n/a | 5 |
| 3 | 501 | 2338 | 3.25 | 2.80 | 3.77 | 12 | 346 | 1707 | 3.06 | 2.57 | 3.64 | 12 |
| 4 | 386 | 3152 | 1.66 | 1.43 | 1.94 | 16 | 364 | 2343 | 2.21 | 1.87 | 2.63 | 16 |
| 5 | 387 | 3423 | 1.52 | 1.30 | 1.77 | 17 | 284 | 2513 | 1.53 | 1.28 | 1.83 | 18 |
| 6 | 380 | 3797 | 1.32 | 1.14 | 1.54 | 19 | 242 | 2589 | 1.24 | 1.03 | 1.49 | 18 |
| 7 | 340 | 4391 | 1 | | | 22 | 250 | 3258 | 1 | | | 23 |
| Subtotal | 1994 | 19 893 | | | | | 1486 | 14 350 | | | | |
| LOS (in weeks) | | | | | | | | | | | | |
| 0 | 1994 | 19 893 | 0.30 | 0.28 | 0.32 | 33 | 1486 | 14 350 | 0.45 | 0.41 | 0.49 | 30 |
| 1 | 1937 | 21 926 | 0.26 | 0.24 | 0.28 | 36 | 1336 | 16 911 | 0.33 | 0.31 | 0.36 | 36 |
| 2 | 1337 | 8743 | 0.49 | 0.45 | 0.53 | 14 | 878 | 7573 | 0.51 | 0.46 | 0.56 | 16 |
| 3 | 901 | 4130 | 0.75 | 0.69 | 0.83 | 7 | 564 | 3544 | 0.74 | 0.66 | 0.82 | 7 |
| 4+ | 1673 | 6182 | 1 | | | 10 | 1015 | 4961 | 1 | | | 10 |
| Teaching hospital | | | | | | | | | | | | |
| Yes | 465 | 4123 | 1 | | | 7 | 317 | 2960 | 1 | | | 6 |
| No | 7377 | 56 751 | 1.18 | 1.06 | 1.30 | 93 | 4962 | 44 379 | 1.05 | 0.93 | 1.18 | 94 |
| All patients | | | | | | | | | | | | |
| Males | 7842 | 60 874 | 1.18 | 1.14 | 1.22 | | | | | | | |
| Females | 5279 | 47 339 | | | | | | | | | | |

**Table continued**

**Table 1** Continued

| Hospital volume in quintiles | Nbr of hospitals | Minimum | Maximum | Range |
|---|---|---|---|---|
| Q1 | 22 | 171 | 510 | 339 |
| Q2 | 21 | 518 | 696 | 178 |
| Q3 | 23 | 697 | 928 | 231 |
| Q4 | 22 | 938 | 1327 | 389 |
| Q5 | 23 | 1359 | 3651 | 2292 |

Dec., deceased; OR, odds ratio; 95% CI, 95% confidence interval, rounded to the nearest unit; Col%, column percent; ARF, acute respiratory failure as principal diagnosis; IMV, invasive mechanical ventilation; n/a, not applicable; NMV, non-invasive mechanical ventilation; Volume, number of admissions with CAP as principal diagnosis.

**Table 2** Adjusted OR and 95% CI (i) of mortality determinants, and (ii) fixed-effects part of the inter-hospital comparison, initial model

| | Determinants of inter-hospital comparison | | | Fixed effects of inter-hospital comparison | | |
|---|---|---|---|---|---|---|
| | OR | 95% CI | | OR | 95% CI | |
| Age class | | | | | | |
| 100+ versus <40 years | 54.68 | 37.57 | 79.58 | 42.58 | 29.34 | 61.79 |
| 90–99 versus <40 years | 32.32 | 25.87 | 40.37 | 24.76 | 19.81 | 30.95 |
| 80–89 versus <40 years | 19.59 | 15.76 | 24.34 | 15.10 | 12.14 | 18.78 |
| 70–79 versus <40 years | 10.61 | 8.54 | 13.19 | 8.49 | 6.82 | 10.56 |
| 60–69 versus <40 years | 5.7 | 4.56 | 7.12 | 4.74 | 3.79 | 5.93 |
| 50–59 versus <40 years | 3.44 | 2.73 | 4.35 | 2.99 | 2.36 | 3.78 |
| 40–49 versus <40 years | 1.97 | 1.51 | 2.56 | 1.81 | 1.39 | 2.36 |
| Gender | | | | | | |
| Males versus females | 1.25 | 1.2 | 1.31 | 1.27 | 1.22 | 1.33 |
| CCI | | | | | | |
| CCI 10 versus 0 | 6.62 | 6.03 | 7.28 | 6.10 | 5.56 | 6.70 |
| CCI 9 versus 0 | 3.28 | 2.76 | 3.89 | 3.08 | 2.61 | 3.65 |
| CCI 8 versus 0 | 3.53 | 3.07 | 4.07 | 3.34 | 2.90 | 3.84 |
| CCI 7 versus 0 | 3.19 | 2.75 | 3.7 | 3.00 | 2.60 | 3.47 |
| CCI 6 versus 0 | 3.07 | 2.74 | 3.45 | 2.84 | 2.53 | 3.18 |
| CCI 5 versus 0 | 2.34 | 2.12 | 2.59 | 2.15 | 1.95 | 2.38 |
| CCI 4 versus 0 | 2.08 | 1.91 | 2.27 | 1.94 | 1.79 | 2.11 |
| CCI 3 versus 0 | 1.82 | 1.69 | 1.97 | 1.69 | 1.57 | 1.82 |
| CCI 2 versus 0 | 1.56 | 1.45 | 1.68 | 1.47 | 1.37 | 1.58 |
| CCI 1 versus 0 | 1.15 | 1.07 | 1.23 | 1.08 | 1.01 | 1.16 |
| Admission | | | | | | |
| Hosp. versus home | 1.92 | 1.71 | 2.16 | 1.91 | 1.70 | 2.14 |
| R and N home versus home | 1.79 | 1.70 | 1.89 | 1.72 | 1.64 | 1.82 |
| Other place versus home | 1.14 | 0.97 | 1.34 | 1.15 | 0.98 | 1.35 |
| Intensity of care | | | | | | |
| IMV versus basic | 14.88 | 13.94 | 15.88 | 14.31 | 13.45 | 15.22 |
| NMV versus basic | 1.29 | 1.20 | 1.39 | 1.24 | 1.16 | 1.33 |
| LOS (weeks) | | | | | | |
| Week '0 versus 4+' | 1.48 | 1.38 | 1.58 | | | |
| Week '1 versus 4+' | 0.61 | 0.57 | 0.65 | | | |
| Week '2 versus 4+' | 0.71 | 0.66 | 0.76 | | | |
| Week '3 versus 4+' | 0.92 | 0.85 | 1 | | | |
| Volume | | | | | | |
| Quintile '1 versus 5' | 1.04 | 0.88 | 1.24 | | | |
| Quintile '2 versus 5' | 1.00 | 0.85 | 1.18 | | | |
| Quintile '3 versus 5' | 1.12 | 0.95 | 1.31 | | | |
| Quintile '4 versus 5' | 1.17 | 0.99 | 1.38 | | | |
| WE admission | | | | | | |
| Yes versus no | 1.03 | 0.98 | 1.09 | | | |
| Teaching | | | | | | |
| Yes versus no | 0.78 | 0.62 | 0.98 | | | |

CCI, Charlson's comorbidity index; R and N home, rest and nursing home; IMV/NMV: invasive/non-invasive mechanical ventilation; OR, odds ratio; 95% CI, 95% confidence interval.

lesser extent, the male sex, weekend admissions and admissions in non-teaching hospitals showed higher mortality ratios. LOS displays a J-shaped relationship with mortality.

## Validity of the data

Comparing MCD's estimate of the in-hospital 30-day rate with the Carenet data, we obtained quite similar overall and age-specific

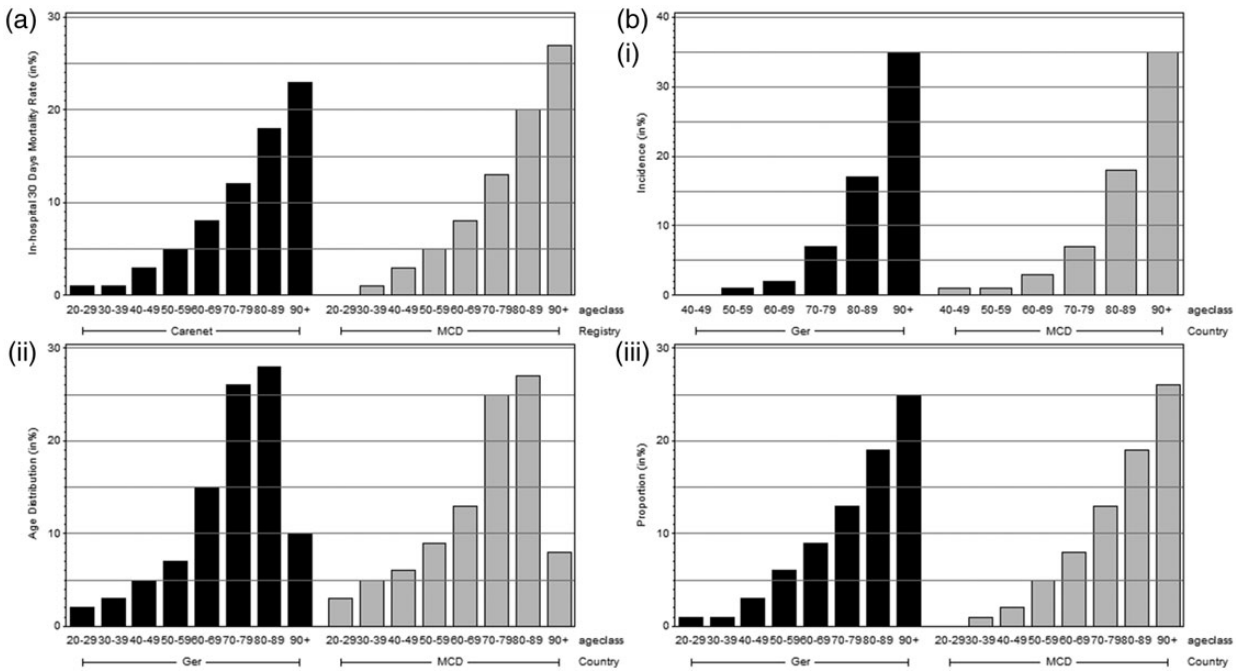


**Figure 1** Validity of the data: (a) comparison of the in-hospital 30 days mortality rate (in %) between the Carenet and MCD registries (upper left panel); (b) comparison of in-hospital CAP admissions between MCD and Germany (Ge) in (i) incidence (in %) (upper right panel), (ii) in age-class distribution (in %) (lower left panel) and (iii) proportion of patients who died during the follow-up period (in %) (lower right panel).

figures (Fig. 1). In addition, striking similarities between the MCD and published German data were observed regarding in-hospital age distribution, age-specific incidence rates, and age-specific proportions of patients who died during the follow-up period (Fig. 1).

Finally, an on-site audit on 4093 medical files concluded that the auditor agreed in the large majority of cases (82%) with the coded hospital diagnosis. In another 14% of cases, the coded hospital diagnosis seemed still to deserve a 'CAP likely' type of conclusion, whereas in 4% of the cases, the auditor assigned a code to the principal diagnosis, corresponding to another, clearly not CAP-related pathology. The type of conclusion, however, considerably varied across hospitals.

### Inter-hospital comparison

According to our definitions, 5 hospitals were classified as 'Possibly better performing', 7 as 'Possibly worse performing', 18 as 'To be assessed' and 81 as 'Normally performing' (Table 3). To somewhat facilitate the interpretation, we also provided the registered intensity of care, by category, as well as the corresponding national percentages. The five hospitals of supposedly 'better' quality found themselves in the sensitivity analyses most often below the inconclusive zone (Fig. 2), suggesting a real survival excess. Six hospitals, labeled as 'possibly worse' performing, presented the opposite image, suggesting a real mortality excess. No single potential starting point for improvement became apparent, with the exception of underuse of IMV/NMV in Hospital 37 combined with a lower SMR in the intensity-of-care-excluded-analysis. A seventh hospital (number 62) deserves a more cautious interpretation since it exclusively treats cancer patients. This may largely explain its extreme position in the basic analysis, as well as the huge SMR in the CCI-excluded-analysis and the less intensive care provided.

## Discussion

In the context of the growing literature [18–21] refining the concept of continuous quality improvement [22], calling for the development of suitable methods and setting standards to stimulate and conduct a quality of care improvement study, the comparison of inter-hospital mortality rates has drawn attention for decades and may give the initial impetus to the conduct of a national clinical audit [12, 21, 23, 24]. Our study demonstrated methods that can be applied to administrative databases in order to unveil considerable inter-hospital differences in CAP-SMR. Our results suggested true differences in quality of care deserving further investigation.

Adding an inconclusive zone to the funnel plot may not only help reveal the presence of both false-positive and false-negative outliers, but, more importantly, it takes into account the magnitude of the SMR's departure. The choice of the limits was inspired by the literature [17] and may find some support in the divergent male versus female ORs in our and the German study [2].

The sensitivity analysis on the other hand allowed us to a certain extent (i) to disentangle therapeutic attitudes from quality of care, and (ii) to remove some of the biases due to inadequate risk-adjustment and to gaming. For instance, a model not adjusted for mechanical ventilation—a resource-intensive procedure shown to be recorded most accurately [25]—may induce a change of quality toward a lesser category than that of the adjusted model. Yet this procedure may be a sign of good quality, in line with the hospital's case-mix and required intensity of care, thus suggesting good quality, a finding that may be confirmed by the registered intensity of care. Similarly, a relatively high proportion of patients with LOS <8 days may be indicative of gaming, by discharging patients who are past saving [26, 27]. Therefore, the exclusion of such patients can reduce bias into the in-hospital mortality comparisons due to differential follow-up [28]. Likewise, through the exclusion of patients aged 80+ or by withdrawing comorbidity

**Table 3** Results of the sensitivity analysis of hospitals, deserving particular attention regarding the quality of care provided: SMR, funnel plot (Plot) and inconclusive zone (Zone), for each of the analyzed models

| Id | Model | | | | | | | | | | | | | | Volume | Quality | Registered intensity of care (%) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Initial model | | | LOSses<8 excluded | | | Patients 80 years+ excluded | | | Intensity of care excluded | | | CCI excl. | | | | | | |
| | SMR | Plot | Zone | SMR | Plot | Zone | SMR | Plot | Zone | SMR | Plot | Zone | SMR | Plot | Zone | | | I | NI | B |
| 16 | 71 | E | Y | 73 | E | Y | 75 | E | N | 74 | E | Y | 82 | E | N | 5 | Better | 8 | 23 | 69 |
| 21 | 72 | E | Y | 76 | G | N | 75 | G | Y | 71 | E | Y | 72 | E | Y | 4 | Better | 3 | 69 | 28 |
| 33 | 61 | E | Y | 62 | E | Y | 68 | G | Y | 57 | E | Y | 63 | E | Y | 4 | Better | 5 | 22 | 73 |
| 34 | 78 | E | N | 85 | G | N | 75 | G | Y | 75 | E | Y | 77 | E | N | 5 | Better | 5 | 29 | 66 |
| 105 | 71 | E | Y | 68 | E | Y | 63 | E | Y | 63 | E | Y | 75 | E | Y | 5 | Better | 4 | 2 | 94 |
| 7 | 126 | Ac | N | 119 | Al | N | 139 | Ac | Y | 134 | Ac | Y | 130 | Ac | N | 4 | Less | 9 | 4 | 87 |
| 15 | 137 | Ac | Y | 150 | Ac | Y | 139 | Al | Y | 151 | Ac | Y | 127 | Al | N | 3 | Less | 7 | 21 | 72 |
| 37 | 211 | Ac | Y | 206 | Ac | Y | 238 | Ac | Y | 143 | Al | Y | 205 | Ac | Y | 1 | Less | 0 | 0 | 100 |
| 52 | 128 | Ac | N | 124 | Al | N | 130 | Al | N | 148 | Ac | Y | 139 | Ac | Y | 4 | Less | 11 | 1 | 89 |
| 62 | 215 | Ac | Y | 223 | Ac | Y | 230 | Ac | Y | 152 | Al | Y | 365 | Ac | Y | 1 | Less | 1 | 4 | 95 |
| 73 | 139 | Ac | Y | 137 | Ac | Y | 138 | Al | Y | 133 | Ac | Y | 152 | Ac | Y | 4 | Less | 7 | 3 | 91 |
| 88 | 130 | Ac | N | 131 | Al | N | 145 | Al | Y | 127 | Al | N | 121 | Al | N | 3 | Less | 5 | 13 | 82 |
| 4 | 125 | Al | N | 117 | N | N | 118 | N | N | 164 | Ac | Y | 127 | Ac | N | 4 | Assess | 15 | 3 | 82 |
| 9 | 95 | N | N | 100 | N | N | 129 | Al | N | 83 | G | N | 96 | N | N | 4 | Assess | 2 | 0 | 97 |
| 10 | 61 | G | Y | 69 | N | Y | 82 | N | N | 56 | E | Y | 54 | E | Y | 1 | Assess | 3 | 16 | 82 |
| 14 | 118 | Al | N | 105 | N | N | 117 | N | N | 133 | Ac | Y | 124 | Ac | N | 5 | Assess | 10 | 30 | 60 |
| 19 | 137 | Ac | Y | 131 | Al | N | 118 | N | N | 153 | Ac | Y | 139 | Ac | Y | 2 | Assess | 9 | 4 | 87 |
| 23 | 117 | Al | N | 115 | N | N | 111 | N | N | 126 | Ac | N | 115 | Al | N | 5 | Assess | 8 | 1 | 91 |
| 29 | 66 | G | Y | 70 | N | Y | 70 | N | Y | 68 | N | Y | 72 | N | Y | 1 | Assess | 6 | 10 | 84 |
| 35 | 70 | E | Y | 77 | G | N | 64 | E | Y | 88 | N | N | 76 | E | N | 5 | Assess | 12 | 9 | 78 |
| 43 | 119 | Al | N | 117 | Al | N | 134 | Ac | Y | 108 | N | N | 114 | Al | N | 5 | Assess | 4 | 2 | 94 |
| 44 | 76 | E | N | 83 | G | N | 80 | N | N | 75 | E | N | 78 | E | N | 5 | Assess | 6 | 1 | 93 |
| 45 | 70 | E | Y | 85 | N | N | 69 | G | Y | 64 | E | Y | 68 | E | Y | 3 | Assess | 3 | 24 | 73 |
| 47 | 114 | N | N | 110 | N | N | 112 | N | N | 128 | Ac | N | 116 | N | N | 4 | Assess | 9 | 5 | 86 |
| 55 | 122 | Al | N | 120 | Al | N | 116 | N | N | 117 | Al | N | 124 | Ac | N | 4 | Assess | 5 | 5 | 90 |
| 71 | 121 | Ac | N | 114 | Al | N | 136 | Ac | Y | 107 | N | N | 123 | Ac | N | 5 | Assess | 4 | 0 | 96 |
| 81 | 114 | N | N | 114 | N | N | 101 | N | N | 129 | Ac | N | 125 | Al | N | 4 | Assess | 8 | 82 | 9 |
| 89 | 124 | Ac | N | 125 | Al | N | 119 | N | N | 111 | N | N | 123 | Ac | N | 5 | Assess | 3 | 0 | 97 |
| 95 | 127 | Al | N | 123 | N | N | 133 | N | N | 141 | Ac | Y | 122 | N | N | 1 | Assess | 9 | 8 | 83 |
| 106 | 119 | N | N | 117 | N | N | 129 | N | N | 145 | Ac | Y | 121 | N | N | 2 | Assess | 12 | 4 | 85 |
| Nat. | | | | | | | | | | | | | | | | | | 6 | 12 | 82 |

Id, anonymous hospital identifier; LOSses < 8 days, the model wherein observations with LOSses of <8 days are excluded; Patients 80 years+ excluded, the model wherein patients aged 80 years or more are excluded; Intensity of care excluded, the model not adjusted for intensity of care; CCI excluded, the model not adjusted for CCI; SMR, standardized mortality ratio, rounded to the nearest unit; P, conclusion based on control limits of the funnel plot (E, excellent; G, good; Al, alarm; Ac, action; N, normal); Zone: SMR outside the inconclusive zone (Y, yes/N, no); Volume, volume in terms of quintiles; Quality: better/less: possibly better/less performing hospital—Assess: to be assessed performance. Registered intensity of care (%), registered intensity of care (expressed in %) as carried out in the individual hospitals and nationally (displayed in the row 'Nat': 'National') (I, invasive mechanical ventilation; NI, non-invasive mechanical ventilation; B, basic care).

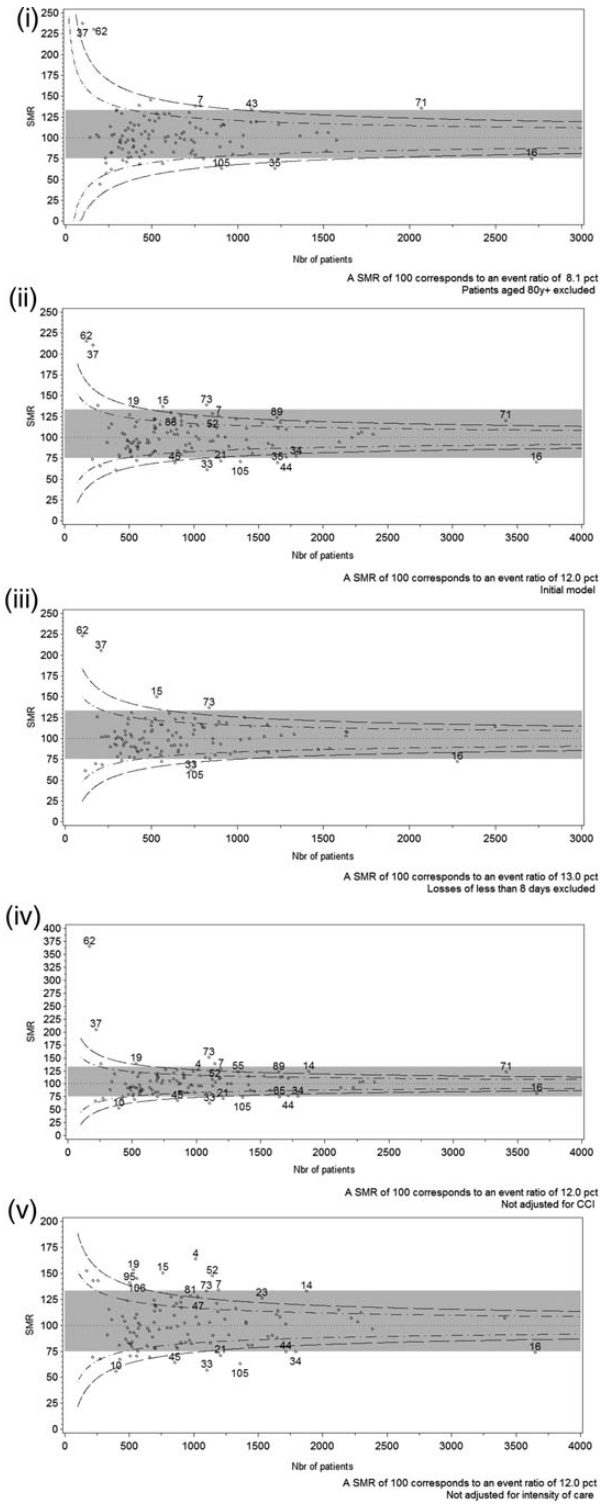


**Figure 2** Funnel plots based on the results of the sensitivity analysis: from top to bottom: (i) the initial model, (ii) the model in which observations with LOS < 8 days are excluded, (iii) the model in which patients aged 80 years or more are excluded, (iv) the model not adjusted for CCI, and (v) the model not adjusted for intensity of care. Inner and outer dashed lines are, respectively, 95 and 99.8% control limits, derived using the 'interpolated', exact binomial distribution. The inconclusive zone (shaded on the graphs) extends from 25% below to 33% above the reference SMR of 100 (dotted line). (Numbered) Circles represent (outlying outside the 99.8% control limits) hospitals.

from the modeling, an attempt was made to assess the possible effects on the SMRs of therapeutic attitudes as well as patient or next of kin wishes [2], respectively, or of up-coding phenomena.

Both techniques added value to the initial model of our inter-hospital analysis, that resulted in the identification of five and six outlying hospitals as providing possibly better or worse care, respectively. These 11 hospitals, identified as outliers, found themselves consistently on the extremes of the SMR distribution and often outside the inconclusive zone. Given the 99.8% control limits and although no over-dispersion was present, the number of 11 hospitals identified as 'out-of-control' is higher than the expected 0.2% risk of a false alarm [29], suggesting we are dealing with a number of truly outlying hospitals within this group. Approaching the problem in this way, we feel to have reckoned with (i) considerable biases due to inadequate risk-adjustment and to gaming, common in administrative data, and (ii) differences in therapeutic attitudes.

Furthermore, our approach disclosed possible quality weaknesses or strengths for some of the 18 hospitals, labeled 'To be assessed'. According to the funnel plot of the initial model (Fig. 2 and Table 3), only three (35, 44, 45) of them could be labeled 'excellent'. However, based on our pre-set definitions (see the Methods section), these hospitals fall in the 'to-be-assessed' category, notwithstanding their barely changing SMRs. This may be due to the influence of the sample size on the control limits by excluding observations (patients aged 80 years and more or LOS <8 days), to therapeutic attitudes regarding the provision of certain types of care to elderly patients (Hospital 44), to discharge practices (Hospital 45) or due to the removal of an adjustment (intensity of care in Hospital 35). For similar reasons, Hospital 10, to be labeled as 'Good' according to the funnel plot and finding itself below the inconclusive zone, is rated 'to-be-assessed'. Although within the inconclusive zone, two large-size hospitals (71 and 89) are labeled as deserving 'Action', which may be due to a suboptimal use of mechanical ventilation, suggested by the 'Intensity-of care-excluded-analysis'. Seven hospitals (4, 14, 19, 23, 43, 55 and 95) designated as 'Alarm' or 'Action' and four hospitals (9, 47, 81and 106) accredited as 'Normal' in the initial model received no clue for improvement from the sensitivity analysis. Hospital 29, labeled 'Normal' on four subanalyses, had four times an 'SMR below the inconclusive zone', suggesting better quality.

Since pneumonia care may be provided in an out-patient setting, selection biases [30] may occur and require a cautious interpretation of the in-hospital findings. However, the striking resemblance between the MCD and German Hospital data, the similarity between MCD and Carenet data concerning both the overall and the age-specific in-hospital mortality and the results of the audit are reassuring for the validity of the data regarding mortality in hospitalized CAP patients and its determinants.

In addition to the already discussed biases, we faced several study limitations, including the lack of laboratory results, of radiological and of clinical findings such as mental confusion and severity of illness [31, 32]. Although 'intensity of care' may perform well as a proxy for severity of illness [8], the completeness of its registration in our administrative data remains uncertain. The preceding encouraged us to label our inter-hospital results a 'screening', that has to be further investigated, rather than 'assessing' quality of care.

Adding a sensitivity analysis and introducing an inconclusive zone in the analysis may be considered strengths of our study. Also the observed adjusted mortality ORs according to age, comorbidity and invasive ventilator support are congruous with the literature [2, 30, 33]. The gender divide in favor of the females is rather small but in accordance with two sizable cohort studies [30, 34] albeit not with a third study [2]. Our finding of a doubled mortality risk in patients admitted

from a rest and nursing home, conceivably at risk of Healthcare Associated Pneumonia (HCAP), is in line with the literature [33, 35]. By selecting pneumonia as principal diagnosis, we avoided the inclusion of cases of nosocomial pneumonia that should be coded as secondary diagnoses. We further excluded short-term fatalities to avoid potential hospital-bias related to early unavoidable deaths [9].

Although, we did not find direct clues to assess specific departures from evidence-based practices, our sensitivity analysis tentatively indicated areas of possible betterment. In addition, the sizeable inter-hospital differences suggested real differences in quality of care. As a first step to quality improvement, monitoring of CAP-SMRs seems needed to assess whether this quality divide is fading away.

## Authors' contributions

W.A. conceived of the study, drafted the manuscript and participated in the design of the study and in the statistical analysis. N.T. carried out the extraction of the MCD data and participated in the analysis. A.B. carried out the extraction and the analysis of the Carenet data, wrote part of the methodology section and of the statistical methodology. G.M. and N.H. advised on the statistical methodology and data analysis. F.D.S. and M.C. acquired and interpreted the Carenet data. P.B. helped in the design of the study, interpreting the results and revising earlier drafts. All authors critically revised and approved the final version of the manuscript.

## Acknowledgements

We are greatly indebted to Drs E. Dewulf, A-M Lambot, D. Desantoine, N. Farhat, F. Proot, S. Van Malderen and K. Wymeersch, who carried out the audits. We also would like to thank professors W. Peetermans and P. Van Damme and Dr P. Reper for stimulating comments.

## Conflict of Interest statement

The author(s) declare that they have no competing interests relevant to the content of this paper. The authors of this article are responsible for its contents. No statement in this article should be construed as an official position of the Federal Service of Health, Food Chain Safety and Environment.

## Supplementary material

Supplementary material is available at *INTQHC* online.

## Funding

A.B.'s participation is funded by the University of Antwerp (UA)'s concerted research action number 23405 (BOF-GOA). N.H. holds the UA Scientific Chair in Evidence Based Vaccinology, financed in 2011–2014 by a gift from Pfizer. G.M., P.B., N.H. and A.B. acknowledge support from a Methusalem research grant from the Flemish government. G.M. and N.H. received funding from IAP research Network P7/06 of the Belgian Government (Belgian Science Policy).

## References

1. Murray CJL, Vos T, Lozano R. *et al*. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012;**380**:2197–223.
2. Ewig S, Birkner N, Strauss R. *et al*. New perspectives on community-acquired pneumonia in 388 406 patients. Results from a nationwide mandatory performance measurement programme in healthcare quality. *Thorax* 2009;**64**:1062–9.
3. Kwok CS, Loke YK, Woo K. *et al*. Risk prediction models for mortality in community-acquired pneumonia: a systematic review. *Biomed Res Int* 2013;**2013**:504136.
4. Wu C, Rosenfeld R, Clermont G. Using data-driven rules to predict mortality in severe community acquired pneumonia. *PLoS ONE* 2014;**9**:e89053.
5. Gibberd R, Hancock S, Howley P. *et al*. Using indicators to quantify the potential to improve the quality of health care. *Int J Qual Health Care* 2004;**16** (Suppl. 1):i37–43.
6. Pneumonia mortality rate. In: Agency for Healthcare Research and Quality (ed). *AHRQ Quality Indicators—Guide to Inpatient Quality Indicators: Quality of Care in Hospitals—Volume, Mortality, and Utilization. Revision* (4th edn). Rockville, MD: Agency for Healthcare Research and Quality, 2004.
7. D'Hoore W, Bouckaert A, Tilquin C. Practical considerations on the use of the Charlson comorbidity index with administrative data bases. *J Clin Epidemiol* 1996;**49**:1429–33.
8. Schneeweiss S. Learning from big health care data. *N Engl J Med* 2014;**370**:2161–3.
9. Garau J, Baquero F, Perez-Trallero E. *et al*. Factors impacting on length of stay and mortality of community-acquired pneumonia. *Clin Microbiol Infect* 2008;**14**:322–9.
10. Glance LG, Dick AW, Osler TM. *et al*. Using hierarchical modeling to measure ICU quality. *Intensive Care Med* 2003;**29**:2223–9.
11. Spiegelhalter D, Sherlaw-Johnson C, Bardsley M. *et al*. Statistical methods for healthcare regulation: rating, screening and surveillance. *J Roy Stat Soc Ser A (Stat Soc)* 2012;**175**:1–47.
12. Ash AS, Shwartz M, Pekoz EA. Comparing outcomes across providers. In: Iezzoni LI (ed). *Risk Adjustment for Measuring Health Care Outcomes* (3rd edn). Chicago, IL: Health Administration Press, 2003.
13. Shahian DM, Normand SL. Comparison of 'risk-adjusted' hospital outcomes. *Circulation* 2008;**117**:1955–63.
14. Spiegelhalter DJ. Funnel plots for comparing institutional performance. *Stat Med* 2005;**24**:1185–202.
15. van Baal PH, Engelfriet PM, Hoogenveen RT. *et al*. Estimating and comparing incidence and prevalence of chronic diseases by combining GP registry data: the role of uncertainty. *BMC Public Health* 2011;**11**:163.
16. Girling AJ, Hofer TP, Wu J. *et al*. Case-mix adjusted hospital mortality is a poor proxy for preventable mortality: a modelling study. *BMJ Qual Saf* 2012;**21**:1052–6.
17. Mark DH. Race and the limits of administrative data. *JAMA* 2001;**285**:337–8.
18. Dixon N. Proposed standards for the design and conduct of a national clinical audit or quality improvement study. *Int J Qual Health Care* 2013;**25**:357–65.
19. Slaghuis SS, Strating MM, Bal RA. *et al*. A measurement instrument for spread of quality improvement in healthcare. *Int J Qual Health Care* 2013;**25**:125–31.
20. Rubenstein L, Khodyakov D, Hempel S. *et al*. How can we recognize continuous quality improvement? *Int J Qual Health Care* 2014;**26**:6–15.
21. Ukawa N, Ikai H, Imanaka Y. Trends in hospital performance in acute myocardial infarction care: a retrospective longitudinal study in Japan. *Int J Qual Health Care* 2014;**26**:516–23.
22. Rubenstein LV, Hempel S, Farmer MM. *et al*. Finding order in heterogeneity: types of quality-improvement intervention publications. *Qual Saf Health Care* 2008;**17**:403–8.
23. Hartz AJ, Krakauer H, Kuhn EM. *et al*. Hospital characteristics and mortality rates. *N Engl J Med* 1989;**321**:1720–5.
24. Ansari MZ, Ackland MJ, Jolley DJ. *et al*. Inter-hospital comparison of mortality rates. *Int J Qual Health Care* 1999;**11**:29–35.
25. Meyer GS, Krakauer H. Validity of the Department of Defense Standard Inpatient Data Record for quality management and health services research. *Mil Med* 1998;**163**:461–5.
26. Pitches D, Burls A, Fry-Smith A. How to make a silk purse from a sow's ear—a comprehensive review of strategies to optimise data for corrupt managers and incompetent clinicians. *BMJ* 2003;**327**:1436–9.
27. Pouw ME, Peelen LM, Moons KG. *et al*. Including post-discharge mortality in calculation of hospital standardised mortality ratios: retrospective analysis of hospital episode statistics. *BMJ* 2013;**347**:f5913.
28. Drye EE, Normand SL, Wang Y. *et al*. Comparison of hospital risk-standardized mortality rates calculated by using in-hospital and 30-day

models: an observational study with implications for hospital profiling. *Ann Intern Med* 2012;**156**(1_Part_1):19–26.
29. Neuburger J, Cromwell DA, Hutchings A. *et al.* Funnel plots for comparing provider performance based on patient-reported outcome measures. *BMJ Qual Saf* 2011;**20**:1020–6.
30. Kaplan V, Angus DC, Griffin MF. *et al.* Hospitalized community-acquired pneumonia in the elderly. Age- and sex-related patterns of care and outcome in the United States. *Am J Respir Crit Care Med* 2002;**165**:766–72.
31. Menendez R, Torres A, Reyes S. *et al.* Compliance with guidelines-recommended processes in pneumonia: impact of health status and initial signs. *PLoS One* 2012;**7**:e37570.
32. Macfarlane JT, Boldy D. 2004 update of BTS pneumonia guidelines: what's new? *Thorax* 2004;**59**:364–6.
33. Kothe H, Bauer T, Marre R. *et al.* Outcome of community-acquired pneumonia: influence of age, residence status and antimicrobial treatment. *Eur Respir J* 2008;**32**:139–46.
34. Thomsen RW, Riis A, Noergaard M. *et al.* Rising incidence and persistently high mortality of hospitalized pneumonia: a 10-year population-based study in Denmark. *J Intern Med* 2006;**259**:410–7.
35. Guidelines for the management of adults with hospital-acquired, ventilator-associated, and healthcare-associated pneumonia. *Am J Respir Crit Care Med* 2005;**171**:388–416.

**Table A1**

| Weights | Conditions | ICD-9 codes |
|---|---|---|
| 1 | Myocardial infarction | 410, 411 |
| | Congestive heart failure | 398, 402, 428 |
| | Peripheral vascular disease | 440–447 |
| | Dementia | 290, 291, 294 |
| | Cerebrovascular disease | 430–433, 435 |
| | Chronic pulmonary disease | 491–493 |
| | Connective tissue disease | 710, 714, 725 |
| | Ulcer disease | 531–534 |
| | Mild liver disease | 571, 573 |
| 2 | Hemiplegia | 342, 434, 436, 437 |
| | Moderate or severe renal disease | 403, 404, 580–586 |
| | Diabetes | 250 |
| | Any tumor | 140–195 |
| | Leukemia | 204–208 |
| | Lymphoma | 200, 202, 203 |
| 3 | Moderate or severe liver disease | 070, 570, 572 |
| 6 | Metastatic solid tumor | 196–199 |

## Appendix 1

Scoring the comorbidity index from secondary diagnoses by the Charlson's comorbidity index, D'Hoore implementation [7] (CCI) (Table A1).

## Appendix 2

Identification of pneumonia cases, intensity of care and ARF by means of ICD-9-CM and ICD-9-related text strings

(i) Definition of CAP-cases in MCD (Table A2)

Building on the work of the Agency for Healthcare Research and Quality (AHRQ) [6], we selected, for the years 2004–07, from the Minimal Clinical Data all stays having CAP as principal diagnosis. We adopted the AHRQ definition of pneumonia (Inpatient Quality Indicator 20) [6]: 'Hospitalized patients with a *principal diagnosis* of pneumonia to the exclusion of patients with missing discharge disposition, transferring to another short-term hospital, Major Diagnostic Categories (MDC) 14 (pregnancy, childbirth, and puerperium), MDC 15 (newborns and other neonates)' and patients <18 years. However, in close consultation with clinicians, we decided to adapt the AHRQ selection of the ICD-9-CM codes, used to identify cases of pneumonia, to the Belgian situation.

Our selection included following codes:

(i) Intensity of care
IMV: 96.7*, 96.04, 97.37
NMV:93.9*
(ii) ARF: 518.81, 518.82
(iii) List of text strings used to identify CAP cases from the discharge field in the CARENET database (Table A3)

## Appendix 3: The Carenet hospitalization database

The Carenet data are constructed under the initiative of the seven sickness funds in Belgium and operates independently from the MCD database. Hospitals that choose to participate in Carenet provide hospitalization records including primary and secondary diagnoses, patient and hospital identifiers and time and date of hospitalization. The proportion of included hospitals in the database has grown

**Table A2**

| | | | |
|---|---|---|---|
| 4808 | VIRAL PNEUMONIA NEC | 48289 | BACT PNEUMONIA NEC |
| 4809 | VIRAL PNEUMONIA NOS | 4829 | BACTERIAL PNEUMONIA NOS |
| 481 | PNEUMOCOCCAL PNEUMONIA | 4830 | MYCOPLASMA PNEUMONIA |
| 4820 | K. PNEUMONIAE PNEUMONIA | 4831 | CHLAMYDIA PNEUMONIA OCT96- |
| 4821 | PSEUDOMONAL PNEUMONIA | 4838 | OTH SPEC ORG PNEUMONIA |
| 4822 | H. INFLUENZAE PNEUMONIA | 4841 | PNEUM W CYTOMEG INCL DIS |
| 48230 | STREP PNEUMONIA UNSPEC | 4843 | PNEUMONIA IN WHOOP COUGH |
| 48231 | GRP A STREP PNEUMONIA | 4845 | PNEUMONIA IN ANTHRAX |
| 48232 | GRP B STREP PNEUMONIA | 4846 | PNEUM IN ASPERGILLOSIS |
| 48239 | OTH STREP PNEUMONIA | 4847 | PNEUM IN OTH SYS MYCOSES |
| 4824 | STAPHYLOCOCCAL PNEUMONIA | 4848 | PNEUM IN INFECT DIS NEC |
| 48240 | STAPH PNEUMONIA UNSP OCT98- | 485 | BRONCOPNEUMONIA ORG NOS |
| 48241 | STAPH AUREUS PNEUMON OCT98- | 486 | PNEUMONIA, ORGANISM NOS |
| 48249 | STAPH PNEUMON OTH OCT98- | 5070 | FOOD/VOMIT PNEUMONITIS |
| 48281 | ANAEROBIC PNEUMONIA | 5100 | EMPYEMA WITH FISTULA |
| 48282 | E. COLI PNEUMONIA | 5109 | EMPYEMA W/O FISTULA |
| 48283 | OTH GRAM NEG PNEUMONIA | 5110 | PLEURISY W/O EFFUS OR TB |
| 48284 | LEGIONNAIRES DX OCT97- | 5130 | ABSCESS OF LUNG |

**Table A3**

| ICD 9 code | Condition | Search strings used in SQL in brackets, % is a wildcard[a] |
|---|---|---|
| 460–486 | Pneumonia excluding influenza | '%PNEUMONI%'; '%LONGONT%'; '480%'; '481%'; '482%; '483%'; '484%'; '485%'; '486%'; '% 480%'; '% 481%'; '% 482%; '% 483%'; '% 484%'; '% 485%'; '% 486%' |
| 510.0 | Empyema within the respiratory system, with mention of fistula | '%EMPYEMA%'; '5100%'; '% 5100%'; '510.0%'; '% 510.0%' |
| 510.9 | Empyema within the respiratory system, without mention of fistula | '%EMPYEMA%'; '5109%'; '% 5109%'; '510.9%'; '% 510.9%' |
| 513.0 | Abscess of the lung | '5130%'; '% 5130%'; '513.0%'; '% 513.0%' |

[a]A wildcard '%' stands for any series of characters. Using wildcards is necessary because of the presence of multiple diagnoses and the combination of text and ICD9 coding within the Carenet discharge field.

substantially during the study period, from 11% of hospital beds in 2004 to 91% in 2007.

For each recorded hospitalization, a discharge record includes a list of primary and optional secondary diagnoses as free text fields, which usually includes ICD9-coding or diagnoses in text. In this diagnosis field, we performed a text string search to identify CAP cases with pneumonia based on ICD9 code or the text string (see Appendix 2, Table A3).

For members of the National Alliance of Christian Sickness Funds (NACSF), one of the seven sickness funds, we linked CAP hospitalizations to patient characteristics (age, gender and survival). The NACSF membership, ~44% of the Belgian population, shows a slight overrepresentation of the older age groups and a small underrepresentation of the unemployed.

All analyses based on the joint information in the Carenet and NACSF internal databases were performed at NACSF under supervision of a social security physician. The other research partners received no personally identifiable information (including small cells) from NACSF.