

Alma Mater Studiorum Università di Bologna  
Archivio istituzionale della ricerca

Risk-Adjusted Mortality Rates as a Quality Proxy Outperform Volume in Surgical Oncology-A New Perspective on Hospital Centralization Using National Population-Based Data

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

*Published Version:*

Baum P., Lenzi J., Diers J., Rust C., Eichhorn M.E., Taber S., et al. (2022). Risk-Adjusted Mortality Rates as a Quality Proxy Outperform Volume in Surgical Oncology-A New Perspective on Hospital Centralization Using National Population-Based Data. JOURNAL OF CLINICAL ONCOLOGY, 40(10), 1041-1050 [10.1200/JCO.21.01488].

*Availability:*

This version is available at: <https://hdl.handle.net/11585/886857> since: 2022-05-24

*Published:*

DOI: <http://doi.org/10.1200/JCO.21.01488>

*Terms of use:*

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).  
When citing, please refer to the published version.

(Article begins on next page)

# Risk-adjusted mortality rates as a quality proxy outperform volume in surgical oncology – a new perspective on hospital centralization using national population-based data

Philip Baum M.D.<sup>1,2</sup>; Jacopo Lenzi PhD<sup>3</sup>; Johannes Diers M.D. MScPH<sup>2</sup>; Christoph Rust M.Sc.<sup>4,5</sup>; Martin E. Eichhorn M.D.<sup>1,6</sup>; Samantha Taber M.D.<sup>7</sup>; Prof. Christoph-Thomas Germer<sup>2,8</sup>; Prof. Hauke Winter M.D.<sup>1,6</sup>; Armin Wiegering M.D.<sup>2,8,9</sup>

<sup>1</sup> Department of Thoracic Surgery, Thoraxklinik at Heidelberg University Hospital, Roentgenstrasse 1, 69126 Heidelberg, Germany

<sup>2</sup> Department of General, Visceral, Transplant, Vascular and Pediatric Surgery, University Hospital Wuerzburg, Oberduerrbacher Strasse 6, 97080 Wuerzburg, Germany

<sup>3</sup> Department of Biomedical and Neuromotor Sciences, University of Bologna, Via San Giacomo 12, 40126 Bologna, Italy

<sup>4</sup> Department of Econometrics, University of Regensburg, Universitaetsstrasse 31, 93040 Regensburg, Germany

<sup>5</sup> Department of Finance, Accounting and Statistics, Vienna University of Economics and Business, Welthandelsplatz 1 1020 Vienna, Austria

<sup>6</sup> Translational Lung Research Center Heidelberg (TLRC), German Center for Lung Research (DZL), Im Neuenheimer Feld 156, 69120 Heidelberg, Germany

<sup>7</sup> Department of Thoracic Surgery, Heckeshorn Lung Clinic, HELIOS Klinikum Emil von Behring, Waltherhöferstrasse 11, 14165 Berlin, Germany

<sup>8</sup> Comprehensive Cancer Center Mainfranken, University of Wuerzburg, Oberduerrbacher Strasse 6, 97080 Wuerzburg Germany.

<sup>9</sup> Theodor Boveri Institute, Biocenter, University of Wuerzburg, Am Hubland, 97074 Würzburg Würzburg, Germany

## Research support:

none

## Corresponding author:

Philip Baum

Department of Thoracic Surgery

Thoraxklinik at Heidelberg University Hospital, Roentgenstraße 1, 69126 Heidelberg, Germany

004915122993638, philip.baum@med.uni-heidelberg.de

**The study as a whole or in part has never been never presented before.**

## Disclaimer:

All authors declare that there are no conflicts of interest.

## Running head:

Risk-adjusted mortality outperforms volume as proxy for quality

## Abstract

### PURPOSE

Despite a long-known association between annual hospital volume and outcome, little progress has been made in shifting high-risk surgery to safer hospitals. This study investigates whether the risk-standardized mortality rate (RSMR) could serve as a stronger proxy for surgical quality than volume.

### METHODS

We included all patients who underwent complex oncologic surgeries in Germany between 2010 and 2018 for any of 5 major cancer types, splitting the data into training (2010–2015) and validation sets (2016–2018). For each surgical group, we calculated annual volume and RSMR quintiles and studied the overlap between the two systems. We modelled a market exit of low-performing hospitals and compared effectiveness and efficiency of volume- and RSMR-based rankings. We compared travel distance/time that would be required to reallocate patients to the nearest hospital with low-mortality rankings.

### RESULTS

Between 2016 and 2018, 158,079 patients were treated in 974 hospitals. At least 50% of high-volume hospitals were not ranked in the low-mortality group according to RSMR grouping. In a RSMR centralization model, an average of 32 patients undergoing complex oncologic surgery would need to relocate to a low-mortality hospital to save one life, while 47 would need to relocate to a high-volume hospital. Mean difference in travel times between the nearest hospital to the hospital that performed surgery ranged from 10 minutes for colorectal cancer to 24 minutes for pancreatic cancer. Centralization based on RSMR compared to volume would ensure lower median travel times for all cancer types, and these times would be lower than those actually observed.

### CONCLUSION

RSMR is a promising proxy for measuring surgical quality. It outperforms volume in effectiveness, efficiency, and hospital availability for patients.

## Introduction

Surgical oncology is a fundamental therapy for the majority of solid cancer types. As shown in our and other previous works, hospitals vary greatly in their ability to perform complex cancer surgery safely, with large inter-hospital differences in rates of complication, failure to rescue, and mortality [1-7]. Annual hospital volume is commonly used as a surrogate parameter for measuring and theoretically improving surgical quality. Despite the establishment of an association between volume and outcome more than 40 years ago, a consistent adverse volume-outcome relationship is still debated [8-11]. The topic remains controversial since almost all evidence comes from retrospective or prospective observational data. Quality improvement programs in Europe and North America recommend implementing clear policies for national health care centralization. To date, however, there has been little progress in moving complex surgical care to high-volume hospitals [12]. This likely reflects several barriers to volume-based realignment. For one, there are relatively few hospitals that qualify as “high-volume” for complex surgery, presenting logistical challenges for many patients [13]. Second, volume is not the only factor in determining hospital quality—other important factors include experience and training of the surgical staff, the frequency with which individual surgeons perform complex operations, the availability of a multidisciplinary team, and also hospital-specific characteristics such as skill in managing complications, academic affiliation, and the experience and training of its nursing staff [1, 14, 15]. Finally, due to the heterogeneity of outcomes at high- and low-volume hospitals, prohibitively large numbers of patients would need to change hospitals to derive a significant patient benefit [15, 16].

An alternative means of assessing quality in surgical oncology is using a measure that considers outcome quality rather than caseload volumes alone. Risk-standardized mortality rates (RSMRs) are a healthcare indicator used for comparing the performance of regions, health systems or hospitals. It accounts for differences in case mix and is validated using administrative claims data [17]. RSMR is defined as the ratio of a given hospital’s “predicted” mortality rate to its “expected” mortality rate, based on hierarchical logistic regression analysis [18]. Thus, RSMR might serve as a stronger proxy for surgical quality than annual hospital volume.

In this study, we investigate the effectiveness and efficiency of RSMR-based ranking of hospitals and compare it to volume-based ranking. The primary endpoint was the number of patients needed to move from low- to high-performing hospitals to save one life, stratified by RSMR- or volume-based regionalization. The secondary endpoint was the difference in travel times required to move patients from low-performing hospitals, depending on whether volume-based or RSMR-based rankings were used as the standard of quality.

## Methods

### Design, setting, population, and data sources

This is a register-based, retrospective cohort study of national hospital discharge data on every adult inpatient who underwent cancer surgery in Germany between January 2010 and December 2018. Diagnosis-related group (DRG) data were analyzed through the Federal Statistical Office in accordance with the legal data protection regulations of Germany (Research Data Centers of the Federal Statistical Office and the Statistical Offices of the States). Due to complete anonymity, no approval from Heidelberg University's ethics committee was required. The DRG statistics represent an almost complete dataset, encompassing all hospital treatments in acute inpatient somatic care in Germany [19]. In addition to demographic data (age, gender), the data contain various administrative information on each treatment case (cause of admission, type of discharge, principal and secondary diagnoses according to the International Classification of Diseases, German Modification [ICD-10-GM], procedures according to the German Operation and Procedure Code and number of hours of mechanical ventilation) [20]. Since regional information, both on hospital location and patient residence, are included, the DRG data can also be used for regional analyses [21]. Clinical variables (e.g., TNM classification, tumor histology, or ASA status) are not available, as these are not billing-relevant.

The administrative data we accessed included surgical procedures, diagnoses, sex, age, in-hospital mortality, and length of stay for individual patients; it also included aggregated localization data of patients and hospitals. Germany's administrative hospital billing data do not contain long-term data or clinical variables like histology, UICC stadium, or adjuvant therapy. Cases that were non-elective or involved patients younger than 18 years were excluded. Hospitals that had only one intervention during the study period were also excluded (Supplemental Figure S1).

Surgeries were stratified into 5 groups, based on the German procedure classification ("Operationen- und Prozedurenschlüssel"). These included oncological resections of the esophagus (*esophageal resection for esophageal cancer*), lung (*lobectomy, bilobectomy, pneumonectomy for lung cancer*), stomach (*partial/subtotal/complete gastric resection for gastric cancer*), pancreas (*partial/complete resection of the pancreas for pancreatic cancer*), and colon/rectum (*partial/total resection of the colon and sphincter, preserving/non-preserving rectum resection for colorectal cancer*). All data were inspected to remove typos, assess proper format variables, and screen for missing values or duplicate patient profiles. The RECORD reporting guidelines for studies based on routinely collected health data were applied [22].

### Training and validation data

To eliminate the risk of overfitting and resubstitution bias, the RSMR- and volume-based rankings described below were developed on a training subsample constituting 2/3 of the whole dataset, while all subsequent analyses (ranking agreement, restricted cubic splines, effectiveness and efficiency of centralization strategies, and travel routes) were conducted on a validation subsample constituting the remaining 1/3. Time of surgery was used as the basic criterion for allocating patients: those operated on in the first 6 years (2010–2015) constituted the training set, while those operated on in the last three years (2016–2019) constituted the validation set. This split, albeit non-random, ensures the correct evaluation of volume cutoffs to define hospital groups and reflects what happens on a routine basis when data-driven policies are put in place, i.e., data from previous years are utilized to address new strategies of intervention in the future.

Hospital performance: based on surgical volume, based on risk-standardized mortality rates

Hospital performance was determined by calculating both the average annual caseloads and the RSMRs for each hospital. Two distinct rankings were produced. A description of the estimation process that we used to obtain the RSMRs is provided in the next subsection.

Hospitals in the lowest quintile and highest quintile of annual caseloads were classified as “low-volume” and “high-volume”, respectively; hospitals in the lowest quintile and highest quintile of RSMR were classified as “low-mortality” and “high-mortality”, respectively. Hospitals lying between the lower and upper quintiles were classified as mid-volume or mid-mortality. For the sake of convenience, hospitals classified as either low-volume or high-mortality are generically referred to as “low-performing”, while those classified as either high-volume or low-mortality are generically referred to as “high-performing”.

Agreement between the two ranking systems was explored visually with the aid of bar charts. Additionally, we performed a restricted cubic spline analysis to model the relationship between individual surgical volumes and RSMRs by using the knot locations recommended by Harrell (2001) [10, 23]. In a sensitivity analysis, cut-points were defined using quartiles in place of quintiles.

### Risk-standardized mortality rates

The RSMRs for each hospital were obtained as the ratio of predicted-to-expected in-hospital mortality, multiplied by the national unadjusted rate. “Predicted” rates were used in place of “observed” rates to avoid several analytical problems that have been cited in the literature [24-26]. Both predicted and expected estimates of in-hospital mortality express the number of in-hospital deaths that would occur if the “standard” event rates, based on the provider case mix, actually occurred. In addition, predicted mortality includes a hospital-specific effect, which represents the baseline mortality risk within the hospital and is a function of the underlying differences in quality of care among healthcare facilities [17]. Operationally, predicted and expected outcomes were estimated via hierarchical logistic regression analysis, which models the log-odds of mortality as a function of patient demographics, clinical characteristics, and a random hospital-specific effect (random intercept). In such models, fixed-effects regression coefficients capture the effect of patient characteristics on the outcome across all hospitals, while hospital-specific effects are a random component arising from a statistical distribution that describes variation in performance among centers [18, 27]. More specifically, expected probabilities were derived for each patient by combining the regression coefficients of the patient covariates selected for risk-adjustment (sex, age [<40, 41–50, 51–60, 61–70, 71–80, >80 y], and the Charlson Comorbidity Index) [28]. These measurements from individual patient records were then added up to derive the expected number of deaths for each hospital. Predicted mortality was obtained in a similar way, but also included the random hospital-effect to reach the final estimate for each individual patient. To avoid over-adjustment, variables that were proxies for quality of care were not included as risk-adjustment covariates in the regression models [18].

All calculations were performed using Stata 15 software (StataCorp. 2017. Stata Statistical Software: Release 15. College Station, TX: StataCorp LLC). A replication of the analyses in the Bayesian framework with the bayesmh command was unfeasible as the sample sizes were too large to run the Markov chain Monte Carlo estimation algorithm.

### Effectiveness and relative efficiency

We evaluated the impact of redistributing cancer patients from low-performing to high-performing hospitals, assuming that these patients would conform to the mortality rates observed in the high-performing class [16]. We evaluated both the volume-based model and the RSMR-based model in terms of destination hospitals, lives saved, patients moved, and patients moved per life saved (relative efficiency). Ninety-five percent

confidence intervals (95% CIs) for lives saved and relative efficiency were calculated using the Wilson score method [29].

### Changes in travel times and distances

A simple closure-scenario was hypothesized to analyze the impact of volume- and RSMR-based redistribution policies on patient travel times and distances. We modelled a simultaneous market exit of all low-performing hospitals, i.e., those not achieving the required threshold of annual caseload or RSMR [21]. Affected patients were then allocated to the next nearest hospital from their place of residence that was still “open” for treatment. Because we were interested in changes in access based on volume and adjusted-mortality regulations, we based our travel time calculations on the surgical oncology hospital that was closest to the patient’s home, irrespective of whether he or she was actually treated there. However, in a secondary analysis we compared minimum versus actual travel times and distances. All comparisons were performed with the paired *t*-test. Patients with missing geocoding were excluded (2.3%) (esophagus: *n* = 144; lung: *n* = 1078; stomach: *n* = 291; pancreas: *n* = 338; colon and rectum: *n* = 1804). Since travel time is dependent on structural and geographical conditions, we conducted a subgroup analysis stratified by all counties in Germany. Stratified results were used as a proxy for effectiveness in patient realignment, reflected in the differences between average volume- and RSMR-based redistribution for the country. Data were visualized using Microsoft Excel 2019.

For the travel time analysis, we calculated driving times and distances in minutes and km, respectively, taking geographic and infrastructural differences into account. This point is often neglected but is especially important for rural areas, because straight-line measurements can underestimate travel time in regions with less comprehensive infrastructure. Methodological details are provided in the next subsection.

### Geocoding and travel routes

The addresses of all German hospitals were obtained from the German hospital directory and were geocoded using the service provided by LocationIQ. Since geocoding is not without errors, we performed data-quality checks based on the following heuristics: (a) the address was matched exactly, or the hospital name was part of the coded result; (b) the name of the match contained “Klinik” or similar wording; (c) a fraction was checked manually by querying Google Maps. The final geocodes were then aggregated at the ZIP code level. If a single ZIP code corresponded to multiple hospitals, we defined the spatial mean of the locations as the coordinate of the corresponding ZIP code. The Stata plugin *osrmtime* was used to estimate travel routes between the hospital coordinates and the aggregated patient coordinates, the latter being based on the German 9-digit community key (AGS “Amtlicher Gemeindeschlüssel”), updated on December 31, 2019 [30]. The result was a final routing matrix of all possible patient–hospital distances in Germany. In a second step, the matrix was merged with the DRG data.

Since the AGS-9 was reformed several times during the observation period, we include the following considerations. (A) The main patient coordinate data were taken from 2019. (B) In cases of multiple entries for one AGS-code with different geolocations, we took the AGS-code from the largest group with respect to WGS84 coordinates (3 digits). (C) In remaining cases of multiple geolocations for single AGSs, we selected the group with the smallest distance to the mean of the three groups, as long as the average distance to the mean was not larger than 1000 meters. Our results of mean distances to closest health care providers are in line with the results given in the current literature, although different methodical approaches have been applied [31].

Geocoding was conducted using Phyton (Python Software Foundation. Python Language Reference, version 3.8) and R (R Core Team. 2020. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing).

## Results

A total of 487,911 patients from the German national administrative database, treated between 2010 and 2018, were included in the study: 329,832 (67.6%) were operated on between 2010 and 2015 and were included in the training set to develop RSMR- and volume-based hospital rankings, while 158,079 (32.4%) were operated on between 2016 and 2018 and were included in validation set used to test their effectiveness and efficiency. The 158,079 patients used in validation analysis had a mean age of  $68.6 \pm 11.4$  years and were treated in 974 different hospitals (Table 1). After stratification according to cancer location, colorectal cancer was the most frequently represented cancer type ( $n = 101,846$  [64.4%]). Overall, in-hospital mortality was 3.8%, ranging from 2.9% for lung cancer surgery to 7.6% for pancreatic cancer surgery. Most patients were hospitalized in urban (50.3%) or suburban (30.3%) areas. As shown in Supplementary Table S1, the characteristics of the 329,832 patients used in training analysis were very similar to those of the validation sample. All the results that follow refer to patients and hospitals included in the validation dataset.

The characteristics of hospital volume quintiles are shown in Supplementary Table S2. Hospitals in the lowest quintile, based on average annual caseload, ranged from a median of 2.0 resections (stomach) to a median of 13.7 resections (colon and rectum) per year. Mortality decreased with increasing hospital volume; in particular, the rate ranged from 4.0% (lung) to 8.9% (esophagus) and 9.9% (pancreas) in low-volume hospitals, and from 1.9% (lung) to 4.7% (pancreas) and 5.2% (esophagus) in high-volume hospitals. Coherent baseline characteristics of RSMR quintiles are presented in Supplementary Table S3. In the RSMR quintile category with lowest mortality, mortality ranged from 1.4% (lung) to 4.4% (pancreas) and 4.6% (stomach), respectively. For any of the 5 cancer types, the range of raw mortality was wider across RSMR quintiles than across volume quintiles.

In our analysis of hospital classification overlap based on either case volume or RSMR, we studied the agreement of the two approaches (Figure 1). The three hospital quintiles lying between the lowest and highest quintile were classified as either mid-volume or mid-mortality. Few hospitals were ranked as high-volume health care providers. We found strong variability in agreement between RSMR-based ranking and volume-based ranking. At least 50% of high-volume hospitals, with some variation dependent on the type of surgery considered, were not ranked within the low-mortality group according to RSMR classes. Also, the percentage of high-mortality hospitals in the medium- and the high-volume classes were comparable (Figure 1). The modest increase in RSMRs in mid- and low-volume hospitals was confirmed by a restricted cubic spline analysis (Supplementary Figure S2). Of note, the mild linear correlation between individual volumes and RSMRs was stronger for esophageal and pancreatic surgery compared to the other surgeries.

Next, we evaluated the effectiveness and efficiency of RSMR-based and volume-based approaches to moving cancer patients away from low-performing hospitals, classified as either low volume or high RSMR. Here we assumed that patients who were moved would achieve the mortality rates observed in the corresponding high-performing class (Table 2). In the RSMR-based centralization model, the number of patients needed to switch hospitals in order to save one life varied according to cancer location. For surgeries of the pancreas, 14 patients would need to move; for esophageal cancers 19 patients; for gastric cancers 26 patients; for lung cancers 33 patients; and for colorectal cancers 38 patients. In the volume-based approach, the numbers also varied according to cancer location, but consistently more patients would need to switch hospitals in order to save one life. Moreover, a realignment based on RSMRs would save a total of 955 lives (95% CI 895–1015), while a realignment based on volumes would save a total of 663 lives (95% CI 606–720) (Figure 2). When we split hospital groups using quartiles in place of quintiles, these results were confirmed (Supplementary Table S5).

Moving patients and restricting operative capacity to fewer hospitals could potentially mean longer travel times for patients and their relatives. For adequate interpretation of this model, we checked the real-world situation by analyzing the time required for patients to travel to the hospital where treatment actually occurred (Supplementary Table S4). Median driving time (by car) ranged from 14 to 27 minutes depending on site of cancer (colon/rectum and esophagus, respectively), suggesting that patients did not intuitively choose the



nearest hospitals. In fact, actual travel time to the hospital chosen by patients compared to travel time to the nearest bypassed hospital differed between 10 and 22/24 minutes on average (colon/rectum and esophagus/pancreas, respectively). Table 3 shows changes in patient travel times and distances when they are reassigned to higher-performing hospitals according to the volume-based centralization model and the RSMR-based centralization model. A RSMR-based approach leads to shorter travel times and distances than a volume-based approach.

Strikingly, adherence to RSMR-based centralization would lower the actual median patient travel time for the complex cancer surgeries investigated here (esophagus: 17 vs. 27 min; lung: 15 vs. 26 min; stomach: 9 vs. 16 min; pancreas: 13 vs 23 min; colon/rectum: 9 vs. 14 min) (all  $P$ -values <0.001).

Finally, we examined hospital availability depending on realignment model, on the basis of individual place of residence. Because of natural variation in traffic infrastructure and hospital density, estimates were stratified by county. Figure 3 presents the comparative effectiveness of the RSMR-based versus the volume-based approaches to centralization in Germany (mean travel duration on the basis of RSMR centralization minus mean travel duration on the basis of volume centralization). There were only a few counties for which the volume-based centralization model would result in shorter travel times (esophagus: 32 [8.0%]; lung: 21 [5.2%]; stomach: 48 [12.0%]; pancreas: 60 [14.7%]; colon and rectum: 70 [17.5%] counties).

## Discussion

This study offers a new perspective on volume-based hospital centralization approaches, underscoring the potential for RSMR-based ranking to succeed where volume-based ranking has failed. Our data reveal only a partial overlap between mortality estimates based on surgical volume vs. RSMR rankings. It suggests that RSMR-based centralization is a more efficient method of classifying hospitals and thus of protecting patients from unnecessary risk in high-mortality hospitals. We demonstrate that a risk-adjusted approach to measuring surgical quality outperforms the currently more widely accepted volume-based ranking approach [32]. RSMRs allow for a clearer differentiation between the quality of care provided by individual hospitals and show that lower-volume institutions can also offer high-quality care. Our study also shows that many patients, seeking specialized cancer surgery, bypass their nearest service provider. Interestingly, patient reassignment based on RSMR rankings seems to theoretically result in patients having shorter travel times than they currently do in almost all German counties.

Finding the optimal proxy for surgical quality is crucial for measuring and comparing surgical outcomes and ultimately for improving the care of patients undergoing complex, high-risk oncologic surgeries. The great challenge, however, is that surgical quality is a multidimensional parameter that cannot be defined by a single best measure [33]. In 2018, the European Surgical Association presented several recommendations for an effective hospital centralization program, aiming for a balanced rather than a purely market driven approach [12]. Despite the effectiveness promise of volume-based approaches, they have been criticized as unrealistic and creating inequality due to increased travel distances [10, 34]. Here, we offer an alternative approach that not only exhibits superior results but is also more feasible for patients and their relatives. RSMRs are an objective measure of hospital quality. A workable analysis and method of ranking, however, must meet two important requirements. First, data should cover a whole population-based healthcare system. This is to minimize type I errors, as otherwise preselection of a nonrepresentative cohort could level out real-world effects [10]. Second, although RSMR calculations may seem trivial at first glance, as we have shown, several potential pitfalls must be taken into account to prevent errors in selecting the optimal method for risk-adjustment from leading to biased results [18] [35]. In addition, our training and validation split ensured the correct evaluation of performance cutoffs for defining hospital groups, thus reflecting what routinely happens when data-driven policies are implemented.

So far, countries with volume-based centralization approaches have achieved promising results, for example with an increase in short-term survival rates of between 39% to 55% for pancreatic surgery in the Netherlands and reduced postoperative mortality after esophagectomies in England (compared to health systems with no volume pledges at the time of comparison) [36] [37]. So far only a few western countries (Austria, France, Germany, Spain, Switzerland, the Netherlands) have implemented legally enforced volume thresholds to address the volume-outcome relationship. Twenty-six esophageal and 10 pancreatic resections per year are required in Germany. Our data do not support the view that volume-pledge is ineffective in reducing mortality, but they highlight the drawbacks of using volume classes to distinguish between “high-performing” and “low-performing” hospitals [34]. Most significantly, we have demonstrated that high-volume hospitals do not per se lead to the best achievable mortality rates in high-risk oncological resections. On the flip side, few low-volume centers were grouped within the low-mortality class. Compared to volume ranking, RSMR incorporates multiple patient and provider characteristics, rather than relying on the singular metric of volume threshold alone. RSMR is procedure specific, and our realignment model showed that the numbers of patients needed to change hospitals to save one life were smaller for all investigated oncologic surgeries.

A number of obstacles to hospital centralization remain. They include structural barriers like lack of specialists or long waiting times, political barriers like regional interests in maintaining local control and financial benefits, and patient barriers like unwillingness to travel greater distances or lack of awareness of outcome differences [12]. One objection to centralization is that it is not patient centered. It is assumed that most patients, after all, prefer to receive care closer to home and do not always want to accept increases in travel distances [34]. From a patient perspective, realignment needs to take these concerns into consideration. Our travel analysis,

however, shows that patients are ordinarily willing to bypass their closest service provider if it is for a high-risk operation. This finding supports known evidence that patients are willingly bypassing their nearest cancer center to receive surgery at a more distant hospitals that they hope will better meet their needs [38]. Moreover, our data suggest that hospital centralization does not automatically result in increased patient travel time, and, interestingly, in many cases it might even lead to patients traveling shorter distances than they currently do. The strength of our geo-analysis is that it reflects real-world distances and driving times by car, taking geographic and infrastructural differences into account. This point is often forgotten but is quite important as straight-line measurements tend to underestimate reality, especially in rural regions where hospital availability is an important factor. Two thirds of the German population can reach the next hospital in 10 minutes or less by car; 97.5% can reach it within 20 minutes [39]. Availability, however, is limited by the fact that not every hospital can treat every specific condition, especially conditions that require complex cancer surgery.

Finding a balance between legal centralization, provider competition and patient choice is difficult and our analysis does not address this problem directly. In theory, RSMR centralization could allow patients to choose lower-mortality hospitals within a certain radius that still would not require them to travel too far from home. In reality, treatment choice is complex and might often be more dependent on reputation or availability of particular treatment technique than on the mortality performance of the actual hospital [40]. For example, men with prostate cancer might be attracted to centers that carry out robotic surgery and employ well-regarded surgeons, whether or not they perform well in more objective measures of quality [41]. How patient choice affects outcomes in elective surgery remains an open question [42]. Evidence is scant but suggests that publicly releasing performance data may stimulate quality improvement activity at the hospital level [43]. Whether public reporting of RSMRs could impact effectiveness, safety, and patient-centeredness remains a future question.

In addition to the well-known pitfalls of observational investigations, our study has additional limitations. Some of these are common to all analyses based on health administrative data. First, we did not have access to important clinical variables such as TNM-stage or histology. Other limitations include potential lack of accuracy and differences in the coding criteria over time and across providers. Moreover, the lack of unique record-linkage information did not allow us to track deaths that occurred after hospital discharge and made our outcome assessment largely dependent on differences in length of stay across institutions. Another important limitation is that travel analyses always remain artificial. Due to data protection rules we had to rely on geographical centroids of patient residence counties and could not account for modes of transportation other than cars and their availability to individual patients. In addition, our results need to be validated internationally, especially in different national health systems and nations with high geographic distances between hospitals and patient homes. Finally, it should be mentioned that high-risk surgery accounts for approximately 5% of all surgeries performed, meaning that any discussion of centralization leaves the vast majority of (low-risk) surgeries untouched [7]. This should be of comfort to surgeons, who may feel threatened by discussions of centralization and worry about losing their patients.

In conclusion, our analysis indicates that RSMR-based quality management is superior to volume thresholds. It outperforms a volume-based approach in terms of effectiveness and hospital availability, making it a better proxy for defining performance in complex oncological surgeries.

**Acknowledgements:**

We thank J. Loske and K. Stockmeyer from the German Federal Statistical Office for their kind help with the data acquisition and management.

## Bibliography

1. Birkmeyer, J.D., et al., *Surgeon Volume and Operative Mortality in the United States*. New England Journal of Medicine, 2003. **349**(22): p. 2117-2127.
2. Diers, J., et al., *Nationwide in-hospital mortality following colonic cancer resection according to hospital volume in Germany*. BJS open, 2019. **3**(5): p. 672-677.
3. Diers, J., et al., *Nationwide in-hospital mortality rate following rectal resection for rectal cancer according to annual hospital volume in Germany*. BJS Open, 2020. **4**(2): p. 310-319.
4. Baum, P., et al., *Nationwide effect of high procedure volume in lung cancer surgery on in-house mortality in Germany*. Lung Cancer, 2020. **149**: p. 78-83.
5. Nimptsch, U., et al., *Nationwide In-hospital Mortality Following Pancreatic Surgery in Germany is Higher than Anticipated*. Ann Surg, 2016. **264**(6): p. 1082-1090.
6. Diers, J., et al., *Hospital volume following major surgery for gastric cancer determines in-hospital mortality rate and failure to rescue: a nation-wide study based on German billing data (2009-2017)*. Gastric Cancer, 2021.
7. Baum, P., et al., *Mortality and Complications Following Visceral Surgery*. Dtsch Arztebl International, 2019. **116**(44): p. 739-746.
8. Luft, H.S., J.P. Bunker, and A.C. Enthoven, *Should operations be regionalized? The empirical relation between surgical volume and mortality*. N Engl J Med, 1979. **301**(25): p. 1364-9.
9. Begg, C.B., et al., *Impact of hospital volume on operative mortality for major cancer surgery*. JAMA, 1998. **280**(20): p. 1747-51.
10. Farjah, F., et al., *Volume Pledge is Not Associated with Better Short-Term Outcomes After Lung Cancer Resection*. J Clin Oncol, 2020. **38**(30): p. 3518-3527.
11. LaPar, D.J., et al., *Hospital procedure volume should not be used as a measure of surgical quality*. Ann Surg, 2012. **256**(4): p. 606-15.
12. Vonlanthen, R., et al., *Toward a Consensus on Centralization in Surgery*. Ann Surg, 2018. **268**(5): p. 712-724.
13. Dimick, J.B., S.R. Finlayson, and J.D. Birkmeyer, *Regional availability of high-volume hospitals for major surgery*. Health Aff (Millwood), 2004. **Suppl Variation**: p. VAR45-53.
14. Karanicolas, P.J., et al., *The more the better?: the impact of surgeon and hospital volume on in-hospital mortality following colorectal resection*. Ann Surg, 2009. **249**(6): p. 954-9.
15. Birkmeyer, J.D. and J.B. Dimick, *Potential benefits of the new Leapfrog standards: effect of process and outcomes measures*. Surgery, 2004. **135**(6): p. 569-75.
16. Chiu, A.S., et al., *Quality Versus Quantity: The Potential Impact of Public Reporting of Hospital Safety for Complex Cancer Surgery*. Ann Surg, 2019. **270**(2): p. 281-287.
17. Krumholz, H.M., et al., *An administrative claims model suitable for profiling hospital performance based on 30-day mortality rates among patients with heart failure*. Circulation, 2006. **113**(13): p. 1693-701.
18. Lenzi, J. and S. Pildava, *Tips for calculating and displaying risk-standardized hospital outcomes in Stata*. Stata Journal, 2019. **19**(2).
19. Busse, R., U. Nimptsch, and T. Mansky, *Measuring, monitoring, and managing quality in Germany's hospitals*. Health Aff (Millwood), 2009. **28**(2): p. w294-304.
20. *Forschungsdatenzentren der statistischen Ämter des Bundes und der Länder: Datenangebot. Fallpauschalenbezogene Krankenhausstatistik (DRG-Statistik)*. Wiesbaden: Forschungsdatenzentren der statistischen Ämter des Bundes und der Länder. [cited 2021 09.06]; Available from: [www.forschungsdatenzentren.de/bestand/drg/index.asp](http://www.forschungsdatenzentren.de/bestand/drg/index.asp)
21. Hentschker, C., et al., *Volume-outcome relationship and minimum volume regulations in the German hospital sector - evidence from nationwide administrative hospital data for the years 2005-2007*. Health Econ Rev, 2018. **8**(1): p. 25.
22. Benchimol, E.I., et al., *The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement*. PLoS Med, 2015. **12**(10): p. e1001885.
23. Gauthier, J., Q.V. Wu, and T.A. Gooley, *Cubic splines to model relationships between continuous variables and outcomes: a guide for clinicians*. Bone Marrow Transplant, 2020. **55**(4): p. 675-680.
24. Goldstein, H. and D.J. Spiegelhalter, *League Tables and Their Limitations: Statistical Issues in Comparisons of Institutional Performance*. Journal of the Royal Statistical Society. Series A (Statistics in Society), 1996. **159**(3): p. 385-443.
25. Normand, S.-L.T., M.E. Glickman, and C.A. Gatsonis, *Statistical Methods for Profiling Providers of Medical Care: Issues and Applications*. Journal of the American Statistical Association, 1997. **92**(439): p. 803-814.

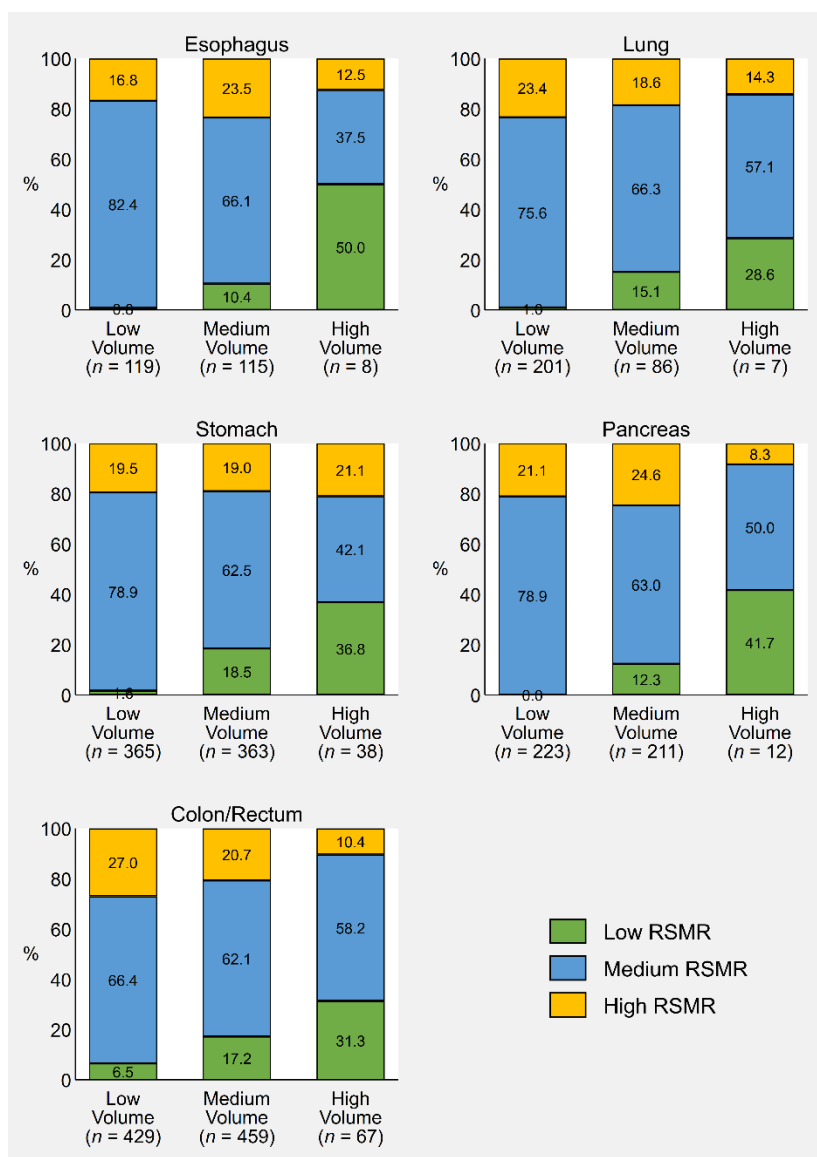
26. Christiansen, C.L. and C.N. Morris, *Improving the statistical approach to health care provider profiling*. Ann Intern Med, 1997. **127**(8 Pt 2): p. 764-8.
27. MacKenzie, T.A., et al., *A primer on using shrinkage to compare in-hospital mortality between centers*. Ann Thorac Surg, 2015. **99**(3): p. 757-61.
28. Quan, H., et al., *Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data*. Med Care, 2005. **43**(11): p. 1130-9.
29. Bender, R., *Calculating confidence intervals for the number needed to treat*. Control Clin Trials, 2001. **22**(2): p. 102-10.
30. Huber, S. and C. Rust, *Calculate travel time and distance with OpenStreetMap data using the Open Source Routing Machine (OSRM)*. Stata Journal, 2016. **16**(2): p. 416-423.
31. Bauer, J., et al., *Spatial accessibility of general inpatient care in Germany: an analysis of surgery, internal medicine and neurology*. Sci Rep, 2020. **10**(1): p. 19157.
32. Birkmeyer, J.D., et al., *Hospital volume and surgical mortality in the United States*. N Engl J Med, 2002. **346**(15): p. 1128-37.
33. Farjah, F. and F.C. Detterbeck, *What is quality, and can we define it in lung cancer?-the case for quality improvement*. Transl Lung Cancer Res, 2015. **4**(4): p. 365-72.
34. Osarogiagbon, R.U., *Volume-Based Care Regionalization: Pitfalls and Challenges*. J Clin Oncol, 2020. **38**(30): p. 3465-3467.
35. Hurley, J.C., *Variation in hospital mortality associated with surgery*. N Engl J Med, 2010. **362**(1): p. 78; author reply 79.
36. Munasinghe, A., et al., *Is It Time to Centralize High-risk Cancer Care in the United States? Comparison of Outcomes of Esophagectomy Between England and the United States*. Ann Surg, 2015. **262**(1): p. 79-85.
37. Gooiker, G.A., et al., *Quality improvement of pancreatic surgery by centralization in the western part of the Netherlands*. Ann Surg Oncol, 2011. **18**(7): p. 1821-9.
38. Aggarwal, A., et al., *Effect of patient choice and hospital competition on service configuration and technology adoption within cancer surgery: a national, population-based study*. Lancet Oncol, 2017. **18**(11): p. 1445-1453.
39. Hentschker, C. and R. Mennicken, *The volume-outcome relationship and minimum volume standards--empirical evidence for Germany*. Health Econ, 2015. **24**(6): p. 644-58.
40. Fotaki, M., *Is patient choice the future of health care systems?* Int J Health Policy Manag, 2013. **1**(2): p. 121-3.
41. Aggarwal, A., et al., *Determinants of Patient Mobility for Prostate Cancer Surgery: A Population-based Study of Choice and Competition*. Eur Urol, 2018. **73**(6): p. 822-825.
42. Bevan, G. and M. Skellern, *Does competition between hospitals improve clinical quality? A review of evidence from two eras of competition in the English NHS*. BMJ, 2011. **343**: p. d6470.
43. Fung, C.H., et al., *Systematic review: the evidence that publishing patient care performance data improves quality of care*. Ann Intern Med, 2008. **148**(2): p. 111-23.

## TABLES AND FIGURES

**Table 1.** Characteristics of the study patients, overall and by site of cancer surgery, Germany, years 2016–2018; values are counts (percentages) or mean  $\pm$  standard deviation.

Characteristic	All ( <i>n</i> = 158 079)	Site of Cancer Surgery				
		Esophagus ( <i>n</i> = 5297)	Lung ( <i>n</i> = 28 647)	Stomach ( <i>n</i> = 11 430)	Pancreas ( <i>n</i> = 10 859)	Colon/Rectum ( <i>n</i> = 101 846)
Sex						
Male	91772 (58.1)	4291 (81.0)	17 317 (60.4)	7006 (61.3)	5595 (51.5)	57 563 (56.5)
Female	66307 (41.9)	1006 (19.0)	11 330 (39.6)	4424 (38.7)	5264 (48.5)	44 283 (43.5)
Age, years	68.6 $\pm$ 11.4	64.4 $\pm$ 9.7	65.9 $\pm$ 9.7	68.5 $\pm$ 12.2	67.7 $\pm$ 10.8	69.7 $\pm$ 11.8
Charlson index						
0	45963 (29.1)	1209 (22.8)	6703 (23.4)	2472 (21.6)	1902 (17.5)	33 677 (33.1)
1	26202 (16.6)	929 (17.5)	6142 (21.4)	1750 (15.3)	1747 (16.1)	15 634 (15.4)
2	15223 (9.6)	693 (13.1)	3321 (11.6)	1227 (10.7)	885 (8.1)	9097 (8.9)
3	8497 (5.4)	369 (7.0)	1722 (6.0)	716 (6.3)	552 (5.1)	5138 (5.0)
$\geq 4$	62194 (39.3)	2097 (39.6)	10 759 (37.6)	5265 (46.1)	5773 (53.2)	38 300 (37.6)
Length of stay, days	18.2 $\pm$ 13.4	28.7 $\pm$ 23.0	15.0 $\pm$ 10.1	20.5 $\pm$ 14.1	24.1 $\pm$ 15.6	17.7 $\pm$ 12.7
In-hospital death						
Yes	6068 (3.8)	381 (7.2)	818 (2.9)	655 (5.7)	824 (7.6)	3390 (3.3)
No	152011 (96.2)	4916 (92.8)	27 829 (97.1)	10 775 (94.3)	10 035 (92.4)	98 456 (96.7)
Region of hospitalization						
Urban	79478 (50.3)	3069 (57.9)	16310 (56.9)	5513 (48.2)	6188 (57.0)	48 398 (47.5)
Suburban	47887 (30.3)	1469 (27.7)	8736 (30.5)	3523 (30.8)	2870 (26.4)	31 289 (30.7)
Rural	30714 (19.4)	759 (14.3)	3601 (12.6)	2394 (20.9)	1801 (16.6)	22 159 (21.8)

**Figure 1.** Percentage distribution of hospitals' RSMR-based mortality classes according to hospital-volume-based classes by site of cancer surgery, Germany, years 2016–2018. Classes are derived from a quintile split of hospital-specific RSMRs and average annual caseloads. Low and high classes are those in the lowest and highest quintile, respectively.



*Abbreviations:* RSMR, risk-standardized mortality rate

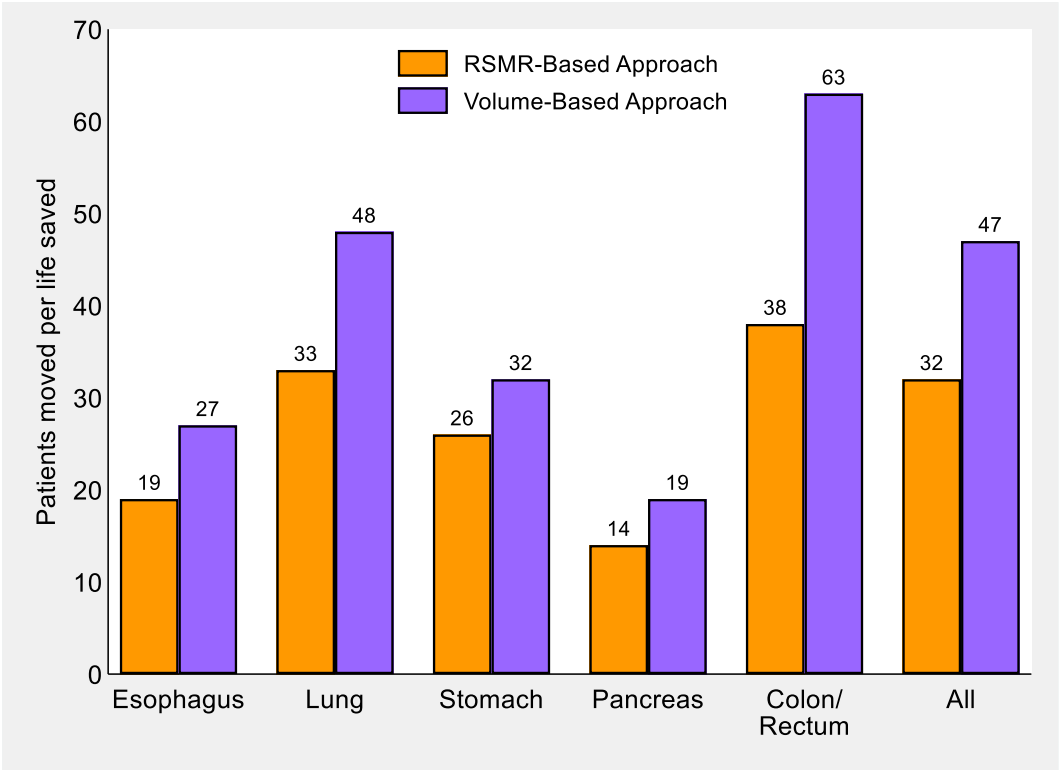


**Table 2.** Effectiveness and efficiency of RSMR-based and volume-based approaches to moving cancer patients to low-mortality (destination) hospitals within Germany, years 2016–2018, overall and by site of surgery.

	RSMR-Based Approach	Volume-Based Approach
<i>Esophagus</i>		
Lives saved (95% CI)	48 (30–68)	38 (14–61)
Patients moved	892	1018
Patients moved per life saved (95% CI)	19 (13–30)	27 (17–71)
Destination hospitals	17	8
<i>Lung</i>		
Lives saved (95% CI)	174 (138–210)	148 (105–190)
Patients moved	5796	7152
Patients moved per life saved (95% CI)	33 (28–42)	48 (38–68)
Destination hospitals	17	7
<i>Stomach</i>		
Lives saved (95% CI)	90 (57–123)	71 (42–101)
Patients moved	2313	2263
Patients moved per life saved (95% CI)	26 (19–41)	32 (23–54)
Destination hospitals	87	38
<i>Pancreas</i>		
Lives saved (95% CI)	136 (105–168)	108 (76–141)
Patients moved	1862	2069
Patients moved per life saved (95% CI)	14 (11–18)	19 (15–27)
Destination hospitals	31	12
<i>Colon and rectum</i>		
Lives saved (95% CI)	507 (438–577)	298 (229–367)
Patients moved	19 390	18 653
Patients moved per life saved (95% CI)	38 (34–44)	63 (51–81)
Destination hospitals	128	67
<i>All</i>		
Lives saved (95% CI)	955 (895–1015)	663 (606–720)
Patients moved	30 253	31 155
Patients moved per life saved (95% CI)	32 (30–35)	47 (42–55)

*Abbreviations:* RSMR, risk-standardized mortality rate; CI, confidence interval.

**Figure 2.** Efficiency of realignment strategies: for each complex-cancer surgery, the number of patients that would need to change hospitals to save a single life is shown for both the RSMR-based and the volume-based realigning approaches.

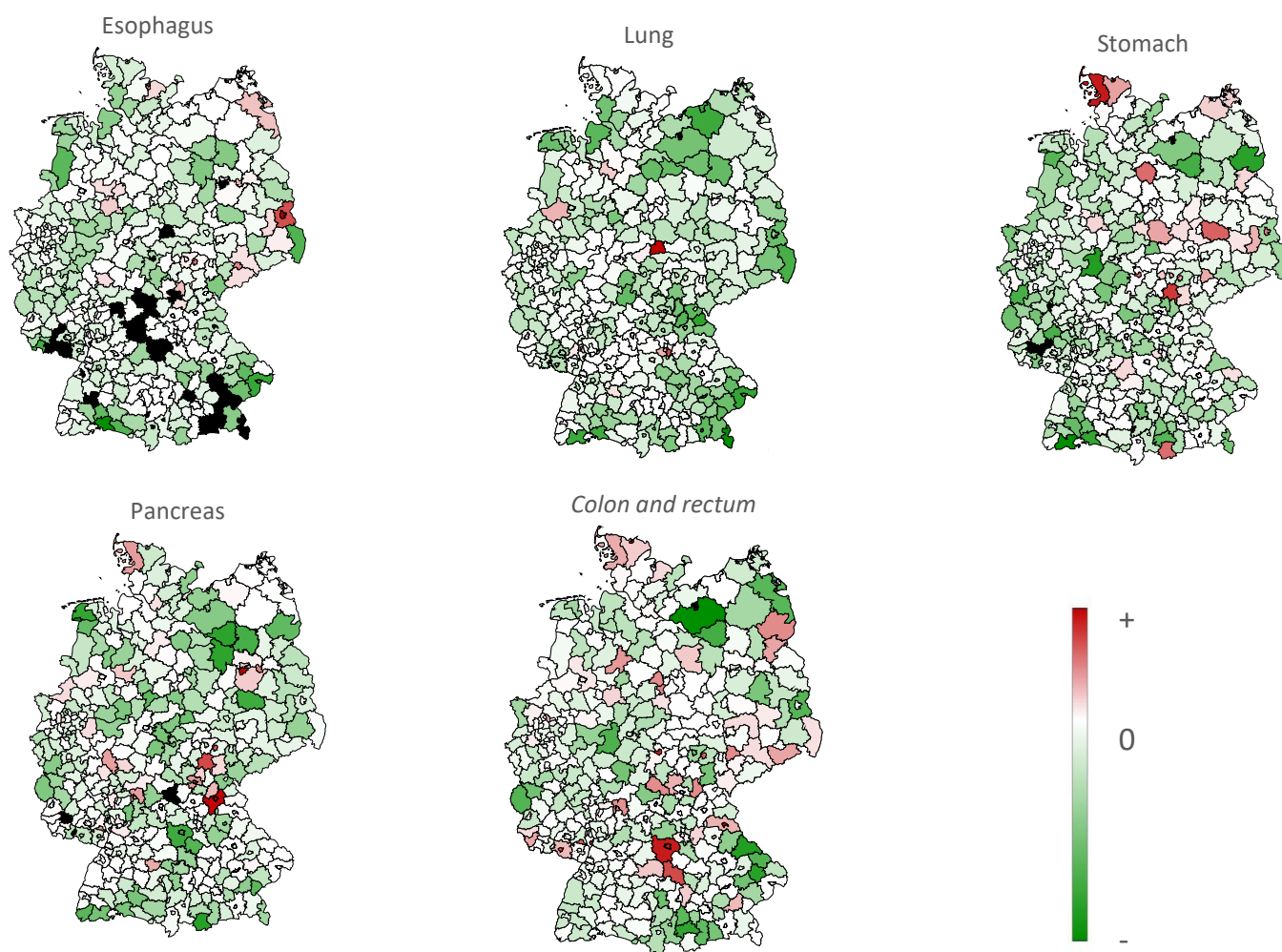


*Abbreviations:* RSMR, risk-standardized mortality rate.

**Table 3.** Changes in patient travel times and distances as a result of closing high-mortality hospitals according to RSMR-based and volume-based rankings, stratified by site of cancer surgery in Germany, years 2016–2018. Travel times and distances to the nearest hospital are the same in both RSMR and volume analysis, as they do not change according to different centralization strategies.

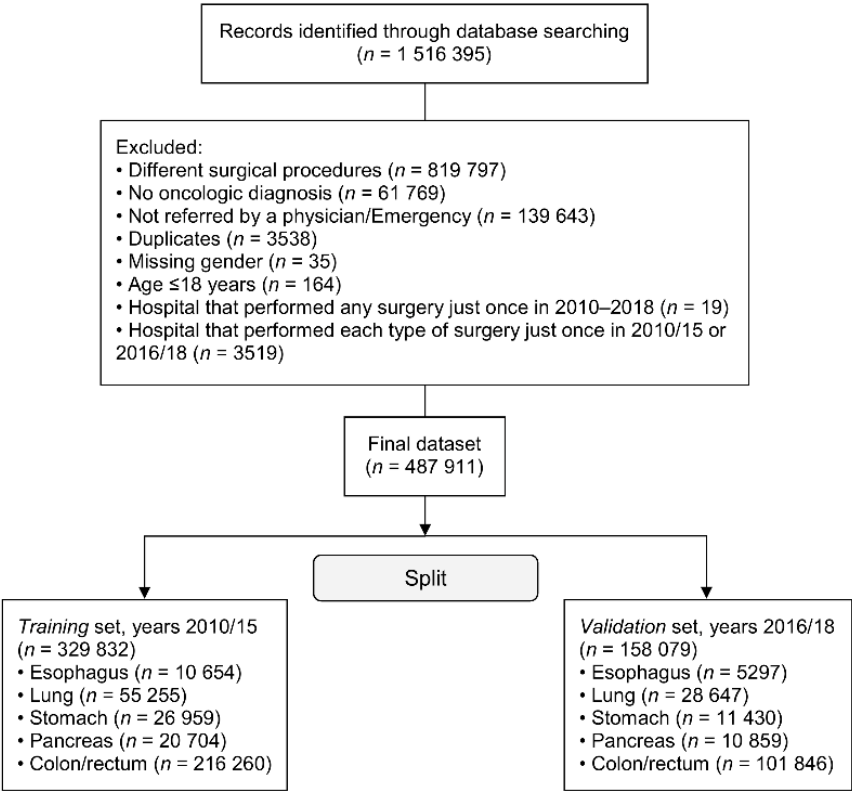
	RSMR-Based Approach		Volume-Based Approach	
	Time (min)	Distance (km)	Time (min)	Distance (km)
<i>Esophagus</i>				
To the nearest low-mortality hosp.				
Mean (95% CI)	19 (19–20)	17.9 (17.5–18.4)	27 (26–27)	27.5 (26.8–28.1)
Median [IQR]	17 [21]	13.7 [23.5]	23 [25]	21.1 [34.2]
To the nearest hospital				
Mean (95% CI)	17 (17–17)	15.1 (14.7–15.5)	17 (17–17)	15.1 (14.7–15.5)
Median [IQR]	14 [18]	11.0 [20.5]	14 [18]	11.0 [20.5]
Mean difference (95% CI)	2 (2–3)	2.8 (2.6–3.1)	10 (9–10)	12.4 (11.8–12.9)
<i>t</i> -test ( <i>P</i> -value)	23.25 (<0.001)	23.32 (<0.001)	49.16 (<0.001)	45.14 (<0.001)
<i>Lung</i>				
To the nearest low-mortality hosp.				
Mean (95% CI)	18 (18–18)	16.1 (15.9–16.3)	27 (27–27)	28.6 (28.3–28.9)
Median [IQR]	15 [20]	11.6 [20.8]	23 [27]	21.1 [36.6]
To the nearest hospital				
Mean (95% CI)	16 (16–16)	13.8 (13.7–14.0)	16 (16–16)	13.8 (13.7–14.0)
Median [IQR]	13 [17]	9.1 [19.1]	13 [17]	9.1 [19.1]
Mean difference (95% CI)	2 (2–2)	2.3 (2.2–2.4)	11 (11–11)	14.8 (14.5–15.1)
<i>t</i> -test ( <i>P</i> -value)	55.74 (<0.001)	53.67 (<0.001)	122.29 (<0.001)	112.99 (<0.001)
<i>Stomach</i>				
To the nearest low-mortality hosp.				
Mean (95% CI)	12 (12–12)	9.8 (9.6–10.0)	16 (15–16)	13.5 (13.3–13.8)
Median [IQR]	9 [13]	6.0 [12.6]	13 [16]	9.8 [18.0]
To the nearest hospital				
Mean (95% CI)	10 (10–10)	7.8 (7.7–8.0)	10 (10–10)	7.8 (7.7–8.0)
Median [IQR]	7 [11]	3.8 [10.5]	7 [11]	3.8 [10.5]
Mean difference (95% CI)	2 (2–2)	2.0 (1.9–2.1)	5 (5–6)	5.7 (5.5–5.9)
<i>t</i> -test ( <i>P</i> -value)	31.74 (<0.001)	30.88 (<0.001)	61.52 (<0.001)	58.77 (<0.001)
<i>Pancreas</i>				
To the nearest low-mortality hosp.				
Mean (95% CI)	16 (15–16)	13.7 (13.4–13.9)	21 (20–21)	19.5 (19.2–19.8)
Median [IQR]	13 [17]	10.1 [18.6]	18 [22]	14.6 [26.0]
To the nearest hospital				
Mean (95% CI)	14 (13–14)	11.3 (11.1–11.5)	14 (13–14)	11.3 (11.1–11.5)
Median [IQR]	11 [14]	7.4 [14.7]	11 [14]	7.4 [14.7]
Mean difference (95% CI)	2 (2–2)	2.4 (2.2–2.5)	7 (7–7)	8.2 (7.9–8.4)
<i>t</i> -test ( <i>P</i> -value)	36.56 (<0.001)	35.70 (<0.001)	65.86 (<0.001)	61.02 (<0.001)
<i>Colon/Rectum</i>				
To the nearest low-mortality hosp.				
Mean (95% CI)	12 (12–12)	9.3 (9.3–9.4)	14 (14–14)	11.6 (11.6–11.7)
Median [IQR]	9 [12]	5.8 [12.4]	11 [14]	8.2 [15.6]
To the nearest hospital				
Mean (95% CI)	10 (10–10)	7.2 (7.2–7.3)	10 (10–10)	7.2 (7.2–7.3)
Median [IQR]	7 [10]	3.7 [9.7]	7 [10]	3.7 [9.7]
Mean difference (95% CI)	2 (2–2)	2.1 (2.1–2.1)	4 (4–4)	4.4 (4.3–4.5)
<i>t</i> -test ( <i>P</i> -value)	115.55 (<0.001)	109.48 (<0.001)	160.20 (<0.001)	154.15 (<0.001)

**Figure 3.** Comparative effectiveness of RSMR-based vs. volume-based hospital centralization as a result of closing high-mortality hospitals, stratified by site of cancer surgery. This graphical analysis represents patients living in a county with their corresponding travel times. Travel time differences (mean duration following RSMR centralization vs. mean duration following volume centralization) in Germany, years 2016–2018. Green counties imply shorter mean travel times for a RSMR approach compared to a volume-based approach. Counties with small patient numbers are marked black to safeguard hospital identities.



SUPPORTING INFORMATION

**Figure S1.** Diagram depicting selection of the study population between 2010 and 2018.



**Table S1.** Characteristics of the study patients used in the training set of data, overall and by site of cancer surgery, Germany, years 2010–2015; values are counts (percentages) or mean  $\pm$  standard deviation.

Characteristic	All ( <i>n</i> = 329 832)	Site of Cancer Surgery				
		Esophagus ( <i>n</i> = 10 654)	Lung ( <i>n</i> = 55 255)	Stomach ( <i>n</i> = 26 959)	Pancreas ( <i>n</i> = 20 704)	Colon/Rectum ( <i>n</i> = 216 260)
Sex						
Male	193835 (58.8)	8692 (81.6)	35 661 (64.5)	15 968 (59.2)	10 773 (52.0)	122 741 (56.8)
Female	135997 (41.2)	1962 (18.4)	19 594 (35.5)	10 991 (40.8)	9931 (48.0)	93 519 (43.2)
Age, years	68.7 $\pm$ 11.5	62.7 $\pm$ 9.9	65.1 $\pm$ 9.9	69.0 $\pm$ 12.1	67.1 $\pm$ 10.5	70.0 $\pm$ 11.7
Charlson index						
0	101952 (30.9)	2788 (26.2)	12 903 (23.4)	6533 (24.2)	4314 (20.8)	75 414 (34.9)
1	55560 (16.8)	1956 (18.4)	11 914 (21.6)	4086 (15.2)	3840 (18.5)	33 764 (15.6)
2	34010 (10.3)	1424 (13.4)	6488 (11.7)	3110 (11.5)	2001 (9.7)	20 987 (9.7)
3	19057 (5.8)	864 (8.1)	3704 (6.7)	1900 (7.0)	1198 (5.8)	11 391 (5.3)
$\geq 4$	119253 (36.2)	3622 (34.0)	20 246 (36.6)	11 330 (42.0)	9351 (45.2)	74 704 (34.5)
Length of stay, days	19.9 $\pm$ 14.4	30.1 $\pm$ 24.0	17.2 $\pm$ 11.6	21.8 $\pm$ 14.7	25.8 $\pm$ 17.3	19.2 $\pm$ 13.6
In-hospital death						
Yes	14777 (4.5)	829 (7.8)	1798 (3.3)	1618 (6.0)	1677 (8.1)	8855 (4.1)
No	315055 (95.5)	9825 (92.2)	53 457 (96.7)	25 341 (94.0)	19 027 (91.9)	207 405 (95.9)
Region of hospitalization						
Urban	136424 (41.4)	5142 (48.3)	25 960 (47.0)	10 609 (39.4)	9914 (47.9)	84 799 (39.2)
Suburban	84288 (25.6)	2537 (23.8)	14 688 (26.6)	6837 (25.4)	4828 (23.3)	55 398 (25.6)
Rural	54872 (16.6)	1305 (12.2)	5960 (10.8)	5082 (18.9)	3055 (14.8)	39 470 (18.3)
Unspecified	54248 (16.4)	1670 (15.7)	8647 (15.6)	4431 (16.4)	2907 (14.0)	36 593 (16.9)

**Table S2.** Average annual caseload and in-hospital mortality in low-, mid- and high-volume hospitals by site of cancer surgery, Germany, years 2016–2018.

Site of cancer surgery	Low Volume 1 <sup>st</sup> Quintile	Medium Volume			High Volume 5 <sup>th</sup> Quintile
		2 <sup>nd</sup> Quintile	3 <sup>rd</sup> Quintile	4 <sup>th</sup> Quintile	
Esophagus					
Median [IQR]	3.0 [2.3]	6.0 [3.3]	8.3 [4.7]	17.7 [13.0]	33.0 [12.5]
Min–Max	0.7–11	1.0–30.3	4.0–32.0	8.3–47.0	4.3–107.3
<i>n</i>	119	61	35	19	8
In-hosp. mortality	8.9%	9.0%	6.3%	6.2%	5.2%
OR (95% CI)	1.00 (·)	1.01 (0.75–0.36)	0.69 (0.50–0.96)	0.67 (0.48–0.92)	0.56 (0.39–0.81)
Lung					
Median [IQR]	8.7 [14.3]	42.3 [21.0]	76.0 [30.0]	109.0 [46.7]	231.7 [121.7]
Min–Max	1.0–102.0	11.3–96.3	51.0–102.7	56.0–162.7	144.3–314.0
<i>n</i>	201	45	27	14	7
In-hosp. mortality	4.0%	2.4%	3.1%	2.3%	1.9%
OR (95% CI)	1.00 (·)	0.60 (0.49–0.74)	0.78 (0.65–0.95)	0.57 (0.46–0.72)	0.47 (0.37–0.60)
Stomach					
Median [IQR]	2.0 [1.5]	4.7 [3.0]	6.0 [3.0]	9.8 [5.3]	17.0 [9.3]
Min–Max	0.7–8.7	1.0–12.0	2.0–13.3	3.3–19.7	7.7–47.7
<i>n</i>	365	175	114	74	38
In-hosp. mortality	6.8%	6.3%	5.7%	6.1%	3.7%
OR (95% CI)	1.00 (·)	0.91 (0.72–1.15)	0.83 (0.65–1.06)	0.88 (0.70–1.12)	0.53 (0.40–0.69)
Pancreas					
Median [IQR]	3.0 [2.3]	6.3 [4.0]	10.8 [6.7]	22.3 [10.3]	43.8 [29.3]
Min–Max	0.7–9.7	1.0–14.3	4.0–25.0	5.7–37.3	22.7–209.7
<i>n</i>	223	114	62	35	12
In-hosp. mortality	9.9%	9.4%	7.3%	6.7%	4.7%
OR (95% CI)	1.00 (·)	0.95 (0.77–1.16)	0.72 (0.58–0.89)	0.65 (0.52–0.82)	0.45 (0.35–0.57)
Colon/Rectum					
Median [IQR]	13.7 [13.0]	31.7 [13.2]	47.5 [14.8]	67.5 [22.3]	100.0 [29.0]
Min–Max	0.7–51.0	10.5–130.0	18.0–93.7	21.7–101.0	55.3–221.3
<i>n</i>	429	209	146	104	67
In-hosp. mortality	4.4%	3.7%	3.0%	2.8%	2.8%
OR (95% CI)	1.00 (·)	0.85 (0.77–0.94)	0.69 (0.62–0.76)	0.64 (0.57–0.71)	0.62 (0.56–0.70)

Abbreviations: IQR, interquartile range; OR, odds ratio; CI, confidence interval.

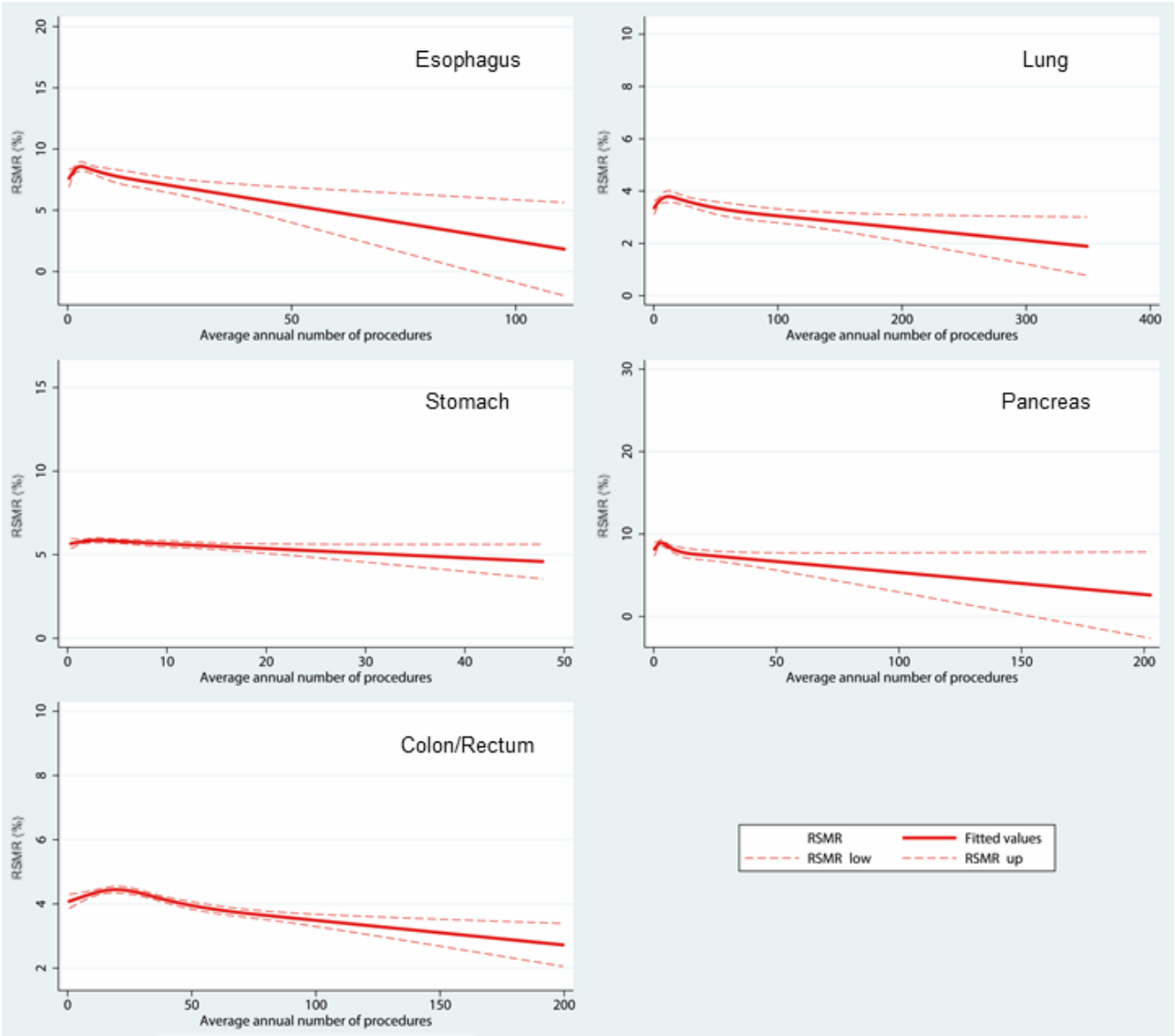
**Table S3.** Average annual caseload and in-hospital mortality in low-, mid- and high-RSMR hospitals by site of cancer surgery, Germany, years 2016–2018.

Site of Cancer	Low RSMR	Medium RSMR			High RSMR
Surgery	1 <sup>st</sup> Quintile	2 <sup>nd</sup> Quintile	3 <sup>rd</sup> Quintile	4 <sup>th</sup> Quintile	5 <sup>th</sup> Quintile
Esophagus					
Median [IQR]	13.3 [19.3]	6.3 [5.7]	3.7 [4.3]	4.7 [5.7]	4.2 [5.2]
Min–Max	2.5–107.3	2.0–47.0	1.0–45.3	0.7–33.3	1.0–33.0
<i>n</i>	17	50	62	65	48
In-hosp. mortality	3.0%	7.7%	8.1%	8.7%	8.4%
OR (95% CI)	0.34 (0.22–0.52)	0.90 (0.66–1.24)	0.95 (0.69–1.32)	1.04 (0.76–1.43)	1.00 (·)
Lung					
Median [IQR]	72.7 [53.7]	36.7 [50.0]	10.3 [25.2]	11.0 [24.7]	19.3 [33.8]
Min–Max	13.3–306.3	3.0–314.0	1.0–231.7	1.0–106.3	1.0–148.0
<i>n</i>	17	35	77	101	64
In-hosp. mortality	1.4%	2.3%	2.4%	3.6%	4.4%
OR (95% CI)	0.31 (0.24–0.40)	0.50 (0.41–0.63)	0.54 (0.44–0.66)	0.81 (0.67–0.97)	1.00 (·)
Stomach					
Median [IQR]	7 [6.3]	3.3 [3.7]	3.0 [3.5]	3.0 [3.5]	3.7 [4.5]
Min–Max	1.7–47.7	0.7–30.3	0.7–46.0	0.7–26.7	0.7–30.3
<i>n</i>	87	186	176	169	148
In-hosp. mortality	4.6%	4.3%	5.3%	6.1%	8.5%
OR (95% CI)	0.52 (0.41–0.67)	0.48 (0.37–0.61)	0.60 (0.47–0.76)	0.69 (0.55–0.87)	1.00 (·)
Pancreas					
Median [IQR]	16.0 [22.7]	6.3 [7.7]	4.0 [4.3]	4.3 [5.7]	5.0 [5.3]
Min–Max	3.0–209.7	1.3–66.7	0.7–31.3	0.7–88.0	0.7–41.7
<i>n</i>	31	83	126	106	100
In-hosp. mortality	4.4%	7.0%	7.5%	8.4%	11.7%
OR (95% CI)	0.35 (0.27–0.44)	0.57 (0.46–0.71)	0.61 (0.49–0.76)	0.69 (0.56–0.85)	1.00 (·)
Colon/Rectum					
Median [IQR]	49.8 [50.3]	36.2 [42.0]	25.7 [34.0]	23.2 [26.5]	24.2 [22.0]
Min–Max	1.7–221.3	1.0–185.3	0.7–199.7	1.0–114.0	2.0–107.7
<i>n</i>	128	170	211	228	218
In-hosp. mortality	2.2%	2.9%	3.1%	3.9%	4.8%
OR (95% CI)	0.44 (0.39–0.49)	0.59 (0.54–0.66)	0.63 (0.57–0.70)	0.80 (0.72–0.88)	1.00 (·)

Abbreviations: RSMR, risk-standardized mortality rate.



**Figure S2.** Restricted cubic spline analysis modelling the correlation between individual hospital volumes and RSMRs by site of cancer surgery, Germany, years 2016–2018. Axis ranges differ across charts; point clouds are concealed to safeguard hospitals identities.



**Table S4.** Travel times and distances to the nearest hospital vs. the hospital where the patient was surgically treated by site of cancer surgery, Germany, years 2016–2018.

	To the Hospital That Performed Surgery		To the Nearest Hospital		Mean Difference (95% CI)	
	Time (min)	Distance (km)	Time (min)	Distance (km)	Time (min)	Distance (km)
<i>Esophagus</i>						
Mean (95% CI)	39 (38–40)	47.2 (45.4–49.1)	17 (17–17)	15.1 (14.7–15.5)	22**	32.2**
Median [IQR]	27 [36]	25.5 [49.0]	14 [18]	11.0 [20.5]	(21–23)	(30.3–34.0)
<i>Lung</i>						
Mean (95% CI)	32 (31–32)	34.3 (33.7–34.8)	16 (16–16)	13.8 (13.7–14.0)	16**	20.5**
Median [IQR]	26 [26]	22.2 [34.1]	13 [17]	9.1 [19.1]	(15–16)	(20.0–21.0)
<i>Stomach</i>						
Mean (95% CI)	24 (24–25)	26.0 (25.1–26.9)	10 (10–10)	7.8 (7.7–8.0)	14**	18.2**
Median [IQR]	16 [21]	12.0 [23.0]	7 [11]	3.8 [10.5]	(13–15)	(17.3–19.1)
<i>Pancreas</i>						
Mean (95% CI)	38 (37–39)	46.1 (44.6–47.6)	14 (13–14)	11.3 (11.1–11.5)	24**	34.8**
Median [IQR]	23 [34]	19.9 [44.2]	11 [14]	7.4 [14.7]	(23–25)	(33.3–36.2)
<i>Colon/Rectum</i>						
Mean (95% CI)	20 (20–20)	19.5 (19.2–19.7)	10 (10–10)	7.2 (7.2–7.3)	10**	12.2**
Median [IQR]	14 [17]	10.4 [17.6]	7 [10]	3.7 [9.7]	(10–10)	(12.0–12.5)

\*\*  $P$ -value  $\leq 0.01$ ; \*  $P$ -value  $\leq 0.05$ .

**Table S5.** Sensitivity quartile-based analysis of effectiveness and efficiency of RSMR-based and volume-based approaches to move cancer patients to low-mortality (destination) hospitals within Germany, years 2016–2018, overall and by site of surgery. Cut-points of RSMRs and volumes were defined using quartiles in place of quintiles; no changes were made to the inclusion and exclusion criteria of the study patients.

	RSMR-Based Approach	Volume-Based Approach
<i>Esophagus</i>		
Lives saved (95% CI)	54 (32–77)	47 (21–73)
Patients moved	1222	1364
Patients moved per life saved (95% CI)	23 (16–39)	29 (19–64)
Destination hospitals	29	12
<i>Lung</i>		
Lives saved (95% CI)	188 (147–230)	162 (117–206)
Patients moved	7163	8653
Patients moved per life saved (95% CI)	38 (31–49)	53 (42–74)
Destination hospitals	26	10
<i>Stomach</i>		
Lives saved (95% CI)	102 (66–138)	77 (44–110)
Patients moved	2850	2845
Patients moved per life saved (95% CI)	28 (21–43)	37 (26–64)
Destination hospitals	125	53
<i>Pancreas</i>		
Lives saved (95% CI)	145 (110–180)	132 (96–169)
Patients moved	2414	2568
Patients moved per life saved (95% CI)	17 (13–22)	19 (15–27)
Destination hospitals	44	18
<i>Colon and rectum</i>		
Lives saved (95% CI)	548 (471–626)	407 (332–484)
Patients moved	24340	23444
Patients moved per life saved (95% CI)	44 (39–52)	58 (48–71)
Destination hospitals	164	91
<i>All</i>		
Lives saved (95% CI)	1037 (975–1099)	825 (766–884)
Patients moved	37 989	38 874
Patients moved per life saved (95% CI)	37 (35–40)	47 (43–53)