Preprint Version

Using exogenous organizational and regional hospital attributes to explain differences in case-mix adjusted hospital costs

Michael M. Havranek[1]*, Josef Ondrej[1], Philippe K. Widmer[2], Stella Bollmann[1], Simon Spika[3], and Stefan Boes[1]

[1] Department of Health Sciences and Medicine, University of Lucerne, Lucerne, Switzerland, [2] Spital Limmattal Zurich, Zurich, Switzerland, [3] University Hospital Zurich, Zurich, Switzerland

* corresponding author: michael.havranek@unilu.ch

Abstract

Diagnosis-related group (DRG) hospital reimbursement systems differentiate cases into cost-homogenous groups based on patient characteristics. However, exogenous organizational and regional factors can influence hospital costs beyond case-mix differences. Therefore, most countries using DRG systems had to incorporate adjustments for such factors into their reimbursement structure. This study investigates structural hospital attributes that explain differences in average case-mix adjusted hospital costs. Using rich patient and hospital-level data containing 4 million cases from 120 hospitals across three years, we show that a regression model using only five variables (number of discharges, ratio of emergency/ ambulance admissions, rate of DRGs to patients, expected loss potential based on DRG mix, and location in large agglomeration) can explain more than half of the variance in average case-mix adjusted hospital costs, capture all cost variations across commonly differentiated hospital types (e.g., academic teaching hospitals, children's hospitals, birth centers, etc.), and is robust in cross-validations across several years (despite differing hospital samples). Based on these findings, we propose an approach to differentiate legitimate from inefficiency-related cost differences across hospitals and discuss the advantages of such an approach as a fair and transparent way to incorporate structural hospital differences into cost benchmarking and the design of payment schemes.

Keywords: Hospital costs, organizational and regional factors, DRG, operating efficiency, cost benchmarking, hospital reimbursement.

1. INTRODUCTION

Diagnosis-related group (DRG) systems are commonly used in hospital reimbursement in developed countries. Implementations of DRG systems vary between countries, but central to all of them is a patient classification that aims to differentiate inpatient cases into cost-homogenous groups. This classification is based on standardized case attributes, such as patients' age, sex, diagnoses, and procedures. With respect to their goal of increasing efficiency, DRG systems critically rely on the assumption that after adjustment for all relevant case attributes, expected treatment costs are identical across efficiently operating hospitals [1]. However, case attributes alone are likely insufficient to model the efficiency frontier in a heterogeneous market like the hospital sector, and consequently, most countries' reimbursement schemes allow for inclusions of additional organizational or regional factors to differentiate hospital costs beyond patient characteristics.

According to a comparison of DRG systems across Europe [1], Estonia is the only country that treats hospitals similarly nationwide. In contrast, Germany and Austria, for instance, use geographic regions to differentiate hospitals, while in Ireland and Portugal, certain hospital peer groups such as academic teaching hospitals or children's hospitals are reimbursed differently from other hospital types. Still other countries, like England and France, use adjustments to compensate for specific structural differences between hospitals (such as salary levels). Despite these mechanisms, recent studies have shown that hospital payments often do not only reflect legitimate cost variations of providers but are also influenced by cost-unrelated factors [2, 3]. If structural differences between hospitals are exogenous (i.e., uncontrollable for hospitals), it is justified to include them in the reimbursement scheme to generate a level playing field. However, if they cannot be considered exogenous, then they should not be included in the reimbursement scheme to prevent gaming of the system and unwarranted payments. This is presumably one reason why some authors have demanded fairer, more transparent, and more objective data-driven approaches to differentiate hospital costs [4].

In Switzerland, which is the focus of our study, DRG-based hospital payments are determined by relative cost weights, informing about the resource requirements of specific patient groups in relation to the average treatment costs of all cases in the country. These relative weights are converted into actual payments by multiplying them with a base rate. By law, this base rate can differ between hospitals and is to be negotiated with health insurers using a national benchmark of what is considered to be an "efficient" hospital. The definition of this benchmark is based on an efficiency threshold (i.e., a specific percentile in the distribution) of the average case-mix adjusted costs (defined as the sum of total costs divided by the sum of cost weights of all patients in a hospital) of all Swiss hospitals. However, this approach has caused debates and lawsuits in the past as to whether payments are fair across different hospitals, whether the impact of organizational and regional factors on hospital costs is sufficiently accounted for, and if not, how additional adjustments can be made in practice (see, e.g., [5, 6]).

The impact of organizational and regional influences on hospital costs is of particular relevance in Switzerland. This is because the Swiss DRG system, in comparison with DRG systems of other countries, is an almost fully prospective, DRG-reimbursed financing system with only few special payments [1]. Despite this practical importance, a consensus on how to incorporate structural hospital differences into cost benchmarking and reimbursement is still lacking. However, this almost fully prospective, DRG-reimbursed nature of hospital financing in Switzerland makes the Swiss DRG system interesting from a scientific perspective of an international readership because relationships between structural factors and hospital costs can be observed more directly than may be possible in other settings.

Previously, two approaches have been suggested on how to use organizational attributes to distinguish hospitals [6, 7]. However, one of these approaches was only able to meaningfully differentiate university (i.e., academic teaching) and children's hospitals from other hospital types, while the remaining hospitals could not be distinguished beyond 1-2% of cost differences [7]. The second approach, on the other hand, proposed a more elaborate (and complex) multi-level hierarchical model using approximately twenty organizational attributes, which restricted its comprehensibility and practical applicability [6]. These limitations of previous approaches motivated us to find an effective but parsimonious model to distinguish hospital costs beyond common case-mix adjustments.

More specifically, the objective of our study was to develop an econometric model for case-mix adjusted hospital costs in Switzerland that uses a small set of variables with good predictive properties. To accomplish this, we used rich patient and hospital-level data to fit a regression model with the best predicting organizational and regional aspects associated with cost variations. Hereby, we specifically focused on factors that are exogenous (i.e., largely uncontrollable by hospitals). Subsequently, we will present how model development and evaluation were performed. Then we will show that our developed model performs well on the given data in terms of goodness-of-fit and that it can fully explain cost variations between various commonly distinguished hospital types across multiple years and samples. Finally, we conclude by discussing the practical applicability of such a model to distinguish legitimate cost differences from operating (in-)efficiencies of hospitals and show how to calculate hospital-specific benchmarks, which can be used in cost benchmarking and the design of hospital reimbursement.

2. METHODS

2.1. Data source

Our analyses build on Swiss inpatient data collected between 2017 and 2019 that was reimbursed as part of the Swiss DRG system.[1] The dataset links routine administrative patient records reported to the Federal Statistical Office (FSO)[2] with organizational data on hospitals[3] as well as specific cost data on the patient[4] level and aggregated cost data on the hospital[5] level, which we enriched by publicly available geographic information from the FSO[6] to capture different regional and spatial characteristics. Data of 2019 was used for model development, while data of 2018 and 2017 was only used for cross-validation (see below) due to smaller hospital sample sizes in 2018 and 2017. In total the data contained 1,360,975 cases from 120 hospitals in 2019, 1,285,251 cases from 101 hospitals in 2018, and 1,305,298 cases from 93 hospitals in 2017.

Focusing on our modeling data from 2019, the patient-level sample covers about 93% of all stationary cases in Switzerland [8]. After excluding cases spanning more than one reporting year (i.e., not concluded cases, *n* = 12,755), cases with missing information in the main variables (DRG codes, cost weights, or costs, *n* = 12,795), psychiatric and rehabilitation cases that are reimbursed outside of the DRG system (*n* = 142,081), and cases from four highly specialized specialty clinics (*n* = 4,452, as has been partly proposed previously [7]), our final patient-level sample contained 1,188,892 cases from 116 hospitals in 2019. This patient-level dataset was used to build the hospital attributes that were examined for inclusion in the model (see Table 1), but all subsequently reported analyses were executed at an aggregated hospital level. At hospital level, the modeling data from 2019 contained 5 university hospitals (i.e., academic teaching hospitals with an affiliation to a medical school), 38 large central general hospitals, 41 medium-sized and small regional general hospitals, 10 birth centers, 3 children's hospitals, and 19 specialty hospitals (including surgical and other specialty clinics).

---

[1] The dataset was provided by the "Verein SpitalBenchmark", which is a hospital association that offers a platform for its member hospitals "on which sensitive and confidential data on internal processes, coding, total costs, and case costs in all sub-areas of hospitals and clinics are objectively processed and exchanged" (see https://www.spitalbenchmark.ch/ueber-uns/ziel-und-zweck/).
[2] "Medizinische Statistik der Krankenhäuser", for further information (only available in German/ French/ Italian) see https://www.bfs.admin.ch/bfs/de/home/statistiken/gesundheit/erhebungen/ms.html.
[3] "Krankenhausstatistik", for further information (only available in German/ French/ Italian) see https://www.bfs.admin.ch/bfs/de/home/statistiken/gesundheit/erhebungen/ks.html.
[4] "Statistik diagnosebezogener Fallkosten", for further information (only available in German/ French/ Italian) see https://www.bfs.admin.ch/bfs/de/home/statistiken/gesundheit/erhebungen/fks.html.
[5] "Kosten- und Leistungsdaten für ITAR_K", see for example https://www.hplus.ch/de/rechnungswesen/itar-k (information available only in German/ French/ Italian).
[6] "Gemeindetypologie", for further information (only available in German/ French/ Italian) see https://www.bfs.admin.ch/bfs/de/home/statistiken/kataloge-datenbanken/karten.assetdetail.2543279.html.

2.2. Model development

Our goal was to build a simple and interpretable model with a limited number of explanatory variables, allowing us to differentiate DRG-reimbursed costs between hospitals and hospital types beyond case-mix adjustments. For this reason, we used average case-mix adjusted hospital costs (calculated as sum of total costs divided by sum of cost weights of all patients in a hospital) as our dependent variable and excluded costs for services reimbursed by private health insurances or by supplementary payments (for certain high-cost medications and treatments)[7].

Based on a literature review and preliminary analyses, we considered 19 potential explanatory variables across four attribute groups for inclusion in our model (see Table 1). These variables have previously been found to be associated with case-mix adjusted hospital cost variations in Switzerland [9].

Table 1 *Description of full list of potential explanatory variables*

Size and complexity[a]

Variable	Operationalization/ Description
Number of beds	as indicator for hospital size
Number of discharges (per year)[b]	as indicator for patient volume
Number of types of services	consisting of acute care, psychiatry, and rehabilitation
Number of departments	e.g., internal medicine, surgery, obstetrics and gynecology, etc., according to the classification of the FSO

Essential healthcare functions[c]

Variable	Operationalization/ Description
Number of trained personnel groups	consisting of residents, medical students, and other health professions
Ratio[d] of deliveries[b]	only counting deliveries without complications or complicating procedures
Ratio[d] of emergency/ ambulance admissions[b]	excluding deliveries without complications or complicating procedures
Ratio[d] of admissions during weekend/ night	excluding deliveries without complications or complicating procedures
Ratio[d] of admissions from nursing homes	including admissions from elderly homes, nursing homes, and other sociomedical institutions
Rate[d] of DRGs to patients[b]	= number of distinct DRGs treated at hospitals / number of discharges
Rate[d] of specialized services to patients	= number of distinct services[e] requiring equipment or expertise not present in all hospitals / number of patients

[7] Q-Q plot and Shapiro-Wilk test (p = 0.97) confirmed that the dependent variable is normally distributed.

Rarity and financial risk of patient mix[f]

Variable	Operationalization/ Description
Case-mix index	= sum of cost weights / number of discharges
Ratio[d] of rare DRGs[b]	using 200 cases across all hospitals per year as thresholds for rare DRGs
Ratio[d] of children[b]	including ages 0-17, but excluding healthy newborns delivered without complications or complicating procedures
Expected loss potential based on DRG mix[b,g]	as indicator for financial risk taken on by hospitals based on their patient mix
Ratio[d] of admissions from outside of canton	as indicator for hospital specialization

Regional attributes[h]

Variable	Operationalization/ Description
Location in large agglomeration[b,i]	hospitals located in municipalities within a large agglomeration
Location in medium-sized/ small agglomeration[i]	hospitals located in municipalities within a medium-sized or small agglomeration
Location in peri-urban/ rural area[i]	hospitals located in municipalities within a peri-urban or rural area

Note. [a] These references motivated us to include the respective variables as potential explanatory variables: [6, 7, 10, 11]. [b] Variables printed in **bold** were preselected for subsequent inclusion in variable selection (see main text below) among the various associated variables within each attribute group. [c] These references inspired us to include the respective variables: [5-7, 12-15]. [d] The ratio/ rate variables are expressed in relation to all discharges/ patients to include their association with costs independent of hospital size (as an inherently present confounding factor) in the modelling. [e] Specialized services include services requiring an emergency room, intensive care unit, computed tomography or magnetic resonance imaging, dialysis, lithotripsy, and radiotherapy equipment. [f] These references motivated us to include the respective variables: [5-7, 16-18]. [g] This variable was generated by first calculating the *expected loss* for each DRG across all cases in the country as the average of actual costs minus reimbursed costs in cost outliers. The *expected loss potential* of a given hospital was then calculated as the average of expected losses across all cases of all DRGs in the hospital. This approach (see [9] for more details) quantifies expected losses from cost outliers in certain (mainly small sample) DRGs that have been shown not to be sufficiently compensated for in Switzerland [9], without including (in-)efficiencies of individual hospitals. [h] These references inspired us to include the respective variables: [10, 15, 19] [i] Hospitals with multiple locations are expressed as a percentage of locations within each category in these variables. Furthermore, the variables also incorporate the percentage of the vicinity around the location of hospitals that belong to the respective category. For [g] and [i], different versions of variable thresholds were tested for best model fit, varying the definition of cost outliers (2, 3, or 5 standard deviations from the mean) and the radius around the location to define vicinity (0, 10, 25, and 50 km). Based on this, the 5 standard deviation threshold was chosen for the definition of cost outliers, whereas the 25 km radius was selected for the vicinity around the hospital location.

As a first step in model development, we excluded variables (the ones that are not printed in bold in Table 1), which were judged as less suitable compared to others based on several criteria. This preselection step was based on expert consensus from focus group meetings and stakeholder consultations. Since we specifically focused on systematic and exogenous (i.e., for hospitals uncontrollable) influences on costs, the most important preselection criteria was that variables are largely uncontrollable for hospitals. For this, we defined “largely uncontrollable variables” as exogenous aspects that cannot be changed by hospitals in the short-term without taking drastic measures (such as, for example, completely relocating their facilities). Our definition of “uncontrollable” also takes into account what aspects cannot be changed by hospitals based on the service mandate that Swiss hospitals are obligated to provide in their service area. For example, while certain variables such as the number of treated DRGs could theoretically be defined by hospitals in order to maximize profits. In practice, this is not possible for most hospitals in Switzerland, because of the service mandate they receive from cantonal authorities.

The number of types of services, number of trained personnel groups, and rate of specialized services to patients were excluded because they were criticized by experts as being relatively easy to manipulate. For instance, because of the operationalization of the variable number of trained personnel groups, hospitals could in theory increase the value of this variable by just employing one single resident (if they did not previously train residents). The number of departments and ratio of admissions from nursing homes were excluded, because the definitions of these variables were compromised by slight regional differences, which led to concerns about their acceptance. For example, the number of departments in hospitals can be counted in various ways depending on how medical specialties are divided into departments. The ratio of admissions from outside of canton was excluded because it was assessed as being potentially influenced by unintended aspects such as hospitals’ location nearby a cantonal border. Finally, location in medium-sized/ small agglomeration and location in peri-urban/ rural area were excluded because the 25 km radius version of the location in large agglomeration variable best captured regional cost differences among hospitals (see also notes of Table 1).

This preselection of variables also served to reduce collinearities that were found between many variables within the four attribute groups. For instance, number of beds correlated highly with number of discharges (r = 0.95) and was excluded in favor of number of discharges, because number of beds was judged to be controllable by hospitals. Likewise, the ratio of admissions during weekend/ night was highly associated with the ratio of emergency/ ambulance admissions (r = 0.96) and was excluded, since ratio of emergency/ ambulance admissions was assessed as being the variable that more directly designates the presumed reason for the association with hospital costs.

After this theoretically driven preselection, nine potential explanatory variables (printed in bold in Table 1) remained. To determine an optimal combination of a limited number of variables among these preselected ones, we used two specifically developed variable selection algorithms. Our first algorithm performed forward stepwise selection on the entire sample of hospitals divided into five subsamples based on 3-fold cross-validation and used R^2 as variable selection criteria. It was stopped after the selection of five variables when additionally chosen variables no longer increased adjusted R^2 and adjusted R^2 started to decrease (from 0.4941 to 0.4940). Using this procedure, the following variables were selected (in this order): 1. expected loss potential based on DRG mix, 2. location in large agglomeration, 3. ratio of emergency/ ambulance admissions, 4. rate of DRGs to patients, and 5. number of discharges.

To confirm the robustness of our variable selection, our second algorithm was designed to forward select a defined number of variables (n = 5, based on the model size selected in step 1 above) for 100 randomly sampled subsamples (n = full sample / 2, with replacement) and test on another 100 random subsamples, the combination of variables showed the lowest mean absolute error (*MAE*) across the test samples. This approach confirmed that the five most frequently included variables in the 100 best models were: location in large agglomeration (selected in 93/100), rate of DRGs to patients (76/100), expected loss potential based on DRG mix (69/100), ratio of emergency/ ambulance admissions (62/100), and number of discharges (56/100).

2.3. Model evaluation

Model validation was performed with regard to the following aspects: goodness of fit was assessed by the proportion of explained variance in hospital costs (R^2) with different weighting approaches (see below), and by putting it into perspective in comparison with a "full" model including all preselected cost-associated independent variables presented above. Collinearity of explanatory variables was assessed via variance inflation factors (*VIF*, i.e., the quotients of the variance in the model with all selected variables together divided by the variance of the model with each variable separately). Model prediction residuals were compared across different hospital types to rule out preferential treatment of certain types of hospitals and examined in relation to patient volume of hospitals (as an indicator of the precision in the point estimates of hospital costs, see results below). Moreover, two "extended" models additionally including dummy variables and interactions for hospital types (university hospitals, large central general hospitals, medium-sized and small regional general hospitals, birth centers, children's hospitals, specialty hospitals) were used to assess whether any differences that were explainable by hospital types remained. Furthermore, all reported models (incl. the full and extended models) are compared with regard to additional measures of model fit that take the number of explanatory variables into account (*adjusted* R^2, *Akaike's Information Criteria* (*AIC*), *Bayesian Information Criteria* (*BIC*)). Generalization performance of the model results within the population of Swiss hospitals was evaluated using cross-validation across several years. Finally, the practical relevance of our results was demonstrated by comparing the distribution of model predictions with the distribution of the observed case-mix adjusted costs across hospitals.

2.4. Statistical analyses

The following weighted least squares (WLS) regression was used to explain case-mix adjusted hospital costs, denoted by y_i, using a set of explanatory variables, denoted by $x_{k,I}$:

$$y_i = \beta_0 + \beta_1 x_{1,i} + \beta_2 x_{2,i} + \beta_3 x_{3,i} + \beta_4 x_{4,i} + \beta_5 x_{5,i} + \varepsilon_i \quad (1)$$

for $i = 1, ..., N_{\mathrm{hosp}}$, where

- $x_{1,i} \ldots x_{5,i}$ are the selected explanatory variables, and
- ε_i is the error term for hospital i, with heteroscedastic variance $\sigma^2_{HC,i}$.

Note that the number of explanatory variables in equation (1) was not pre-determined but turned out to be the best fitting model in the analysis. Similar WLS regressions were used for the “full” model (including all preselected explanatory variables) and the two “extended” models (including dummy variables and interactions for hospital type). Regression weights were chosen to capture an observed relationship between the standard deviations of hospital costs across years and patient volume of hospitals[8] in order to quantify the precision in the estimation of hospital costs (for a discussion of this topic, see, e.g., [20]). In addition, results will also be reported for patient volume-weighted and unweighted regressions. Cross-validation across years was performed by partitioning the data into three complementary subsets that each consisted of two years as a training set (e.g., 2017 and 2018), on which the model was estimated, and one year as testing set (e.g., 2019), on which model predictions were evaluated. All analyses were conducted using Python (Version 3.8.8) and results were considered significant if $p < 0.05$.

[8] As weights w_i we use $w_i = 1 / \sigma^2_{HC,I}$, where $\sigma_{HC,i}$ is quantified as $\sigma_{HC,i} = a / volume_i^b$, and a and b are parameters estimated by maximum likelihood based on volume and standard deviation of hospital costs across the years 2017-2019, which is a function that fitted variations in hospital costs across years very well in our data.

3. RESULTS

Table 2 presents descriptive statistics for our dependent variable and the set of explanatory variables to summarize the main features of the data and to demonstrate that considerable heterogeneity exists between and within different hospital types. As can be observed, case-mix adjusted hospital costs vary widely between CHF 8,133.00 and CHF 12,905.00, which means that deviations from mean costs range from CHF -2,301.00 (CHF 8,133.00 - 10,434.00, -22%) to CHF 2,471.00 (CHF 12,905.00 - 10,434.00, 24%).

Table 2 *Descriptive statistic (means, standard deviations (SD), and ranges) of selected variables*

	Full Sample	University	Central	Regional	Birth	Children	Specialty
n	116	5	38	41	10	3	19
Hospital costs							
Mean (SD)	10,433.73 (866.81)	11,748.17 (676.83)	10,393.26 (540.63)	10,518.03 (772.77)	9,524.71 (892.91)	11,801.24 (573.43)	10,249.33 (1,014.56)
Range	8,133.33-12,904.84	11,144.9-12,904.84	9,091.9-11,397.2	8,970.81-12,478.39	8,133.33-11,163.95	11,249.41-12,394.06	8,970.81-11,976.14
Number of discharges							
Mean (SD)	10,400.19 (11,920.13)	46,544.4 (6,921.79)	18,256.05 (9,793.9)	4,867.88 (2,691.18)	376.5 (286.61)	6,275.33 (2,205.67)	3,041.89 (2,145.49)
Range	110.0-56,157.0	37,503.0-56,157.0	3,177.0-42,706.0	406.0-9,629.0	110.0-892.0	4,162.0-8563.0	481.0-7,798.0
Ratio of emergencies							
Mean (SD)	0.39 (0.25)	0.45 (0.05)	0.5 (0.15)	0.5 (0.21)	0.01 (0.01)	0.61 (0.05)	0.06 (0.08)
Range	0.0-0.89	0.39-0.51	0.0-0.68	0.01-0.89	0.0-0.02	0.55-0.65	0.0-0.26
DRGs to patients							
Mean (SD)	0.06 (0.05)	0.02 (0.0)	0.04 (0.01)	0.1 (0.06)	0.03 (0.02)	0.08 (0.02)	0.05 (0.03)
Range DRGs	0.01-0.3	0.02-0.02	0.02-0.06	0.03-0.3	0.01-0.05	0.06-0.09	0.02-0.11
Expected loss potential							
Mean (SD)	93.16 (90.26)	342.06 (80.88)	103.06 (31.37)	66.92 (34.57)	4.63 (1.39)	400.42 (30.25)	62.6 (75.82)
Range	1.99-462.52	252.54-462.52	35.65-193.36	17.57-199.16	1.99-6.59	380.54-435.22	7.53-329.52
Location in large agglomeration							
Mean (SD)	0.11 (0.12)	0.24 (0.15)	0.12 (0.12)	0.05 (0.09)	0.1 (0.11)	0.19 (0.16)	0.16 (0.15)
Range	0.0-0.47	0.1-0.45	0.0-0.32	0.0-0.31	0.0-0.3	0.0-0.31	0.0-0.47

Note. n = sample size, University = university hospitals, Central = central general hospitals, Regional = regional general hospitals, Birth = birth centers, Children = children's hospitals, Specialty = specialty hospitals

3.1. Presentation of selected variables and model results

As outlined above, our main regression model for case-mix adjusted hospital costs consisted of the following five structural hospital attributes:

1.) number of discharges to describe the size or more precisely the complexity within hospitals;
2.) ratio of emergency/ ambulance admissions, which stands for the extent of emergency and first point of care functions that hospitals perform in their service area;
3.) rate of DRGs to patients, which illustrates the broadness of the treatment spectrum that hospitals provide (in relation to their patient volume);
4.) expected loss potential based on DRG mix, which characterizes the financial risk taken on by hospitals based on their patient mix; and
5.) location in large agglomerations, which is the regional factor that best captured cost variations across geographic areas among the hospitals

(see also the discussion for a more detailed reflection on the selected variables).

Table 3 presents the WLS results of the main model, including estimated coefficients, standard errors, t-statistics, and p-values for the test of zero coefficients. Since the dependent variable (average case-mix adjusted hospital costs) is measured in CHF, the estimated *β* coefficients can be interpreted in CHF as well. Expected hospital costs for a hypothetical hospital with all selected variables equal to zero are predicted at CHF 8'823.95 (the intercept in Table 3). Non-zero values in the explanatory variables of a specific hospital can be multiplied with the *β* coefficient of that variable (e.g., 0.5 x CHF 667.68 for a hospital that treats 50% emergency patients) and added up across all variables (including the intercept) to obtain predictions of the expected case-mix adjusted costs of that hospital. In that way, each hospital can calculate its expected costs based on the estimated model coefficients.

It should be noted that *β* coefficients are not standardized, i.e., the magnitude of the coefficients does not express their relative importance but rather is related to the values within that variable observed in the sample (for standardized model results, see Table S1 in the Supporting Information). Therefore, for example, the coefficient of the rate of DRGs to patients is relatively high because the values of this variable are relatively small since most hospitals treat many patients within a DRG. Similarly, the coefficient of location in large agglomeration is relatively high because only a few hospitals have comparably large values in that variable. This is the case because only a few hospitals have all their locations directly situated in the center of a large agglomeration so that the entire vicinity radius (25km according to our specifications) around their locations would be situated within that large agglomeration.

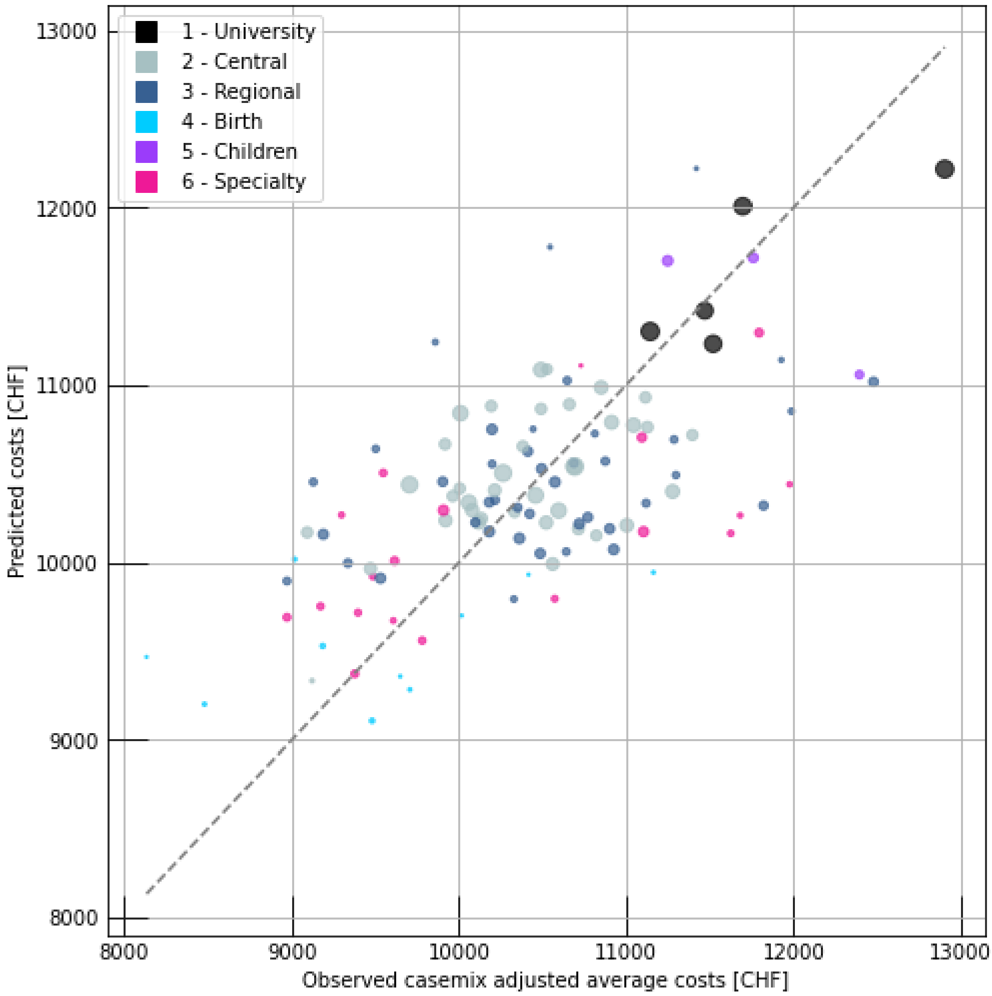
Table 3 *WLS regression results of the model using the selected variables to explain hospital costs (in CHF)*

	β	*Std Err*	*t*	*p*	
Intercept	8,968.50	213.02	42.10	0.000	***
Number of discharges	0.02	0.01	2.83	0.006	**
Ratio of emergency/ ambulance admissions	554.29	308.97	1.79	0.076	‘
Rate of DRGs to patients	8,933.36	2,500.39	3.57	0.001	**
Expected loss potential based on DRG mix	2.18	0.71	3.08	0.003	**
Location in large agglomeration	2,868.60	450.60	6.37	0.000	***

Note. Dependent variable = average case-mix adjusted hospitals costs, *n* = 116, *df(residuals)* = 110, *df(model)* = 5, R^2 = 0.516, *β* = coefficient, *Std Err* = standard error, t = *t*-value, *p* = p-value, *** = p-value below 0.001, ** = p-value below 0.01, ‘ = p-value below 0.1.

The R^2 goodness-of-fit measure reported in Table 3 shows that our model can explain 52% of the variance in hospital costs with only these five explanatory variables. Figure 1 illustrates these good predictive properties of the model by depicting the relationship between model predictions of hospital costs and observed costs across hospital types. This result on goodness-of-fit can further be put into perspective by comparing it to a “full” model, including all nine preselected explanatory variables listed in Table 1. Including all these nine variables (see Table S2 in the Supporting Information) increases R^2 marginally to 53%, but at the cost of more variables, higher complexity, and multicollinearity within the model. To compare issues of multicollinearity between the two different models, *VIF* values of the five-variables model are 2.11 (number of discharges), 1.61 (ratio of emergency/ ambulance admissions), 1.74 (rate of DRGs to patients), and 1.62 (expected loss potential based on DRG mix), 1.22 (location in large agglomeration), which is not indicative of issues of multicollinearity. In contrast, in the nine-variables model, the four *VIF* values of expected loss potential based on DRG mix (11.09), ratio of rare DRGs (9.27), case-mix index (7.61), and ratio of children (5.10) are above the commonly accepted threshold of 5.0, which demonstrates issues of multicollinearity.

Fig. 1 *Scatterplot of cost predictions from the model presented above and observed hospitals costs*



Note. Dotted line = bisector line, Dot colors denote hospital types (see legend), and dot sizes illustrates hospitals' patient volume. University = university hospitals, Central = central general hospitals, Regional = regional general hospitals, Birth = birth centers, Children = children's hospitals, Specialty = specialty hospitals.

Furthermore, comparing model results of our five-variables model estimated with different weights shows that patient volume weights somewhat increase R^2 to 59%, whereas calculating an unweighted model slightly decreases R^2 to 42%. However, the model summaries of these two differently weighted models in Tables S3 and S4 of the Supporting Information reveal that the impact of all variables remains relevant across different weighting approaches, which indicates that the uncovered relationships are important independent of how smaller and larger hospitals are weighted in the model.

3.2. Comparisons across hospital types and years

Table 4 shows that mean absolute percentage differences and standard deviations of residuals were comparable across different hospital types but slightly higher for birth centers and specialty hospitals, which exhibit lower patient volumes compared to other hospitals. This (expected) relationship between lower patient volume and larger residuals is further illustrated in Figure S1 in the Supporting Information (see also the discussion section below). These results demonstrate that the selected model does not systematically over, or under-predict costs of certain hospital types.

Table 4 *Mean and standard deviations (SD) of patient volume and model prediction residuals for different hospital types*

	Mean (SD) **Residuals**	***Mean (SD)*** **Volume**
University	0.025 (0.018)	43,359 (4,311)
Central	0.039 (0.026)	17,528 (9,223)
Regional	0.055 (0.041)	4,701 (2,549)
Birth	0.07 (0.045)	376 (286)
Children	0.05 (0.053)	6,182 (2,092)
Specialty	0.062 (0.04)	2,840 (2,087)

Note. Residuals are reported in % of hospital costs (from 0 to 1.0). University = university hospitals, Central = central general hospitals, Regional = regional general hospitals, Birth = birth centers, Children = children's hospitals, Specialty = specialty hospitals.

In addition, the model summary in Table 5 shows that the inclusion of dummy variables for hospital types in an "extended" model did not explain significant additional variance in hospital costs (R^2 = 53%), indicating that cost predictions based on the five selected variables are capturing all differences between hospital types. This finding is further confirmed by the model summary presented in Table S5 of the Supporting Information, which displays model results for a second extended model that includes not only dummy variables for hospital types but also interactions between hospital types and cost predictions. Finally, comparisons across all the above-mentioned models (with five variables, nine variables, with dummy variables, and with interactions) are summarized in Table 6, which demonstrates that our five-variables model shows the best properties across several measures of model comparison (*adjusted* R^2, *AIC/BIC*, and *VIF*), particularly in consideration of its limited number of variables.

Table 5 *WLS regression results of the "extended" model using predictions from the model presented above and dummy variables for hospital types to explain hospital costs (in CHF)*

	β	*Std Err*	*t*	*p*	
Intercept	1,552.91	1,560.63	1.00	0.322	
Predicted Costs	0.86	0.15	5.72	0.000	***
University	222.14	258.95	0.86	0.393	
Central	-132.22	128.10	-1.03	0.304	
Birth	-257.32	509.24	-0.51	0.614	
Children	271.02	404.14	0.67	0.504	
Specialty	-22.18	214.06	-0.10	0.918	

Note. Dependent variable = average case-mix adjusted hospitals costs, n = 116, *df(residuals)* = 109, *df(model)* = 6, R^2 = 0.532, *β* = coefficient, *Std Err* = standard error, t = *t*-value, p = p-value, *** = p-value below 0.001. University = University hospitals, Central = Central general hospitals, Birth = Birth centers, Children = Children's hospitals, and Specialty = Specialty hospitals, whereas regional general hospitals were used as reference category.

Table 6 *Model comparison results across all reported models*

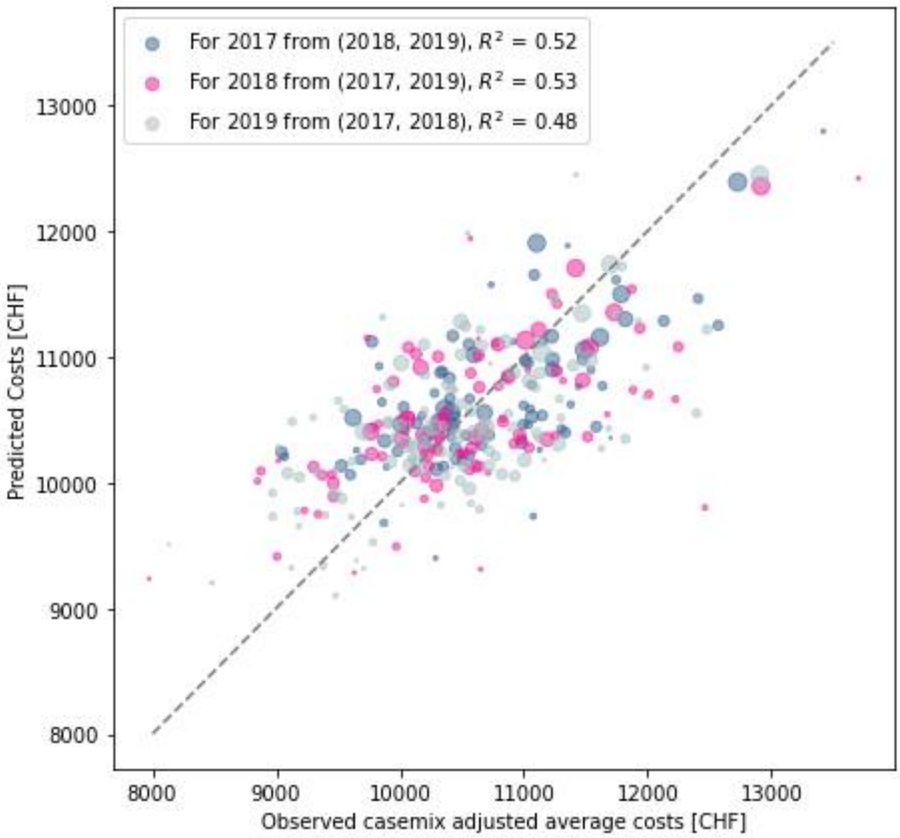
	Our Model	**Full Model**	**Extended 1**	**Extended 2**
***n* variables**[a]	5	9	10	15
R^2	0.52	0.53	0.53	0.56
Adjusted $R^{2,a}$	0.49	0.49	0.5	0.52
AIC[a]	1,841.37	1,845.76	1,843.75	1,845.01
BIC[a]	1,857.89	1,873.29	1,874.04	1,889.07
***n VIF* > 5.0**	0	4	0	10
max VIF	2.11	11.09	1.96	1,408.38

Note. Our model refers to the five-variables version, full model to the nine-variables version, extended 1 to the model with additionally included dummy variables, and extended 2 to the model with additionally included interaction variables. *n* Variables = number of variables in model, R^2 = coefficient of determination, *adjusted* R^2 = coefficient of determination adjusted for number of variables, *AIC* = Akaike's Information Criteria, *BIC* = Bayesian Information Criteria, *n VIF* > 5.0 = number of VIF values above 5.0, *max VIF* = maximum VIF value. [a] *n* variables, *adjusted* R^2, *AIC*, and *BIC* are adjusted for the number of variables that are additionally included in the predicted costs variable of the extended models.

Cross-validation results presented in Figure 2 reveal that the proportion of explained variance in hospital costs remains constant across train-test splits of different years (2017-2019), indicating consistency and robustness of the model within the population of Swiss hospitals captured by our dataset (despite differing sample sizes across the years, see method section).

In a final step, Table 7 compares the distribution of observed case-mix adjusted hospital costs with the distribution of predicted costs (using our five-variables model). This brings us back to the cost heterogeneity presented at the beginning of this chapter and illustrates the practical relevance of our outlined approach. By simply comparing observed case-mix adjusted hospital costs with some efficiency threshold (e.g., mean costs used in Table 7 to calculate "residuals observed") without taking structural differences into account, the cost residuals (i.e., deviations from chosen threshold) of individual hospitals vary widely between CHF -2,301.00 (-22%) and CHF 2,471.00 (24%). Such large cost residuals cannot realistically be expected to be compensated by efficiency gains of hospitals. However, by providing an individual benchmark for each hospital in the form of its predicted costs (taking its structural characteristics into account), the range of the residuals of individual hospitals ("residuals predicted" in Table 7) is compressed. This makes overcoming the remaining cost differences by efficiency increases much more realistic and attainable for individual hospitals.

Fig. 2 *Scatterplot of cross-validation results across several years (2017-2019)*



Note. Dotted line = bisector line, Dot colors denote years (see legend), and dot sizes illustrate hospitals' patient volume.

Table 7 *Comparison of the distributions of observed and predicted costs (in CHF)*

	Observed Costs	Residuals Observed	Predicted Costs	Residuals Predicted
Mean	10,433.73	0	10,428.05	5.68
SD	866.81	866.81	598.61	660.89
Min	8,133.33	-2,471.12	9,106.86	-1,384.66
10th	9,316.41	-1,140.57	9,708.57	-815.9
25th	9,906.24	-512.16	10,120.48	-398.56
50th	10,451.4	-17.67	10,346.18	-14.46
75th	10,945.89	527.48	10,751.55	447.96
90th	11,574.3	1,117.31	11,126.79	800.27
Max	12,904.84	2,300.39	12,219.69	1,535.12

Note. Residuals observed depict cost differences (in CHF) of individual hospitals from the mean, whereas residuals predicted stand for the cost differences (in CHF) of individual hospitals by subtracting predicted from observed costs. *SD* = standard deviation, *Min* = minimum, *10th - 90th* = 10th - 90th percentile, *Max* = maximum.

4. DISCUSSION

Previous research has identified several organizational and regional factors explaining justifiable cost variations among hospitals [5-7]. However, these findings have not yet been incorporated into policy and reimbursement practice in Switzerland. Apart from possible barriers in the legislative process and the vested interests of different stakeholders, we suspect that earlier proposals of including structural factors have not been pursued nationally because they either did not sufficiently differentiate hospital costs, or their underlying models were too complex to be implemented in practice. The approach we present herein aims to take a middle way by building an econometric model that can well explain observed cost variations, while at the same time being parsimonious in its model specification and includes only factors that can hardly be manipulated by hospitals. Our proposed regression model uses only five variables to explain over 50% of the observed variations in average case-mix adjusted hospital costs, it fully explains the otherwise observed differences in costs across hospital types, and it produces robust results across several years.

4.1. Selected model and variables

The selected explanatory variables for our model were chosen from different structural characteristics of hospitals that have been shown to be associated with case-mix adjusted costs in previous research (for a more detailed discussion of these and other variables, see [9]). Number of discharges describes the size and complexity within hospitals, which has been found to be positively correlated with costs both in Switzerland and internationally [6, 7, 10, 11]. The cost-increasing complexity that this and similar variables (like number of beds or number of departments, see [9]) describe appear to exceed the potential cost-decreasing effects of the larger size of hospitals due to economies of scale. Ratio of emergency/ ambulance admissions conveys the extent of emergency and first point of care functions that hospitals provide in their service area, which has also been related to higher costs in Switzerland and abroad [6, 7, 12-14]. It was argued that the provision of these services requires additional upfront investments and maintenance of sufficient reserve capacity, for which compensations by the Swiss DRG system are not specifically enough directed towards the hospitals that predominantly provide them. Rate of DRGs to patients illustrates the fact that some hospitals are required to treat a broader spectrum of conditions within their area of service compared to others. This variable has been found to explain cost variations well within the group of small regional hospitals [9], suggesting that particularly in small regional hospitals, the costs needed to provide a broad spectrum of services cannot be distributed across a high enough number of patients.

Expected loss potential based on DRG mix represents the fact that certain hospitals treating a higher proportion of rare and/ or more complicated cases in Switzerland tend to have more frequent high-cost outliers, whose losses cannot be compensated sufficiently by gains from low-cost outliers. Previous research has shown that this is characteristically the case in university and children's hospitals [5, 16-18]. In contrast, birth centers (performing low-risk deliveries exclusively without complications or complicating procedures) typically lie on the opposite end of this spectrum. It has been argued that the more severe and complex high-cost cases that university and children's hospitals treat require specialized skills and/ or equipment, for which higher wages and upfront capital investments are not sufficiently compensated by the Swiss DRG system. Finally, among the investigated regional differences, location in large agglomerations was the variable that best captured regional cost variations among hospitals. More precisely, hospitals with locations in large agglomerations showed higher costs than hospitals in smaller agglomerations or rural areas. Similar findings between urban and rural hospitals have been found in other countries [10, 15, 19], presumably as a result of differences in salary levels, living costs, and real estate prices in large agglomerations. As was noted in the method section (see also [9]), the calculation of our regional variables also takes different locations of hospitals with multiple locations and the vicinity around the location of hospitals into account.

Although several of these aspects have been proposed in previous approaches aimed at including structural characteristics into cost benchmarking and hospital payments in Switzerland [6, 7], we carefully designed and selected our variables to overcome weaknesses in these previous approaches. One previously proposed approach [7] used several collinear variables (such as the number of discharges and the number of DRGs) and applied principal component analysis, which made comprehensibility and interpretation of model results difficult. In contrast, we specifically designed and selected variables to be independent from each other and to avoid the necessity of using principal component analysis. For instance, by using rate of DRGs to patients (rather than number of DRGs), we made this variable independent from the otherwise related variable number of discharges. The second previously proposed approach [6], on the other hand, used the actual rates of high-deficit (i.e., high-loss) patients in hospitals to portray financially relevant differences in patient mix that are not sufficiently compensated by the Swiss DRG system. However, by using actual rates of high-deficit patients of hospitals as variable to explain cost variations, we cannot rule out that efficiency differences of hospitals impact the frequency or losses of such high-deficit patients. In contrast, the calculation of our variable expected loss potential based on DRG mix is only based on probabilities and costs from all cases in the country without incorporating (in-)efficiencies of individual hospitals (see method section or [9] for more details).

Our model containing only these five structural hospital attributes explained over 50% of the variance in average case-mix adjusted costs. This demonstrates that a large part of the variance in the costs of Swiss hospitals can be explained by systematic exogenous factors rather than (in-)efficiencies of hospitals. Furthermore, predictions of the model are able to capture all differences across the different hospital types, which was demonstrated by the result that none of the included hospital type dummy variables in the “extended” models provided additional explanatory power. Moreover, cross-validation results across several years showed the consistency and robustness of model estimates, even though hospital samples differed across the years 2017-2019 (e.g., with 22.5% fewer hospitals present in the sample of 2017 compared to 2019). This finding reinforces our interpretation that our model captures key structural and exogenous attributes explaining cost variations across hospitals rather than endogenous aspects that can be manipulated by individual hospitals, which could potentially bias estimates in a certain direction.

As can be observed in Figure 1 and Table 4, the presented model does not systematically over, or under-predict costs of certain hospital types, but the relationship between patient volume and residual size shows that variation in residuals becomes larger with lower patient volume in hospitals. However, the reason for this finding is not lower model accuracy for smaller hospitals, but rather the increasing variance in their hospital costs, which is in line with reports from other studies (see, e.g., [20]) and has led us to utilize the reported weighting approach. This means that not only model predictions of costs but also the observed and annually reported costs are less accurate in hospitals with smaller patient volumes. This is relevant especially for birth centers, which generally have a much lower patient volume than other hospitals, and to a lesser degree also in some specialty clinics. In addition, in specialty hospitals, lower model prediction accuracy is expected because this group contains a highly diverse set of hospitals, some of which are specialized in a single field. This uniqueness makes it impossible to identify systematic factors explaining cost differences, even though some of these differences may be due to exogenous factors.

4.2. Practical application and relevance

A modeling approach as presented herein may be applied in several practical ways. First, it can be used to differentiate justifiable cost differences based on uncontrollable exogenous factors from cost differences due to (in-)efficiencies of hospitals by comparing observed average case-mix adjusted costs of hospitals with their expected costs based on model predictions. In practice, this can be achieved by simply dividing observed costs by expected costs, as in the field of quality indicators, where such "observed-to-expected" ratios are commonly used to assess hospital performance (see, e.g., [21]). In our case, a resulting ratio above 1.0 would indicate that a hospital operates less efficiently than would be expected in comparison to other hospitals practicing under similar conditions, while a ratio below 1.0 would reveal a more efficient operation than expected. In that sense, hospitals' expected costs based on model predictions could act as an individual benchmark for hospitals.

Second, by multiplying the observed-to-expected ratios with a "national" base rate, calculated across all cases of all hospitals in a country, a form of risk-adjusted costs could be calculated for each hospital. Such risk-adjusted costs would inform about hospitals' operating efficiency without confounding influences of structural differences. Because of that, the risk-adjusted costs would provide a more suitable and objective alternative as benchmarking measure of operating efficiency than comparing average case-mix adjusted costs that are unadjusted for these structural differences (as it is currently practiced in Switzerland). Hereby, risk-adjustment does not only decrease cost variations significantly across hospitals (by roughly 25%, as has been explained with regard to the predicted costs presented in Table 7 above), but the remaining cost variations are also easier interpretable and actionable for hospitals and policymakers because of the adjustments made for key structural differences between the hospitals. This practice is again not new and an established approach in quality indicators comparing hospital performance (see, e.g., [21]).

Third, model predictions informing about expected costs of different hospitals based on structural differences could even be used as a basis to determine individual base rates for hospitals. Most countries with DRG systems use approaches to incorporate certain structural differences among hospitals into their payment rates. This is, for example, achieved by using different base rates for certain hospital peer groups [1]. In Switzerland, base rates are legally required to be negotiated between hospitals and health insurers. Consequently, these negotiations would potentially allow for a systematic inclusion of organizational or regional factors related to cost variations among hospitals. However, a consensus on which structural factors to consider is lacking, and negotiations instead are often biased by particular interests of stakeholders. A modeling approach could provide objective support here. By agreeing beforehand on certain organizational and regional attributes that should be included in the prediction of expected hospital costs, hospitals and insurers could abstract discussions away from the interests of individual stakeholders. In addition, subsequent negotiations (if even necessary) could start out from a well-established and agreed-upon foundation, from which at most small adaptations would need to be made (for instance, for certain specialty clinics, whose structural differences could not be sufficiently incorporated in the model).

In our opinion, a data-based approach, as suggested here, that can derive predictions of expected or (in other words) justified costs for individual hospitals, is also superior to graduating base rates by treating certain peer groups (such as academic teaching hospitals or children's hospitals) separately from others [1]. On the one hand, because a modeling approach allows for the inclusion of several cost-associated attributes simultaneously, and on the other hand, because it enables a comparison of many hospitals across a broad spectrum of values. The latter is of particular importance in Switzerland and in general settings with limited sample sizes of certain hospital peer groups. In Switzerland, for instance, there are only five university and three children's hospitals. Such sample sizes of five or even only three observations are much too small to make comparisons within the peer groups of university and children's hospitals. Because of that, treating university and children's hospitals separately from other hospitals in Switzerland would defeat the entire purpose of benchmarking. In contrast, by including hospitals' expected loss potential based on DRG mix, we are portraying the fundamental reason that sets university and children's hospitals apart from other hospitals and enable benchmarking of them together with other hospital types. Consequently, we can compensate them for such justified structural differences, but without excluding other hospitals that face similar conditions, which would happen if university and children's hospitals were to be treated as separate peer groups. That way, by not treating certain hospital peer groups separately from other hospitals, there are no regulation-based limitations of a free market and hospitals can compete across a broader spectrum of business models.

As pointed out above, a limitation of our approach is the fact that not all justifiable differences across hospitals can be considered, but only attributes observable across a large enough number of hospitals. Thus, certain justifiable cost differences of individual hospitals may remain unadjusted for. However, this limitation applies to all empirically derived approaches (such as, for example, indicator-based quality comparisons as well). Moreover, even though the proposed approach could be easily adapted and applied in other countries, the results and the selected variables may be specific to the regulations and market situation in Switzerland.

Finally, we would like to reiterate that if an approach as suggested here is used for cost benchmarking and to design reimbursement schemes, the utilized model variables have to be largely exogenous (i.e., uncontrollable for hospitals). If variables cannot be considered as exogenous, then they should not be included in modelling to prevent gaming of the system and unwarranted payments. Admittedly, there is probably no variable that is strictly exogenous, but we have used a literature review, careful consideration of previous approaches, stakeholder dialogues, and finally an expert consensus to ensure that our chosen variables are largely uncontrollable for Swiss hospitals. This means that while certain variables (such as the number of discharges or the rate of DRGs to patients) may in theory be influenced by a hospital's actions, in practice large changes are not possible because of the service mandate that Swiss hospitals are obligated to provide in their service area (see also method section).

In addition, the proposed modeling approach is not restricted to the exact set of variables that we used. Rather certain variables could be removed, or added, or exchanged for others if deemed necessary (e.g., number of discharges for number of departments, which describes the same underlying relationship, see [9]). In fact, from a practical standpoint regarding the Swiss healthcare system, we recommend that a national committee composed of representatives from different stakeholders (policymakers, hospitals, health insurers, etc.) should be built to agree upon a specific set of structural variables stipulated to be considered in cost benchmarking and reimbursement. This committee could also oversee continuous development and updating of the modelling approach to take mid- and long-term changes in the Swiss healthcare system into account. Moreover, changes in the utilized modelling variables could be monitored over time and simple measures could be employed to mitigate undesirable incentives and attempts of gaming. For example, modelling variables could be calculated retrospectively over several years and/ or could be restricted to suffice certain requirements (e.g., a minimal number of cases per DRG for all DRGs included in the variable rate of DRGs to patients).

To summarize, we presented an econometric approach to compare hospital cost variations, which takes cost-determining structural aspects into account in a parsimonious and data-driven way to explain case-mix adjusted hospital costs. We showed that a large proportion of the variation in hospital costs can be explained by a relatively small number of systematic exogenous factors and that the resulting model predictions fully account for observed differences across the main types of hospitals. With that, our results provide novel evidence for an ongoing policy debate in Switzerland on how to include such uncontrollable aspects into cost benchmarking and payments of hospitals to generate a level playing field.

5. References

1. Busse, R., et al., *Diagnosis-Related Groups in Europe - Moving Towards Transparency, Efficiency and Quality in Hospitals.* . 2011.
2. Augurzky, B. and H. Schmitz, *Wissenschaftliche Untersuchung zu den Ursachen unterschiedlicher Basisfallwerte der Länder als Grundlage der Krankenhausfinanzierung*. 2013, RWI - Leibniz-Institut für Wirtschaftsforschung.
3. Krinsky, S., et al., *Variation in Payment Rates under Medicare's Inpatient Prospective Payment System.* Health Serv Res, 2017. **52**(2): p. 676-696.
4. Klein-Hitpaß, U. and D. Scheller-Kreinsen, *Policy trends and reforms in the German DRG-based hospital payment system.* Health Policy, 2015. **119**(3): p. 252-7.
5. Widmer, P., S. Spika, and H. Telser, *Leistunsgorientierte Vergütung mit dem Fallpauschalensystem SwissDRG*. 2015, Polynomics.
6. Widmer, P., M. Trottmann, and H. Telser, *Das Fallpauschalenmodell: Leistungsbezogene Basispreise unter SwissDRG*. 2015, Polynomics.
7. Einkaufsgemeinschaft HSK, *Benchmark SwissDRG Tarifjahr 2021*. 2020, HSK.
8. Verein SpitalBenchmark, *Benchmarkingverfahren der Schweizer Spitäler*. 2020, Verein SpitalBenchmark.
9. Havranek, M.M., et al., *Identification and assessment of a comprehensive set of structural factors associated with hospital costs in Switzerland.* PLoS One, in press.
10. Skolasky, R.L. and L.H. Riley, 3rd, *Medicare Charges and Payments for Cervical Spine Surgery: Association With Hospital Characteristics.* Spine (Phila Pa 1976), 2015. **40**(16): p. E936-42.
11. Watanabe, Y. and Y. Nakata, *Association between outpatient orthopedic surgery costs and healthcare facility characteristics.* Int J Health Care Qual Assur, 2018. **31**(3): p. 265-272.
12. Meacock, R., T. Doran, and M. Sutton, *What are the Costs and Benefits of Providing Comprehensive Seven-day Services for Emergency Hospital Admissions?* Health Econ, 2015. **24**(8): p. 907-12.
13. Solberg, B.C., et al., *Changes in hospital costs after introducing an intermediate care unit: a comparative observational study.* Crit Care, 2008. **12**(3): p. R68.
14. Thompson, R.T., et al., *Increased length of stay and costs associated with weekend admissions for failure to thrive.* Pediatrics, 2013. **131**(3): p. e805-10.
15. Wang, L. and D.B. Stewart, *Increasing hospital costs for Clostridium difficile colitis: type of hospital matters.* Surgery, 2011. **150**(4): p. 727-35.
16. Merenstein, D., B. Egleston, and M. Diener-West, *Lengths of stay and costs associated with children's hospitals.* Pediatrics, 2005. **115**(4): p. 839-44.
17. Raol, N., et al., *Inpatient Pediatric Tonsillectomy: Does Hospital Type Affect Cost and Outcomes of Care?* Otolaryngol Head Neck Surg, 2016. **154**(3): p. 486-93.
18. Widmer, P., P. Hochuli, and H. Telser, *Theoretische und empirische Analyse zu den Mehrkosten der Kinderspitäler unter SwissDRG*. 2017, Polynomics.
19. Robinson, B.R., et al., *Cost discrepancies for common acute care surgery diagnoses in Ohio: influences of hospital characteristics on charge and payment differences.* Surgery, 2014. **156**(4): p. 814-22.
20. Grenda, T.R., R.W. Krell, and J.B. Dimick, *Reliability of hospital cost profiles in inpatient surgery.* Surgery, 2016. **159**(2): p. 375-80.
21. Agency for Healthcare Research and Quality, *Quality Indicator Empirical Methods*. 2019, Agency for Healthcare Research and Quality.

6. Statements

6.1 Data availability

The data that support the findings of this study are owned by a third-party organization, which collects and confidentially benchmarks financial data among Swiss hospitals. Restrictions apply to the availability of these data, which were used under strict confidentiality and data protection agreements. However, a similar dataset is available by the Swiss Federal Office of Statistics (contact via gesundheit@bfs.admin.ch) for researchers who meet the criteria for access.

6.2 Funding

This study was funded by the Research Grant 38988.1 IP-SBM of the Swiss Innovation Agency (Innosuisse), which promotes science-based innovations commissioned by the Swiss Confederation.

6.3 Conflict of interest

The authors have no conflicts of interest to declare that are relevant to the content of this article. PKW and SS are employed at two Swiss hospitals (see affiliations), but the hospitals were in no sense involved in this study.

6.4 Acknowledgments

The authors thank the members of the board of directors of "Verein SpitalBenchmark" and in particular its director, Ernst Frank, as well as Fabian Gfeller from SwissDRG for various critical discussions and valuable suggestions.