Health policy 172 (2026) 105703

Contents lists available at ScienceDirect

Health policy

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

Does practice make perfect? Panel evidence from hip fracture surgery in Italy

Sahar Paktinat, Cristina Ugolini, Rossella Verzulli *

Department of Economics, University of Bologna, Piazza Scaravilli 2, 40126, Bologna, Italy

ARTICLE INFO

JEL Classifications:
C20
I10
I18
H51

Keywords:
Surgical volumes
Learning-by-doing
Surgeons experience
Patient outcomes

ABSTRACT

Background: A large body of evidence documented a positive association between hospital volume and patient outcomes, but the underlying causal mechanisms remain debated. While the "practice makes perfect" hypothesis suggests quality gains from centralization, the "selective referral" hypothesis raises concerns about access and equity.

Objective: To examine the causal effect of surgical volumes on outcomes for hip fracture patients in a non-competitive healthcare system. In this setting, patients are routinely referred to the nearest hospital for urgent surgeries, and mandatory minimum-volume thresholds have been in place since 2016. These institutional features offer a robust context to test the "practice makes perfect" hypothesis.

Methods: The analysis is based on a cohort of emergency patients admitted to publicly funded hospitals in the Emilia-Romagna region of Italy between 2017 and 2019, after implementation of Ministerial Decree DM70/2015, which introduced volume thresholds for hip fracture surgeries. Using multi-way fixed effects models and a comprehensive set of covariates to control for confounders, we estimate the causal effect of surgical volumes on 30-day mortality.

Results: Our findings document a robust volume-outcome relationship: higher lagged surgical volumes are significantly associated with lower 30-day mortality rates. The fact that the effect strengthens with longer lag periods is consistent with a learning-by-doing mechanism, whereby accumulated experience improves outcomes.

Conclusions: In non-competitive systems, volume-based standards for urgent procedures are associated with better outcomes and provide a rationale for further centralization, as long as ensuring universal access remains central.

1. Introduction

Over the past decades, an extensive empirical literature, starting with the study by Luft et al. (1987) [1], has documented a positive correlation between hospital volume and patient outcomes. This pattern, observed across various healthcare contexts, is typically associated with lower mortality and fewer post-surgical complications [2]. Two main mechanisms have been proposed: the "practice makes perfect" hypothesis, which attributes better outcomes to learning-by-doing and economies of scale; and the "selective referral" hypothesis, which suggests that hospitals perceived as higher quality attract more patients, increasing volume. Evidence for the latter highlights the role of reputation in influencing patient choice and driving volume growth [3,4].

The existence of alternative explanatory mechanisms introduces potential bias in estimating the volume–outcome relationship, mainly due to the endogeneity of volume. This is particularly relevant for policymaking, as understanding the direction of causality is crucial. If higher volumes improve outcomes ("practice makes perfect"), centralizing procedures in high-volume centres may enhance efficiency and quality. If instead better outcomes attract more patients ("selective referral"), centralization may have limited effects and could reduce access, especially in underserved areas [5,6].

Selective referral is more plausible for elective, non-urgent surgeries within competitive systems, where patients can choose among providers and reputation plays a key role. Conversely, "practice makes perfect" is more convincing for emergency procedures, where patient choice is minimal and volume-driven learning effects are stronger. Based on this rationale, several countries have implemented centralization policies for high-risk or emergency surgeries, including minimum volume thresholds for hospitals, wards, and individual surgeons [7–10]. However,

* Corresponding author at: Piazza Scaravilli 2, 40126, Bologna, Italy.
E-mail address: rossella.verzulli@unibo.it (R. Verzulli).

https://doi.org/10.1016/j.healthpol.2026.105703
Received 31 July 2025; Received in revised form 16 June 2026; Accepted 26 June 2026
Available online 14 July 2026
0168-8510/© 2026 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

volume-based restrictions may also have unintended consequences, such as the loss of specialized expertise in discontinued services and related areas, particularly in acute care. They may also widen geographic disparities in access, raising concerns about equity for patients in remote regions [11,12]. Consistent with these concerns, a growing literature shows that longer travel times can significantly worsen survival and health outcomes in time-sensitive emergencies, such as acute myocardial infarction (AMI), stroke and childbirth [13–16]. These findings foreground the policy trade-off between learning-by-doing gains and timely access for time-critical care.

Our study estimates the effect of surgical volume on health outcomes for emergency hip fracture surgeries performed between 2017 and 2019 in Emilia-Romagna, a northern Italian region with a centrally planned healthcare system characterized by strong inter-hospital coordination and no market-based competition. In this regulated context, regional policymakers promote provider cooperation through service integration and centralized capacity planning. Hip fractures, mainly affecting the elderly due to falls, are the leading cause of emergency orthopaedic admissions in Italy [17]. Within this framework, as part of national quality improvement efforts, Ministerial Decree 70/2015 introduced a minimum annual volume of 75 hip fracture surgeries per hospital ward for patients over 65, starting in 2016, based on the assumption that higher volumes improve outcomes.

A distinctive feature of the setting we examine is its institutional organization: regional coordination, rather than patient sorting, shapes provider volumes. Once stabilised, hip fracture patients are transferred by ambulance to the nearest emergency department (ED). This allows us to analyse a context where ward-level volumes are plausibly policy-driven rather than demand-driven, thus limiting selective referral and providing a cleaner test of the practice-makes-perfect mechanism.

We assess whether the 2016 minimum volume standards improved outcomes and whether stricter thresholds could deliver additional gains. The results are twofold. First, we find evidence of a robust volume-outcome relationship: higher lagged ward-level volumes are associated with lower 30-day mortality, consistent with a learning-by-doing mechanism whereby accumulated clinical experience improves outcomes. Second, the current threshold appears effective, but findings suggest that a higher procedural minimum could further enhance results, reinforcing the case for additional geography-calibrated centralization.

2. Background

An extensive medical literature has investigated the relationship between volume and health outcomes across various procedures and countries, finding a generally positive correlation [2,18]. Yet, the evidence emerging from the economic literature is mixed, with differences across conditions, institutional contexts and the empirical designs used to address the endogeneity of hospital volume.

Focusing on orthopaedic surgery, where surgical procedures are not strictly time-sensitive and travel distance is therefore less salient for immediate survival, a study from Germany [19] reports a causal reduction in mortality following emergency hip fracture surgery in high-volume hospitals. In contrast, other studies from Canada and the UK [20,6] find no significant effect of hospital volume on outcomes once selection biases and unobserved heterogeneity are accounted for. At the surgeon level, empirical research from an economic perspective is still at its infancy, with recent contributions reporting short-term improvements in performance associated with increased procedural activity [21, 22]. Although endogeneity remains a central methodological challenge in the economic literature, the problem appears to be particularly relevant for elective procedures and within institutional settings where patients are free to choose their hospital, as selective referral mechanisms are more likely to influence observed volumes and outcomes.

Even though a robust volume–outcome relationship in orthopaedics has not been definitively established, policymakers have long operated under the assumption that such a relationship exists, introducing minimum volume regulations aimed at improving the quality of care. Several studies have examined the effects of these policies in different national contexts. For example, in Germany, such policies showed no clear effect on elective total knee arthroplasty outcomes [9], while in the Netherlands, centralizing surgical elective and urgent procedures led to better outcomes with limited access concerns due to geographic factors [12]. Overall, the existing body of empirical research predominantly pertains to healthcare systems where hospital competition is prevalent and/or there is substantial patient sorting. In such environments, given the widespread availability of hospital facilities equipped to perform surgeries, it is plausible that patients, even when making quick decisions, incorporate the reputational signals of nearby facilities into their choices. This introduces potential endogeneity into estimations of the volume–outcome relationship, primarily due to the possibility of simultaneous causation.

Our study contributes to the existing literature by examining the volume-outcome relationship within the context of a National Health System (NHS), where centralized planning, strong vertical integration, and the presence of coordinated hospital networks are defining institutional features. In this setting, the allocation of surgical activity is determined by health authorities, rather than market forces. Moreover, the implementation of national minimum volume standards imposes additional constraints on service delivery. This institutional configuration allows urgent surgical volumes to be treated as an exogenous variable, offering a favourable empirical setting for evaluating the effects of minimum volume thresholds on healthcare quality and outcomes.

The Italian NHS provides universal coverage through tax-based financing, and regions have broad autonomy in organizing and managing their healthcare systems [23]. In Emilia-Romagna - where this study is based - healthcare is centrally planned, with a strong emphasis on coordination among providers and integration between hospital and local services. With a population of 4.5 million, the region's system is comparable in scale and standards to those of small to medium-sized Western European countries.

Orthopaedic care in Emilia-Romagna is organized through an integrated network of public and accredited private healthcare facilities, aiming to provide high-quality and accessible treatment for the entire population. At the regional level, a technical-scientific commission for orthopaedic care is responsible for monitoring the quality of care, developing proposals for clinical-organizational innovation and improvement, and supporting regional planning by determining the need for orthopaedic procedures and surgeries. At provincial level, orthopaedic services are further coordinated through an Osteoarticular Department, which is tasked with ensuring consistent orthopaedic clinical governance across the network. This department coordinates orthopaedic and trauma care activities, ensuring appropriate, timely, and patient-centred assistance. Overall, the organization of orthopaedic care in Emilia-Romagna is based on a network of specialized facilities, centralized planning, and strong inter-institutional collaboration, with the aim of providing high-quality orthopaedic care to all residents. Within this coordinated organizational framework, the management of emergency cases, such as hip fractures, follows a standardized and centralized protocol to ensure timely and appropriate care. When a patient sustains a hip fracture, the care pathway is initiated with an emergency response, involving transport to the nearest appropriate facility equipped to handle such emergency. This system reflects the region's commitment to equitable access and prompt intervention, optimizing the use of available resources within a structured and integrated care system. Since 2016, the region has restructured the orthopaedic network for urgent procedures to meet the threshold of 75 hip fracture surgeries per year. This goal is part of regional health planning and is linked to performance-based incentives for Local Health Authority managers. For hospitals below the threshold, accessibility is assessed to guide case concentration strategies aimed at improving outcomes.

Table 1
Summary statistics.

	Mean	Std. Dev.	Min	Max
Patient characteristics				
30-day mortality	0.034	0.183	0	1
Age	83.90	7.437	65	100
Female	0.755	0.430	0	1
Pre-operative wait > 48 h	0.205	0.404	0	1
Post-surgical length of stay	9.896	6.244	0	175
CCI 0	0.659	0.474	0	1
CCI 1	0.195	0.396	0	1
CCI 2	0.081	0.274	0	1
CCI Over 2	0.064	0.245	0	1
ASA I	0.010	0.101	0	1
ASA II	0.233	0.423	0	1
ASA III	0.646	0.478	0	1
ASA IV	0.109	0.312	0	1
ASA V	0.001	0.034	0	1
Admission during weekend	0.272	0.445	0	1
Operation during weekend	0.124	0.330	0	1
Type of surgery				
Partial or total hip replacement	0.497	0.500	0	1
Fracture reduction	0.502	0.500	0	1
Type of hip fracture				
Pertrochanteric	0.438	0.496	0	1
Transcervical	0.535	0.499	0	1
Other type of fracture	0.027	0.163	0	1
Observations	10,377			
Ward volumes for hip fracture surgeries				
Surgeries in the past 4 weeks	17.825	0.325	0	60
Surgeries in the past 12 weeks	53.149	28.985	0	158
Surgeries in the past 24 weeks	105.260	56.369	0	298
Surgeries in the past 52 weeks	224.453	116.918	0	595

Notes: Sample period: years 2017–2019. ASA: American Society of Anesthesiologists.

3. Methods

3.1. Data

The primary data source for this study is the regional administrative hospital discharge dataset which is linked to data from the mortality register. The unit of analysis is the individual patient. A key advantage of using patient-level data relies on the capacity to control for individual health characteristics, thereby reducing unobserved heterogeneity across observations and improving the accuracy of outcome estimates. The database provides comprehensive information on patient demographics, clinical variables as well as specifics about their hospital stay. In addition to patient-centred information, the dataset includes a physician identifier for each intervention, allowing for the identification of the primary surgeon, secondary surgeon, and anaesthesiologist involved in the intervention. We also integrated information from the regional surveillance system for surgical site infections (SIChER), which collects clinical data for surgeries performed in inpatient settings. This additional source enabled us to enhance our ability to adjust for patient risk factors through the inclusion of the ASA Physical Status Classification System (ASAPS), a widely used proxy for perioperative risk [24].

Our dataset includes patients over 65 with a diagnosis of femoral neck fracture (ICD-9-CM: 820.0–820.3) treated either with fracture reduction (ICD-9-CM: 79.15, 79.35) or with partial or total hip replacement (ICD-9-CM: 81.51, 81.52). Following the National Guidelines, we excluded patients who were not residing in Italy at the time of the operation, those younger than 65 or older than 100, polytrauma patients, patients diagnosed with malignant tumours in the previous two years, and patients transferred from another institution [25]. The final dataset comprises 11,308 hip fracture patients who received a surgery in one of the 39 regional hospitals (65 hospital wards, 466 first surgeons) during the period from 2017 to 2019.

As outcome measures, we had several mortality windows available: in-hospital mortality, 30-day mortality, and 60-day mortality. While most studies in this field primarily focus on in-hospital mortality due to data limitations, we chose to prioritize 30-day mortality, as it aligns with the outcome monitored at the national level [25]. Furthermore, we opted not to include 60-day mortality in our analysis, as this measure requires careful interpretation. Given the elderly nature of the patient cohort under study, external factors increasingly impact mortality as the time from the surgical intervention extends. This made 30-day mortality a more appropriate and reliable indicator for evaluating outcomes in this context.

Table 1 shows a 30-day mortality rate of 3.4% in the overall sample. At the national level, the Italian data documented a higher 30-day mortality rate for hip fractures, around 5% for the period 2017–2019 [25]. This comparison suggests that the outcome in our dataset is favourable relative to the Italian benchmark, potentially reflecting effective perioperative care and patient management.

In addition to outcome data, Table 1 provides a summary of patient characteristics. The average patient age is 84 years, with females being most of the sample (75%). Approximately 27% of admissions occurred over the weekend, and 12% patients underwent surgery during that time. The data indicates that 20% of patients experienced a preoperative wait exceeding 48 h. This proportion falls well below the national maximum threshold of 40%, suggesting that preoperative management in this context is both timely and consistent with national quality standards. Clinically, 53% of cases were transcervical and 44% pertrochanteric fractures. These types differ in location - femoral neck vs. trochanteric area – with implications for surgical approach and prognosis due to differing risks of vascular disruption.

Surgical interventions were evenly distributed between fracture reductions (50%) and partial or total hip replacements (50%). To account for patient comorbidities, we use the Charlson Comorbidity Index (CCI) as a categorical variable based on grouped score ranges. Approximately 66% of sample patients were recorded as having no comorbidities, while the remaining 34% have at least one. Since the CCI may not fully capture the clinical complexity of the patient population, we also include a clinician-assigned measure of preoperative risk by using the American Society of Anesthesiologists (ASA) Physical Status Classification System, as additional control for the underlying health status of the population studied. Based on ASA scores, 65% of patients were classified as having a severe systemic disease with moderate limitations (ASA III), and an additional 11% as having more critical conditions (ASA IV-V). The remaining 24% of patients were classified as ASA I or II, indicating the absence of, or only mild, systemic disease.

Our key explanatory variable is the volume of hip fracture surgeries performed within each hospital ward over time. We employ lagged volume measures, transformed using the natural logarithm, as an aggregate proxy for accumulated clinical experience. We analyse the effect of lagged procedural volumes on patient mortality over different temporal horizons by constructing separate indicators of prior surgical activity over the previous month, quarter, half-year, and full year. This multi-interval framework allows us to assess how varying levels of surgical experience impact patient outcomes.

Table 1 also presents summary statistics at the ward level. The average number of emergency hip fracture surgeries performed is 18 per month, 53 per quarter, 104 per half-year, and 222 per year, with substantial standard deviation indicating considerable variability in procedural volumes across wards. To assess the impact of minimum volume requirements, we categorize ward volumes into categories based on their annual procedure counts. Specifically, we distinguish wards that fall below the mandated threshold and then stratify the remaining wards into evenly sized groups to ensure balanced observation counts across volume categories. The top panel of Fig. 1 provides a more granular view of ward activity by plotting a kernel density estimate of the distribution of hip–fracture surgeries performed per ward over the preceding 52 weeks. The horizontal axis represents the number of procedures, while the vertical axis indicates the density, reflecting the frequency of wards operating at different volume levels. Notably, 7.4% of observations lie

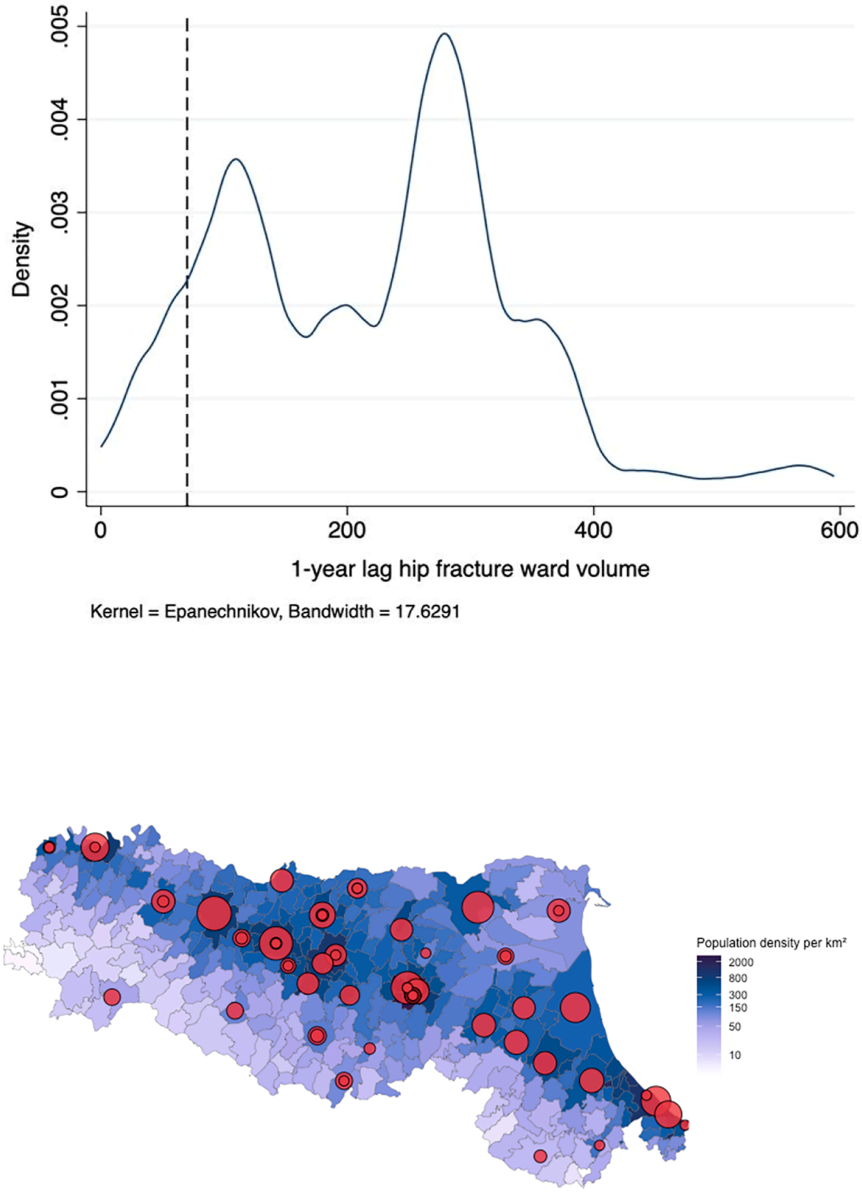

Fig. 1. Kernel density for 1-year lag hip fracture ward volume (top panel), and geographical distribution of hospital wards performing hip fracture surgeries in Emilia-Romagna (bottom panel). Notes: Red circles indicate the wards that performed hip fracture surgery during the study period; overlapping circles reflect multiple wards located within the same hospital, with smaller volumes drawn on top.

below the 75-procedure annual threshold. The bottom panel of Fig. 1 provides a map showing the location of hospital wards performing hip fracture surgery during the study period, together with their annual volumes and the population density per km^2 of the municipalities. High-volume wards cluster largely along the Via Emilia axis and in major urban centres. Lower-volume wards appear in rural and mountain municipalities as well as in some more densely populated areas, indicating that small ward size is not limited to geographically remote settings.

3.2. Empirical methodology

We identify the causal effect of increasing hip fracture volumes on patient outcomes by estimating the following multi-way fixed effects specification:

$$y_{iswht} = \beta_0 + \beta_1 \ln(volume)_{iwh,t-1} + X'_{iswht}\beta_2 + \alpha_s + \delta_h + \gamma_t + \varepsilon_{iswht} \quad (1)$$

where y_{iswht} is a binary indicator of 30-day mortality for patient *I*, treated by surgeon *s* in ward *w* at hospital *h* during year *t*. The principal explanatory variable is $\ln\ (volume)_{iwht-1}$ the natural logarithm of the lagged count of hip fracture surgeries performed in ward *w* of hospital *h* prior to the admission of patient *i*. X_{iswht} is a set of controls for patient characteristics, including age, gender, comorbidity measures, type of fracture and binary indicators for weekend hospital admissions and interventions. $\alpha_s,\ \delta_h$ and γ_t are fixed effects for surgeon, hospital and year, respectively, absorbing all time-invariant factors at those levels. The parameter of interest is β_1, capturing the elasticity of 30-day mortality with respect to prior ward-level procedural volumes. The inclusion of both hospital and surgeon fixed effects is particularly crucial for addressing the mobility of surgeons who practice across multiple hospitals. By including fixed effects at both levels, our specification isolates within-surgeon and within-hospital variation, thereby accurately capturing the impact of cross-institutional practice patterns on patient outcomes.

We complement this analysis by assessing the impact of minimum volume thresholds on patient outcomes. Specifically, we re-estimate Eq.

Table 2
Results of the effect of lagged surgical volumes on 30-day mortality.

	(1)		(2)		(3)		(4)	
Ln (Ward volume) $_{t-4}$	−0.00633 (0.00502)	−0.00578 (0.00487)						
Ln (Ward volume) $_{t-12}$			−0.00922 (0.00560)	−0.00643 (0.00545)				
Ln (Ward volume) $_{t-24}$					−0.0212** (0.00915)	−0.0167* (0.00886)		
Ln (Ward volume) $_{t-52}$							−0.0333*** (0.0122)	−0.0291** (0.0113)
Patient controls	N	Y	N	Y	N	Y	N	Y
Surgeon FE	Y	Y	Y	Y	Y	Y	Y	Y
Hospital FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
N	10,377	10,377	9996	9996	9327	9327	7550	7550
R2	−6373.8	−6684.1	−6173.3	−6468.2	−5508.5	−5804.7	−4429.8	−4660.5
Akaike crit. (AIC)	−6149.1	−6445.0	−5985.8	−6230.2	−5330.0	−5569.1	−4270.5	−4431.9
Bayesian crit. (BIC)	0.0571	0.0853	0.0521	0.0810	0.0574	0.0885	0.0648	0.0953

Notes: Sample period: years 2017–2019. Standard errors (in parentheses) are clustered on hospitals.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

Table 3
Effect of past year hospital ward volume thresholds, without/with patient controls.

	(1)	(2)
Past year hospital ward volume (baseline: ≤ 75)		
76–118 surgeries	−0.0556 (0.0441)	−0.0558 (0.0395)
119–186 surgeries	−0.0690 (0.0421)	−0.0657* (0.0378)
198–260 surgeries	−0.0776* (0.0432)	−0.0733* (0.0389)
261–285 surgeries	−0.0910** (0.0428)	−0.0849** (0.0384)
286–330 surgeries	−0.0954** (0.0435)	−0.0898** (0.0392)
≥ 331 surgeries	−0.103*** (0.0461)	−0.0940** (0.0432)
Patient controls	N	N
Surgeon FE	Y	Y
Hospital FE	Y	Y
Year FE	Y	Y
N	7550	7550
R^2	0.0640	0.0949
Akaike crit. (AIC)	−4413.7	−4657.3
Bayesian crit. (BIC)	−4219.7	−4428.7

Notes: Sample period: years 2017–2019. Standard errors (in parentheses) are clustered on hospitals.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

(1) by replacing $\ln(volume)_{iwh,t-1}$ with a vector of dummy variables corresponding to different volume categories. By differentiating between wards falling below the mandated threshold and those that exceed it, and by further stratifying the latter into higher-volume brackets, we aim at evaluating how varying levels of procedural volume influence surgical outcomes.

All regressions are estimated using a linear probability model (LPM) with clustered standard errors at the hospital level to account for potential geographical correlation among patients within the same hospital.

4. Results

Table 2 reports estimates considering different lag windows across columns. We find that increasing ward volumes (measured by the number of surgeries performed in previous periods) reduces 30-day mortality rates. As the number of lags increases, the negative coefficients grow larger and more significant, reinforcing the idea that accumulated experience in a ward could play a critical role in affecting patient outcomes. The results reported in column (4) reveal that as the number of surgeries performed in the previous year increases, the likelihood of 30-day mortality decreases significantly. Specifically, the results indicate that a 10% increase in the previous year's surgical volume decreases the average patient's probability of 30-day mortality by 0.3 percentage points approximately. This is consistent with a 9% reduction, relative to the baseline 30-day mortality rate of 3.4%. These results confirm that the higher past surgical volume in a ward lead to better patient outcomes, potentially reflecting factors such as increased provider experience, familiarity with procedures, or improved staff coordination in high-volume settings. The estimated effect is not statistically significant when lagged volumes refer to the previous month and quarter but becomes significant starting from the semester lag. Both the AIC and BIC criteria indicate that the model with 1-year lagged volumes (column 4) fits the data best.

As the analysis points to the presence of a learning-by-doing effect driven by cumulative annual procedural experience, we shift our focus to the minimum volume thresholds introduced in 2016. Specifically, we categorized wards into seven groups based on their lagged annual volumes of hip fracture surgeries. The reference group consists of hospital wards performing fewer than 75 hip fracture surgeries annually, and the other volume brackets are compared against this baseline. The results are reported in Table 3. Our estimates suggest that hospital wards performing 76 to 118 surgeries per year do not exhibit statistically significant differences in 30-day mortality compared to the reference group. In contrast, wards performing 119 or more surgeries annually show a statistically significant reduction in 30-day patient mortality. Precisely, wards in the 119–186 vol range achieve a lower mortality rate by about 6.6 percentage points compared to low-volume wards, and this mortality advantage further increases with higher procedural volumes.

To mitigate concerns about the presence of structural and geographic factors driving our findings, we re-estimated our models by excluding hospitals located in mountain municipalities along with, separately, those located in rural areas [26]. In both cases, the volume-mortality gradient remained negative and of similar magnitude to the baseline, suggesting that our results appear to be not driven by geographically "isolated" providers facing lower population density. Detailed results are available from the Authors upon request.

Research in context box

Extensive empirical research has shown a positive association between hospital volume and patient outcomes for various surgical procedures. However, the causal direction of this relationship remains debated, with two main hypotheses - “practice makes perfect” and “selective referral” - offering alternative explanations. Understanding the underlying mechanism is crucial for informing volume-based policies, especially for urgent procedures.

This study provides robust evidence on the volume–outcome relationship in emergency hip fracture surgery within a non-competitive, highly regulated healthcare system where the “practice makes perfect” mechanism offers a plausible explanation for the observed patterns. It shows that higher lagged surgical volumes are significantly associated with lower 30-day mortality, confirming the effectiveness of Italy’s current minimum-volume threshold and suggesting its clinical relevance.

The findings support the rationale for maintaining and potentially increasing minimum-volume thresholds for urgent procedures, such as hip fracture surgery, to enhance patient outcomes. However, while the results suggest that further centralization could yield additional quality gains, such strategies should be implemented with caution. Excessive concentration of surgical services may compromise access for patients in remote or underserved areas. Policymakers should carefully balance the benefits of higher procedural volumes with the need to ensure timely and equitable access across the population, tailoring thresholds to regional capacities and geography.

5. Discussion

Leveraging a setting where selective referral for urgent surgeries is absent, our multi-way fixed-effects analysis documents a robust negative relationship between lagged ward-level surgical volumes and 30-day mortality after emergency hip fracture surgery. The growing magnitude and statistical significance of the coefficients with longer lags are consistent with a causal effect of accumulated clinical experience on patient outcomes. These results, together with our assessment of the 2016 threshold, reinforce the rationale for volume-based regulatory standards.

Beyond learning-by-doing, other mechanisms may potentially contribute to the negative volume-outcome association. Higher volume hospitals may attract more skilled surgeons [27]. Yet, surgeon fixed effects absorb time-invariant heterogeneity in intrinsic ability, making staffing composition an unlikely driver of the estimated effects. Economies of scale in technology may also play a role. However, hospital fixed effects limit concerns about structural differences, and the relatively short study period makes the possibility of within-hospital technological change limited. We may also expect waiting times to play a mediating role in the volume-outcome nexus, as delays tend to increase in higher-volume hospital wards and adversely affect care quality [6]. In our context, patients require urgent surgery within 48 h of admission [28] and about 80% received surgery within this target window. By controlling for pre-surgery waiting times, estimated volume effects are net of treatment delays.

Nonetheless, policy regulations toward increasing centralization should be carefully evaluated in light of local healthcare capacity and geographic conditions, with particular attention to ensuring equity in access to timely and high-quality care. While our findings support the rationale for volume-based standards, any further centralisation should be calibrated to geography and service capacity, with particular attention to ensuring equity in access to timely and high-quality care. Evidence from mountainous territories and low-density areas points to substantial spatial inequalities in potential access [29,30], and international studies indicate clinical risks linked to longer travel times in obstetrics and cardiology [16,31]. However, importantly for interpreting our results, hip fracture surgery is urgent, but not strictly time-critical: clinical standards target surgery within 2 days, rather than within a few hours as for AMI. Morphology and infrastructure compress travel times, with most residents living on the Po Valley plain and major hospitals aligning along the Via Emilia corridor, served by a dense road and rail network. This configuration mitigates the access penalties that centralization can create in other geographies and for time-sensitive procedures.

6. Conclusions

This study aimed to estimate the impact of surgical volumes on patient outcomes within the Italian region of Emilia-Romagna, where policymakers have prioritized provider cooperation and coordination through the central planning of production capacity. In 2016, Ministerial Decree (MD) 70/2015 introduced minimum-volume requirements of 75 cases per hospital ward for selected surgical procedures including surgeries for emergency hip fractures. Focusing on this latter procedure, which represents the primary cause of admission to emergency departments for orthopaedic interventions, we estimate the effect of increasing lagged volumes on 30-day post-operative mortality.

The results provide empirical support for the rationale behind volume-based standards, showing that higher procedural volumes are associated with better outcomes. By exploiting within-hospital variation over time and conditioning on surgeon and hospital fixed effects as well as pre-surgery waiting times, our estimates largely net out compositional, technological, and congestion-related channels. We interpret our findings as consistent with the dynamic of learning-by-doing, while not fully ruling out other volume-related mechanisms. Our assessment of the 2016 threshold suggests that the current minimum of 75 procedures per year may fall short of fully capturing potential quality gains. Where accessibility is not a limiting factor, higher thresholds may warrant consideration as part of broader efforts to centralise emergency orthopaedic services and further reduce mortality risk.

At the same time, policymakers should weigh the benefits of increased volume against the imperative of equitable access, adjusting thresholds in line with local healthcare capacity and geographic conditions. Overconcentration of surgical services risks reducing access for patients in remote or underserved areas, especially for procedures for which timely intervention is critical. In geographically isolated, sparsely populated areas, low volumes may persist due to unavoidable access constraints rather than inefficiency, and policy solutions may therefore need to rely more on networking, teleconsultation, and shared protocols than on full centralisation.

These considerations imply that volume standards should be region-specific and aligned with travel-time maps and local capacity. For minute-dependent conditions, such as AMI or stroke, centralisation requires efficient transfer and coordination systems. Conversely, for procedures with wider treatment windows, such as hip fracture repair, higher-volume reconfigurations are less likely to compromise timely access aligned with clinical standard targets. Accordingly, the definition of volume thresholds should be procedure-sensitive and explicitly integrate regional geography and capacity assessments, so as to secure quality gains without eroding equity.

Ethics approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. The study, based on routine administrative information, was carried out in compliance with Emilia-Romagna Regional Authority data processing regulations and the Italian Data Protection Act, which has been harmonized with the European General Data Protection Regulation 2016/679 by means of the Legislative Decree 101/2018. Administrative data were anonymized prior to the analysis at the regional statistical office, where each patient is assigned a unique identifier. This identifier does not allow to trace the patient's identity and other sensitive information. Anonymized regional administrative data may be used for retrospective studies, with no specific written patient consent, when the aim is health-care quality evaluation and improvement, which was the primary objective of this analysis. Given the characteristics of the study, no ethical approval was required.

CRediT authorship contribution statement

Sahar Paktinat: Writing – original draft, Software, Methodology, Data curation, Conceptualization. **Cristina Ugolini:** Writing – review & editing, Supervision, Project administration, Methodology, Conceptualization. **Rossella Verzulli:** Writing – review & editing, Supervision, Methodology, Conceptualization.

Declaration of competing interest

The authors have no conflict of interest to declare.

Acknowledgements

The project is the result of the partnership “Economic evaluation of clinical pathways and prescribing appropriateness” between the School of Health Policies-University of Bologna and the Regional Agency for Health and Social Care of the Emilia-Romagna Region (ASSR-RER), with the cooperation of the Local Health Authority of Parma. The ASSR-RER provided the databases used in this study. The opinions expressed here are the sole responsibility of the authors and do not represent the views of the University of Bologna and the ASSR-RER. Access to the SICHER dataset, which facilitated the detailed analyses presented in this study, was granted as part of Paktinat's internship at the Rizzoli Orthopaedic Institute. The dataset was anonymized prior to analysis to ensure data privacy. The resources and expertise provided by the Rizzoli Orthopaedic Institute, coupled with the guidance of esteemed collaborators, were instrumental in advancing this work. We extend our gratitude to Barbara Boldrini of the Rizzoli Orthopaedic Institute for their valuable suggestions. The usual disclaimers apply.

Funding sources: Sahar Paktinat received a PhD scholarship from the University of Bologna, co-financed by the Rizzoli Orthopaedic Institute (grant code 38–033–03-DOT22k2552-3150, CUP J33C22002040009).

Data availability

The data used in this study were released by the data holders (Regional Agency for Health and Social Care of the Emilia-Romagna Region, Italy and Rizzoli Orthopaedic Institute, Italy) under specific data sharing agreements, only for the purpose of this study, and are not publicly available due to privacy restrictions. The data sharing agreements do not permit further sharing or publication of the data. Those interested may seek to obtain data directly from the data holders.

References

[1] Luft HS, Hunt SS, Maerki SC. The volume-outcome relationship: practice-makes-perfect or selective-referral patterns? Health V Res 1987;22(2):157.

[2] Halm EA, Lee C, Chassin MR. Is volume related to outcome in health care? A systematic review and methodologic critique of the literature. Ann Intern Med 2002;137(6):511–20. https://doi.org/10.7326/0003-4819-137-6-200209170-000.

[3] Brekke K, Gravelle H, Siciliani L, Straume O. Patient choice, mobility and competition among health care providers. In: Levaggi R, Montefiori M, editors. Health care provision and patient mobility: health integration in the european union. Milan, Milano: Springer; 2014. p. 1–26.

[4] Gutacker N, Bloor K, Cookson R, Gale CP, Maynard A, Pagano D, et al. Hospital surgical volumes and mortality after coronary artery bypass grafting: using international comparisons to determine a safe threshold. Health V Res 2017;52(2): 863–78. https://doi.org/10.1111/1475-6773.12508.

[5] Blanco B, Kothari A, Blackwell R, et al. “Take the volume pledge” may result in disparity in access to care. Surgery 2017;161(3):837–45. https://doi.org/10.1016/j.surg.2016.07.017.

[6] Rachet-Jacquet L, Gutacker N, Siciliani L. Scale economies in the health sector: the effect of hospital volume on health gains from hip replacement surgery. J Econ Behav Organ 2021;190:704–29. https://doi.org/10.1016/j.jebo.2021.08.014.

[7] Vogel JF, Barkhausen M, Pross CM, Geissler A. Defining minimum volume thresholds to increase quality of care: a new patient-oriented approach using mixed integer programming. Eur J Health Econ 2022;23(7):1085–104. https://doi.org/10.1007/s10198-021-01406-w.

[8] Morche J, Renner D, Pietsch B, Kaiser L, Broenneke J, Gruber S, Matthias K. International comparison of minimum volume standards for hospitals. Health Policy 2018;122(11):1165–76. https://doi.org/10.1016/j.healthpol.2018.08.016.

[9] de Cruppé W, Ortwein A, Kraska RA, Geraedts M. Impact of suspending minimum volume requirements for knee arthroplasty on hospitals in Germany: an uncontrolled before–after study. BMC Health V Res 2020;20:1–13. https://doi.org/10.1186/s12913-020-05957-1.

[10] Rombey T, Goossen K, Breuing J, Mathes T, Hess S, Burchard R, Pieper D. Hospital volume-outcome relationship in total knee arthroplasty: protocol for a systematic review and non-linear dose-response meta-analysis. Syst Rev 2020;9:1–8. https://doi.org/10.1186/s13643-020-01295-9.

[11] Birkmeyer JD, Stukel TA, Siewers AE, Goodney PP, Wennberg DE, Lucas FL. Surgeon volume and operative mortality in the United States. N Engl J Med 2003; 349(22):2117–27. https://doi.org/10.1056/NEJMsa035205.

[12] Mesman R, Faber MJ, Berden BJ, Westert GP. Evaluation of minimum volume standards for surgery in the Netherlands (2003–2017): a successful policy? Health Policy 2017;121(12):1263–73. https://doi.org/10.1016/j.healthpol.2017.09.017.

[13] Avdic D. Improving efficiency or impairing access? Health care consolidation and quality of care: evidence from emergency hospital closures in Sweden. J Health Econ 2016;48:44–60. https://doi.org/10.1016/j.jhealeco.2016.02.002.

[14] Avdic D, Lundborg P, Vikström J. Mergers and birth outcomes: evidence from maternity ward closures. iza discussion papers 11772. Institute of Labor Economics (IZA); 2018.

[15] Avdic D, Lundborg P, Vikstrom J. Estimating returns to hospital volume: evidence from advanced cancer surgery. J Health Econ 2019;63:81–99. https://doi.org/10.1016/j.jhealeco.2018.10.005.

[16] Ghislandi S, Renner AT, Varghese NE. The impact of budget cuts on individual patient health: causal evidence from hospital closures. J Health Econ 2025;101: 102975. https://doi.org/10.1016/j.jhealeco.2025.102975.

[17] Shichman I, Roof M, Askew N, et al. Projections and epidemiology of primary hip and knee arthroplasty in Medicare patients to 2040-2060. JB JS Open Access 2023; 8(1). https://doi.org/10.2106/JBJS.OA.22.00112.

[18] Chowdhury MM, Dagash H, Pierro A. A systematic review of the impact of volume of surgery and specialization on patient outcome. Br J Surg 2007;94(2):145–61. https://doi.org/10.1002/bjs.5714.

[19] Hentschker C, Mennicken R. The volume–outcome relationship revisited: practice indeed makes perfect. Health V Res 2018;53(1):15–34. https://doi.org/10.1111/1475-6773.12696.

[20] Hamilton BH, Hamilton VH. Estimating surgical volume—outcome relationships ap- plying survival models: accounting for frailty and hospital fixed effects. Health Econ 1997;6(4):383–95. https://doi.org/10.1002/(SICI)1099-1050(199707)6:4≤383::AID-HEC278≥3.0.CO;2-L.

[21] Rachet-Jacquet L. Do breaks from surgery improve the performance of orthopaedic surgeons? J Health Econ 2022;85:102667. https://doi.org/10.1016/j.jhealeco.2022.102667.

[22] Van Gestel R, Broekman N, Mueller T. Surgeon supply and healthcare quality: are revision rates for hip and knee replacements lower in hospitals that employ more surgeons? Health Econ 2023;32(10):2298–321. https://doi.org/10.1002/hec.4727.

[23] Di Novi C, Piacenza M, Robone S, Turati G. Does fiscal decentralization affect regional disparities in health? Quasi-experimental evidence from Italy. Reg Sci Urban Econ 2019;78(10):103465. https://doi.org/10.1016/j.regsciurbeco.2019.103465.

[24] American Society of Anesthesiologists. Statement on ASA physical status classification system. https://www.asahq.org/standards-and-practice-parameters/statement-on-asa-physical-status-classification-system; 2020.

[25] National Agency for Regional Healthcare Services (Agenas) (2020). Agenas report pne 2020. Accessed: August 21, 2024.

[26] Eurostat. (2018). Methodological manual on territorial typologies—2018 edition, https://doi.org/10.2785/930137. Available from: https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-GQ-18-008.

[27] Di Giacomo M, Piacenza M, Salmasi L, Turati G. Understanding labour productivity in maternity wards. Labour Econ 2025;96:102760. https://doi.org/10.1016/j.labeco.2025.102760.

[28] NICE. Hip fracture management, clinical guidelines. Available at, www.nice.org.uk/guidance/cg124; 2011.

[29] Perucca G, Piacenza M, Turati G. Spatial inequality in access to healthcare: evidence from an Italian Alpine region. Reg Stud 2019;53(4):478–89. https://doi.org/10.1080/00343404.2018.1462481.

[30] AGENAS. Accessi in Pronto Soccorso e Implementazione DM 77/2022 per una migliore presa in carico dei pazienti (in Italian). Available at, https://www.agenas.gov.it/images/agenas/comunicati/2024/Comunicato_stampa_22_aprile.pdf; 2024.

[31] Malouf R, Henderson J, Alderdice F. Impact of obstetric unit closures, travel time and distance to obstetric services on maternal and neonatal outcomes in high-income countries: a systematic review. BMJ Open 2020;10:e036852. https://doi.org/10.1136/bmjopen-2020-036852.