Health Policy 121 (2017) 842–852

Contents lists available at ScienceDirect

Health Policy

journal homepage: www.elsevier.com/locate/healthpol

Hospital quality variation matters – A time-trend and cross-section analysis of outcomes in German hospitals from 2006 to 2014

Christoph Pross[a], Reinhard Busse[a,b], Alexander Geissler[a,*]

[a] Berlin University of Technology, Department of Health Care Management, Germany
[b] European Observatory on Health Systems and Policies, Berlin Centre of Health Economics Research, Germany

ARTICLE INFO

Article history:
Received 16 February 2017
Received in revised form 19 June 2017
Accepted 25 June 2017

Keywords:
Hospital quality of care
Outcome variation
Quality assessment
AMI
Stroke
PCI
Cholecystectomy
Hip replacement
Pacemaker implantation
Germany

ABSTRACT

Awareness of care variation and associated differences in outcome quality is important for patients to recognize and leverage the benefits of hospital choice and for policy makers, providers, and suppliers to adapt initiatives to improve hospital quality of care. We examine panel data on outcome quality in German hospitals between 2006 and 2014 for cholecystectomy, pacemaker implantation, hip replacement, percutaneous coronary intervention (PCI), stroke, and acute myocardial infarction (AMI). We use risk-adjusted and unadjusted outcomes based on 16 indicators. Median outcome and outcome variation trends are examined via box plots, simple linear regressions and quintile differences. Outcome trends differ across treatment areas and indicators. We found positive quality trends for hip replacement surgery, stroke and AMI 30-day mortality, and negative quality trends for 90-day stroke and AMI readmissions and PCI inpatient mortality. Variation of risk-adjusted outcomes ranges by a factor of 3–12 between the 2nd and 5th quintile of hospitals, both at the national and regional level. Our results show that simply measuring and reporting hospital outcomes without clear incentives or regulation – "carrots and sticks" – to improve performance and to centralize care in high performing hospitals has not led to broad quality improvements. More substantial efforts must be undertaken to narrow the outcome spread between high- and low-quality hospitals.

1. Introduction

In recent years, substantial variation in the quality of hospital process and outcome indicators has been observed in and across health systems [1–5]. This variation is reported for emergency conditions, such as acute myocardial infarction (AMI) or stroke [4,7], as well as for elective procedures such as joint replacements and pacemaker implantation [8].

Across all conditions, outcome variation is often caused by medical errors and adverse events, which can be attributed to structural, process, and/or training deficiencies [9,10]. Data from European Union member states show that medical errors and adverse events occur in 8% to 12% of hospitalizations [11]. Studies analyzing medical errors show that 50–70% of these are preventable through comprehensive patient safety measures, better communication, and quality management systems [12].

Reducing outcome variation and improving outcomes overall are priorities for both policy makers and clinical leaders. In many countries, significant financial and time resources are currently being invested to build or enhance quality measurement and reporting systems, and to operate these systems in clinical practice [13,14]. Some of these systems were first implemented more than a decade ago. Examining recent time trends of hospital quality and inter-hospital variation is critical to evaluate and improve these measurement and reporting programs since their implementation.

Studies examining United States (US) hospital outcome data have reported mixed outcome quality trends. A study analyzing administrative data from 2001 to 2005 and multiple databases found improved risk-adjusted mortality, but mixed results for risk-adjusted complications and morbidity results [15]. Likewise, an analysis using Hospital Compare data from 2003 to 2009 for hospitals in 13 US states deduced no significant improvements in risk-adjusted mortality or complication ratios in Coronary Artery Bypass Graft (CABG) surgery or hip and knee replacements [16]. Similarly, no improvements in risk-adjusted death or re-hospitalization for ischemic stroke between 2003 and 2006 were identified [7]. In contrast, Cohen et al. found significant improvements in adverse, risk-adjusted surgical outcomes such as

* Corresponding author at: Berlin University of Technology, Dept. of Health Care Management, Administrative Office H80, Str. des 17. Juni 135, 10623, Berlin, Germany.
E-mail address: a.geissler@tu-berlin.de (A. Geissler).

http://dx.doi.org/10.1016/j.healthpol.2017.06.009
0168-8510/

mortality, morbidity, and infection between 2006 and 2013 in hospitals participating in the American College of Surgeons National Surgical Quality Improvement Program [17]. And a significantly improved overall standardized in-hospital AMI mortality was identified in data from the US Nationwide Inpatient Sample between 2001 and 2011 [18]. For a shorter timeframe, Werner and Bradlow observed improved hospital performance for process and outcome measures for AMI, heart failure, and pneumonia in CMS Hospital Compare data between 2004 and 2006 [19]. Results in European countries are similarly varied [20–24].

Since 2004, all German acute care hospitals are obliged to report structural, process, and outcome indicators as part of a national quality monitoring program [25]. The program collects clinical data for more than 350 indicators in 30 treatment areas, and as such places substantial administrative and resource-intensive burdens on hospitals [26]. To expand and improve national quality monitoring, the German government has recently created the Institute for Quality Assurance and Transparency in Healthcare (IQTIG). Annual national hospital quality reports summarize quality measures at a national average and relate them to those from the year prior for comparability and benchmarking purposes [27]. A study of the underlying hospital-level data for the 2004–2008 period found an overall improvement of 50% based on 204 analyzed indicators; however, this was primarily shaped by process indicators and many of the outcome indicators showed deterioration [28]. An investigation of national inpatient diagnosis-related group (DRG) statistics found AMI mortality rates to remain constant between 2005 and 2009 [29]. In a more recent analysis, Kraska et al. examined two indication appropriateness, three process, and one unadjusted outcome indicator from the mandatory hospital reporting from 2006 to 2012 and found significant quality improvements across all six measures. The data was divided into two-year periods, and he main improvement was observed to have occurred within the first measurement interval from 2006 to 2008 [30].

Using data for more substantial time periods (i.e. >5 years) and a greater number of hospitals (e.g. all acute hospitals for one country) has only recently become possible as many outcome measurement systems worldwide were only initiated in the last ten years [31]. Yet the evidence of improved (or deteriorated) outcomes, especially risk-adjusted outcomes, is mixed in Germany and other countries. Stroke and AMI outcome indicators in Germany, for example, have not yet been analyzed comprehensively, nor over time. In the international context, the analysis of inter-hospital variation over time is also scarce.

To address some of these research gaps, we analyze outcomes and their variation between 2006 and 2014 in German hospitals for the inpatient procedures cholecystectomy, pacemaker implantation, hip joint replacement, and percutaneous coronary intervention (PCI), and two emergency conditions, stroke and AMI. We examine three dimensions: 1) whether raw andy, pacemaker implantation, hip joint replacement, and percutaneous coronary intervention (PCI), and two emergency conditions, stroke and AMI.

2. Methods

2.1. Dataset

We analyze a hospital-level panel dataset with repeated observations, and integrate 16 outcome indicators for six treatment areas from the two premier provider quality reporting systems in Germany. These are the national mandatory quality monitoring system of the Federal Joint Committee (G-BA) and the Quality Assurance with Administrative Data (QSR) from Germany's largest sickness fund, the AOK (25.3 Mio insured persons, ~35% within the statutory health insurance scheme in 2016 [32]). We focus on outcome indicators as they are relevant end-points for patients. Further, the scientific evidence for the direct relationship between process and outcome indicators is unclear [33,34].

Indicators from the national quality monitoring system of G-BA are self-reported by hospitals, documented at the patient level, and publicly available for research and reporting at an aggregated hospital level. They comprise indication appropriateness, process, and outcome indicators for 30 treatment areas, from which we have selected a subset of ten (see Table 1). The selection considers data restrictions both in terms of indicator comparability across years and size of hospital samples.

The ten selected G-BA outcome indicators cover the four inpatient procedures only. They include six unadjusted outcome indicators such as mortality and re-intervention rates and four risk-adjusted outcome indicators such as risk-adjusted mortality and risk-adjusted re-intervention ratios. Risk-adjusted indicators compare the number of observed events (e.g. mortalities) with the number of expected events, with the latter calculated through a logistic regression that accounts for patient risk-factors such as age, comorbidities, gender and patient volume [35]. Annual risk-adjustment is undertaken centrally and ensures comparability of outcomes across hospitals and their respective patient samples.

The remaining six outcome indicators are from QSR and apply to the two emergency medical conditions stroke and AMI. They are centrally calculated by the WIdO Research Institute (a scientific body of the AOK), based on administrative data of AOK-insured patients. Stroke includes ICD diagnoses (i) intracerebral hemorrhage (ICD Code I61), (ii) ischemic stroke (I63), and (iii) stroke not specified as hemorrhage or ischemic (I64). AMI includes the diagnoses ST and non-ST elevation myocardial infarction (STEMI and NSTEMI) (I21) and subsequent STEMI and NSTEMI myocardial infarction (I22). The 30-day standardized mortality ratio (SMR) includes events up to 30 days after hospital discharge, comparing number of observed with number of expected events. Expected events for all data years are calculated based on the latest 2014 risk-adjustment model, via a logistic regression based on patients treated in each year and their respective risk-factors such as age, gender and comorbidities [36]. 30-day mortality and 90-day readmission rates are unadjusted. The QSR stroke and AMI data is not available publicly, access was granted only for this study and outcomes were provided at an aggregated level for each hospital. We link QSR indicators with the national quality monitoring data through standardized hospital IDs and hospital address data.

2.2. Statistical analysis

To examine outcome trends over time, we use box and whisker plots. For each indicator, respective patient volumes are included as analytical weights. The median illustrates the outcome trend. The 75th and 25th percentiles (box) specifies the interquartile range (IQR), the upper and lower adjacent values (whisker) are the most extreme values within 1.5 × IQR of the respective quartile. Both methods capture outcome variation.

To validate graphically identified time trends, we model indicator time trends using linear, longitudinal regressions. The log of the respective outcome indicator QI_{it} for hospital i in year t is the dependent variable and $YEAR_t$ and a log_dummy_{it} are predictors. The time variable is an exogenous variable and the year coefficient summarizes the outcome time trend. Outcomes are log transformed to reduce skew and approximate normality [37]. We correct for zero values in the dependent variable, i.e. the optimal level of realized quality – a zero standardized mortality ratio (SMR) or zero non-risk-adjusted outcome ratios – on both sides of the equation [38]. On the right hand of the regression equation, we include a dummy variable that takes on the value of 1 if the outcome (the dependent variable) is zero. On the left hand side, we add 1 to the outcome if it