Int. J. Nonlinear Anal. Appl. In Press, 1–13
ISSN: 2008-6822 (electronic)
http://dx.doi.org/10.22075/ijnaa.2023.29842.4274

Costs of quality assurance in the German medical market

Majid Tehrani

Department of Business Administration, Universidad Catlica San Antonio de Murcia, Murcia, Spain

(Communicated by Javad Vahidi)

Abstract

The central subject of this study is the economic analysis of Quality Assurance (QA) in the German healthcare system. With the help of the study, an estimate of the total costs in the German healthcare market shall be given. The focus of the analysis is on Companies, Political Bodies, Liberal Professions and Public Corporations which are part of the healthcare system including Certification Bodies, which lead to costs from using QA and interacting with the healthcare system. First, a systematic literature search was conducted to determine the costs. It was found that there are no articles or publications that address the topic of the total cost of QA. A continuation/update of existing studies was therefore not possible. To be able to estimate the total costs of QA in the German healthcare market, the QA costs were surveyed using a bottom-up analysis. After identifying organizations and collecting relevant data, the total costs of QA in the German healthcare market were estimated using a mathematical calculation. In general, ensuring quality is an original part of the actions of all professional groups and institutions working in the health care system. Due to this importance, it is remarkable that an economic analysis of the total costs has never taken place before. One reason for this may be the "complexity of the German health care system". Furthermore, the costs of QA are not listed separately, but as part of general administrative expenses. Controlling and transparent presentation of the costs is therefore not possible. The cost estimation and the database created for this study about the parties involved in QA in the German healthcare market can be a useful support for further studies in this field of research.

Keywords: quality assurance, costs of quality assurance, German medical market, health care, quality management, administrative costs
2020 MSC: 91G15, 91B54

1 Introduction

In the German statutory health insurance system (Gesetzliche Krankenversicherung-GKV), there is not only an efficiency requirement, but also a quality requirement: the quality and effectiveness of services must correspond to the generally recognized state of medical knowledge and take medical progress into account (§2 (1) sentence 3 SGB V). To ensure this, hospitals, contracted physicians and medical care centers are obliged to participate in measures for inter-institutional quality assurance (QA) and to introduce and further develop quality management (QM) within the institution. In a narrower sense, these measures for QA across facilities and for QM within facilities represent the measures for quality development of patient care in the GKV system that have been subsumed under the umbrella term "quality assurance" to date.

In the meantime and against the backdrop of this quality requirement, an increasingly complex set of regulations for mandatory QA has emerged. As an interim result, however, it must be stated that mandatory QA is increasingly failing

Email address: info@m-tehrani.de (Majid Tehrani)

Received: November 2022 *Accepted:* February 2023

to fulfill its original purpose of promoting quality. This is due, on the one hand, to a backlog in system maintenance and continuous further development of QA procedures, and, on the other hand, to an increasing neglect of QM within facilities [6]. If, for example, against the backdrop of the Corona pandemic, the QA procedure “community-acquired pneumonia” is perceived only as a burden and not as useful, both for the patients concerned and for the hospitals, this must be seen as a serious alarm signal for the internal state of the facility-internal quality culture in German hospitals.

Even the newer goals of QA - cross-sectoral QA and quality-oriented care management – have not yet been achieved. However, the reasons for this delay in a quality push in the healthcare system cannot be found in QA alone, but also in the framework conditions of the healthcare system [6].

Although scientific research has been late in addressing the topic of QA, the number of publications has grown rapidly [7, 10, 5, 2].

In fact, in recent years there have been many proposals to ensure and manage the quality of treatment in medical institutions more rigorously. It is often argued that QA and QM have been introduced late in the health care sector compared to industry [3, 11].

Other authors point out that QA is of course not new in medicine, but takes another approach and has other indications to comparable projects in other industries. This includes, for example, the extremely long training of providers, the manifold medical certificates or the posthumous diagnosis confirmation by pathological examination [9].

Against this background, a wealth of methods, measures and organizations have been established in Germany with the aim of maintaining or improving the quality of medical care and patient safety. In view of the effort and costs involved, there have been repeated calls for established QA or QM programs to be reviewed for their practicality, effectiveness and cost appropriateness [4].

For some years now, an increased influence of non-medical experts on the health care system and especially its organizations has been observed. On the one hand, this is evident under the term Evidence Based Medicine (EBM), which pushes health teaching too strongly in the direction of science with a plethora of regulations, guidelines, directives and standards. On the other hand, under various concepts such as QA, quality promotion, QM, quality control, total quality management, benchmarking, balanced score card, accreditation, certification, process management, etc., a sector, managed by experts, has become institutionalized [8].

The self-proclaimed goal is to improve care performance and provide necessary management processes. While the Model of Business Excellence of the European Foundation for Quality Management (EFQM) is usually oriented to the management concept as well as the German Institute for Standardization (DIN ISO 9000 ff.), which develops content and formal standards for the evaluation of organizations, certification procedures of various.

With the help of the study an estimate of the total costs in the German health care market shall be given. The focus of the analysis is on companies, organizations, associations and institutions which are part of the health care system including certification bodies, which lead to costs from using QA and interacting with the health care system.

Methodological approach

The aim of the empirical study was to provide an estimate of the total costs of QA in the German health care market.

First, an attempt was made to address this question within the framework of systematic research.

The original intention was to examine and analyze articles and publications that dealt with this question, e.g. for subsectors of the German health system at national level.

The literature search was carried out in the Medline databases via `https://pubmed.ncbi.nlm.nih.gov/`, in the Cochrane Library via `http://www.cochranelibrary.com` and additionally via other hand search sources. These sources included references from the bibliographies of relevant publications, “related citations” of Medline hits and books. In order to be able to record further literature, including “grey literature”, a search was also conducted on selected websites of topic-relevant national institutions that are responsible for QA/QM programs and by using Google Scholar as a search engine.

1. **Medline via** `https://pubmed.ncbi.nlm.nih.gov/`

   The search strategy documented in the table below included the keywords: “quality assurance healthcare”, “quality management healthcare”, and “administrative costs healthcare”. Medline already provided MeSH

Terms for the selected keywords.

Due to the high hit rate and imprecise results in searches #1 and #2, the search keywords had to be combined. The hit rate was reduced considerably, but the result was not satisfactory.

Narrowing down by time and selecting German titles significantly reduced the hits and finally made the results more precise. Search #5 returned 181 results. The first 100 hits were used for the analysis.

Nr.	Search Keyword	Results
#5	#1 AND #2 Filters: Publication date from 2011/01/01 to 2020/12/31; German	176
#4	#1 AND #2 Filters: Publication date from 2011/01/01 to 2020/12/31; English; German	11,346
#3	#1 AND #2 (quality assurance healthcare [MeSH Terms] OR quality management healthcare [MeSH Terms] OR quality assurance program* OR quality management program*) AND ("costs" OR "administrative costs" healthcare [MeSH Terms])	24,531
#2	"costs" OR "administrative costs" healthcare [MeSH Terms]	132,171
#1	quality assurance healthcare [MeSH Terms] OR quality management healthcare [MeSH Terms] OR quality assurance program* OR quality management program*	523,772

2. **Cochrane Library**

Analogous to the search under Medline, the same procedure was applied, which is documented below. However, in order to achieve corresponding results, the term "quality assurance program" and "quality management program" were added in addition to Medline.

Again, the terms from searches #1 and #2 had to be combined to achieve more consistent results.

Search #3 returned 406 hits. The first 100 hits were evaluated for the analysis.

Nr.	Search Keyword	Results
#3	#1 AND #2 Limit 2011 – 2020 Cochrane Database of Systematic Reviews	406
#2	"costs" OR "administrative costs" healthcare	1,464
#1	quality assurance healthcare OR quality management healthcare OR quality assurance program* OR quality management program*	692

3. **Google Scholar**

Google Scholar is not a medical database like Medline, but a web crawler. Therefore, it is not possible to search along a thesaurus. It is also not recognizable which exact algorithms Google Scholar uses searching. Which search terms are successful can therefore only be matter of trial and error.

The appropriate terms for the search were: "quality assurance healthcare" "quality management healthcare" "certifications healthcare" "costs" and "administrative costs healthcare".

As expected, there was an enormous hit rate with undesired results in searches #1 and #2. Only the combination of the two searches reduced the hit rate significantly and provided a better result.

Search #3 returned 422 hits. Of these, the first 100 hits were analyzed.

Nr.	Search Keyword	Results
#3	#1 AND #2 Filters: Publication date from 2011 to 2020; German	422
#2	allintitle: "costs" OR "administrative costs" healthcare	2,800,000
#1	allintitle: "costs" quality assurance healthcare OR "costs" quality management healthcare OR "costs" certifications healthcare	30,400

In this analysis, however, it turned out that there is no corresponding publication or summary study dealing with the total costs of QA in the German health care market.

To be able to estimate the total costs of QA in the German health care market, the QA costs were surveyed using a bottom-up analysis. For this study, the research question was broken down into two further sub-topics:

1. Which organizations exist in the German healthcare market that deal with the topic of QA?
2. What are the costs incurred for QA in these organizations?

The identification of organizations in the German health system that carry out QA measures was done exclusively by web search.

Governmental organizations, health insurance companies or specific areas of the health care system, such as political bodies, which proactively promote the topic of QA, could be identified and documented systematically for the use of this study, via the relevant homepage.

Explanation of the procedure using the example of health insurance funds:

- Identification of all health insurance funds in Germany via various sources, e.g. www.krankenkassen.de
- Analysis of each individual health insurance fund via the relevant homepage with regard to QA measures
- Investigation of possible cooperation partners
- Documentation in case of a positive hit in the database

As explained, institutional and state-owned companies were identified for the empirical analysis in terms of the research question via web search.

The desired data and corresponding cost structures of these organizations were collected and documented with the help of Statistisches Bundesamt (The Federal Statistical Office).

For the study of other organizations and companies, the approach had to be adapted.

In order to obtain relevant information, two questions were formulated and the identified parties were contacted by mail:

1. How many employees are employed in your company to deal with the topic of QA?
2. How are the total costs incurred in your company for QA estimated?

After re-checking and validating the collected data, the dimensions to be analysed were structured and divided into the four following clusters based on Thielscher [12]:

1. Political Bodies, QGD
2. Public Corporations and their Organizations
3. Companies
4. Liberal Professions and other

Cluster analysis

Internal evaluation

When a clustering result is evaluated based on the data that was clustered itself, this is called internal evaluation. These methods usually assign the best score to the algorithm that produces clusters with high similarity within a cluster and low similarity between clusters. One drawback of using internal criteria in cluster evaluation is that high scores on an internal measure do not necessarily result in effective information retrieval applications. Additionally, this evaluation is biased towards algorithms that use the same cluster model. For example, k-means clustering naturally optimizes object distances, and a distance-based internal criterion will likely overrate the resulting clustering. Therefore, the internal evaluation measures are best suited to get some insight into situations where one algorithm performs better than another, but this shall not imply that one algorithm produces more valid results than another [5]. Validity as measured by such an index depends on the claim that this kind of structure exists in the data set. An algorithm designed for some kind of models has no chance if the data set contains a radically different set of models, or if the evaluation measures a radically different criterion. For example, k-means clustering can only find convex clusters, and many evaluation indexes assume convex clusters. On a data set with non-convex clusters neither the use of k-means, nor of an evaluation criterion that assumes convexity, is sound. More than a dozen of internal evaluation measures exist, usually based on the intuition that items in the same cluster should be more similar than items in different clusters. For example, the following methods can be used to assess the quality of clustering algorithms based on internal criterion:

- **Davies-Bouldin index** The Davies-Bouldin index can be calculated by the following formula:

$$DB = \frac{1}{n}\sum_{j=1}^{n}\max_{j\neq i}\left(\frac{\sigma_i + \sigma_j}{d(c_i, c_j)}\right)$$

  where n is the number of clusters, c_i is the centroid of cluster i, σ_i is the average distance of all elements in cluster i to centroid c_i and $d(c_i, c_j)$ is the distance between centroids c_i and c_j. Since algorithms that produce clusters with low intra-cluster distances (high intra-cluster similarity) and high inter-cluster distances (low inter-cluster similarity) will have a low Davies-Bouldin index, the clustering algorithm that produces a collection of clusters with the smallest Davies-Bouldin index is considered the best algorithm based on this criterion.

- Dunn index
  The Dunn index aims to identify dense and well-separated clusters. It is defined as the ratio between the minimal inter-cluster distance to maximal intra-cluster distance. For each cluster partition, the Dunn index can be calculated by the following formula:

$$D = \frac{\min_{i\leq j\leq n} d(i,j)}{\max_{1\leq k\leq n} d'(k)}$$

  where $d(i,j)$ represents the distance between clusters i and j, and $d'(k)$ measures the intra-cluster distance of cluster K. The inter-cluster distance $d(i,j)$ between two clusters may be any number of distance measures, such as the distance between the centroids of the clusters. Similarly, the intra-cluster distance $d'(k)$ may be measured in a variety ways, such as the maximal distance between any pair of elements in cluster K. Since internal criterion seek clusters with high intra-cluster similarity and low inter-cluster similarity, algorithms that produce clusters with high Dunn index are more desirable.

- Silhouette coefficient
  The silhouette coefficient contrasts the average distance to elements in the same cluster with the average distance to elements in other clusters. Objects with a high silhouette value are considered well clustered, objects with a low value may be outliers. This index works well with k-means clustering, and is also used to determine the optimal number of clusters.

External evaluation

In external evaluation, clustering results are evaluated based on data that was not used for clustering, such as known class labels and external benchmarks. Such benchmarks consist of a set of pre-classified items, and these sets are often created by (expert) humans. Thus, the benchmark sets can be thought of as a gold standard for evaluation. These types of evaluation methods measure how close the clustering is to the predetermined benchmark classes. However, it has recently been discussed whether this is adequate for real data, or only on synthetic data sets with a factual ground truth, since classes can contain internal structure, the attributes present may not allow separation of clusters or the classes may contain anomalies. Additionally, from a knowledge discovery point of view, the reproduction of known knowledge may not necessarily be the intended result. scenario of constrained clustering, where meta information (such as class labels) is used already in the clustering process, the hold-out of information for evaluation purposes is non-trivial.

A number of measures are adapted from variants used to evaluate classification tasks. In place of counting the number of times a class was correctly assigned to a single data point (known as true positives), such pair counting metrics assess whether each pair of data points that is truly in the same cluster is predicted to be in the same cluster.

As with internal evaluation, several external evaluation measures exist, for example:

- **Purity:**
  Purity is a measure of the extent to which clusters contain a single class. Its calculation can be thought of as follows: For each cluster, count the number of data points from the most common class in said cluster. Now take

the sum over all clusters and divide by the total number of data points. Formally, given some set of clusters M and some set of classes D, both partitioning NV data points, purity can be defined as:

$$\frac{1}{N}\sum_{m\in M}\max_{d\in D}|m\cap D|$$

This measure doesn't penalize having many clusters, and more clusters will make it easier to produce a high purity. A purity score of 1 is always possible by putting each data point in its own cluster. Also, purity doesn't work well for imbalanced data, where even poorly performing clustering algorithms will give a high purity value. For example, if a size 1000 dataset consists of two classes, one containing 999 points and the other containing 1 point, then every possible partition will have a purity of at least 99.9.

- **Rand index**
The Rand index computes how similar the clusters (returned by the clustering algorithm) are to the benchmark classifications. It can be computed using the following formula:

$$RI = \frac{TP + TN}{TP + FP + FN + TN}$$

where TP is the number of true positives, TN is the number of true negatives, FP is the number of false positives, and FN is the number of false negatives. The instances being counted here are the number of correct pairwise assignments. That is, TP is the number of pairs of points that are clustered together in the predicted partition and in the ground truth partition, FP is the number of pairs of points that are clustered together in the predicted partition but not in the ground truth partition etc. If the dataset is of size N, then $TP+TN+FP+FN = \binom{N}{2}$.

One issue with the Rand index is that false positives and false negatives are equally weighted. This may be an undesirable characteristic for some clustering applications. The F-measure addresses this concern,[citation needed] as does the chance-corrected adjusted Rand index.

- **F-measure**
The F-measure can be used to balance the contribution of false negatives by weighting recall through a parameter $8 > 0$. Let precision and recall (both external evaluation measures in themselves) be defined as follows:

$$P = \frac{TP}{TP + FP}$$
$$R = \frac{TP}{TP + FN}$$

where P is the precision rate and R is the recall rate. We can calculate the F-measure by using the following formula.

$$F_\beta = \frac{(\beta^2 + 1)\cdot P \cdot R}{\beta^2 \cdot P + R}$$

When $\beta = 0$, $F_0 = P$. In other words, recall has no impact on the F-measure when $\beta = 0$, and increasing allocates an increasing amount of weight to recall in the final F-measure.

Also TN is not taken into account and can vary from 0 upward without bound.

- **Jaccard index**
The Jaccard index is used to quantify the similarity between two datasets. The Jaccard index takes on a value between 0 and 1. An index of 1 means that the two dataset are identical, and an index of 0 indicates that the datasets have no common elements. The Jaccard index is defined by the following formula:

$$J(A,B) = \frac{|A\cap B|}{|A\cup B|} = \frac{TP}{TP + FP + FN}$$

This is simply the number of unique elements common to both sets divided by the total number of unique elements in both sets. Also T'N is not taken into account and can vary from 0 upward without bound.

- **Dice index**
The Dice symmetric measure doubles the weight on TP while still ignoring TN:

$$DSC = \frac{2TP}{2TP + FP + FN}$$

- **Fowlkes-Mallows index**
  The Fowlkes-Mallows index computes the similarity between the clusters returned by the clustering algorithm and the benchmark classifications. The higher the value of the Fowlkes-Mallows index the more similar the clusters and the benchmark classifications are. It can be computed using the following formula:

$$FM = \sqrt{\frac{TP}{TP+FP} \cdot \frac{TP}{TP+FN}}$$

  where TP is the number of true positives, FP is the number of false positives, and FN is the number of false negatives. The FM index is the geometric mean of the precision and recall P and R, and is thus also known as the G-measure, while the F-measure is their harmonic mean.

  Moreover, precision and recall are also known as Wallace's indices BI and BI Chance normalized versions of recall, precision and G-measure correspond to Informedness, Markedness and Matthews Correlation and relate strongly to Kappa.

- The mutual information is an information theoretic measure of how much information is shared between a clustering and a ground-truth classification that can detect a non-linear similarity between two clusterings. Normalized mutual information is a family of corrected-for-chance variants of this that has a reduced bias for varying cluster numbers.

Results

The data sets used for the analysis, which consist of responses from the organizations and self-researched data, are summarized in the table below.

Group	Number of Organizations	EMP	EMP in QA	AC	AC in QA
Political Bodies	8	5	0	6	1
Public Corporations	32	20	13	3	2
Companies	66	33	12	3	2
Liberal Professions	56	25	10	1	0
Overall	162	83	35	13	5

- Number of Organizations: The column indicates how many organizations were contacted in the respective cluster. A total of 162 organizations were contacted and researched.
- EMP (Employees), EMP in QA: The columns indicate how many organizations could be researched in terms of total number of employees and employees in QA.
- AC (Administrative Costs), AC in QA: The columns indicate how many organizations could be researched in relation to administrative costs in general and in the area of QA.

The statistical key figures that allow a valid calculation and extrapolation of the data sets in the formed clusters are "linear regression", "standard deviation" and "coefficients of variation".

Thus, a representative value for the "cost per employee" can be determined using linear regression. When plotting the total costs against the number of employees, the slope reflects the value for the "costs per employee". However, in order to be able to appropriately examine the deviations within the selected clusters, which come about through the results of the linear regression, and to ensure a valid comparison of the clusters, the coefficient of variation (deviation coefficient) was chosen as a supplementary key figure. The advantage of the coefficient of variation over the standard deviation is that the coefficient of variation is indifferent to the scale on which the data were measured [1].

In this case, the coefficient of variation is calculated from the quotients of the standard deviation and costs per employee.

This described calculation basis was applied for clusters "Political Bodies", "Public Corporations" and "Companies". Insufficient data sets in the cluster "Liberal Professions" do not allow for a calculation of the coefficient of variation.

For the "Freelance Professions" cluster, the mean value of the other three clusters had to be used to calculate the "costs per employee" due to the lack of important data sets.

1. **Analysis of Cluster: Public Corporations**

For the calculation "costs per employee", two entities were used for which the data sets EMP total and total administrative costs were available. The result was used as a basis for the preliminary extrapolation of the missing data and for calculating the administrative costs of other entities in this cluster.

A closer look at this cluster reveals that a number of organizations have a low number of employees or the number of EMP-QA is exactly the same as the total number of EMP. A holistic view of the cluster with all available data in the matrix would therefore distort the statistical significance. In order to nevertheless obtain a more valid result, an additional key figure "percentage weighting" was chosen so that the entities mentioned do not affect the overall result too drastically.

Despite the introduction of the percentage weighting, it was determined after several calculations that IQTIG, with its 100 employees all working in QA at the same time, could not be excluded from "diluting" the overall result. Thus, IQTIG was excluded from the calculation.

The result of the cluster can be mapped as follows:

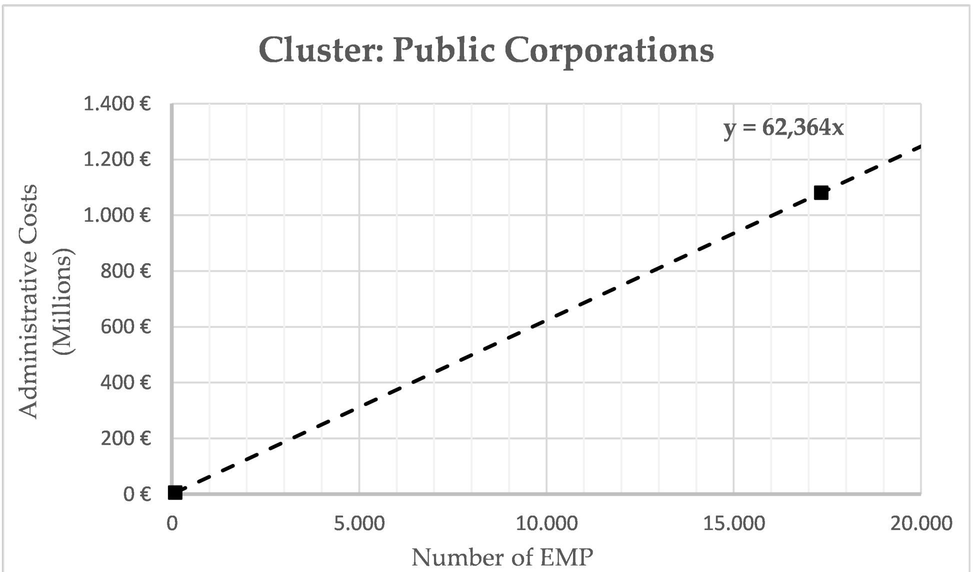
- The slope from the linear regression yields a value of: €62,364.00 as "cost per employee".

Due to the missing data sets, the key figures standard deviation and coefficient of variation cannot be calculated.

With the existing data records, the key figure "percentage weighting" was used to create the basis for extrapolating the percentage of EMP-QA. Of a total of 2,338 employees, 291 work in QA. Thus, the ratio of employees working in QA is 12.4%.

With the assumption that 12.4% of the employees work in QA, the preliminary extrapolation for this cluster results in:

- 20,537 EMP-QA with a volume of €1,280,392,920.00 incurred for QA administration costs.
- The largest share of this sum is made up by the GKV with over €1,037,924,052.00.



Cluster: Public Corporations				
Organizations	**#EMP**	**#EMP in QA**	**% of EMP in QA**	**Weighting [%]**
AQUA-Institute	90	60	66,7	3,8
AZQ	30	8	26,7	1,3
BAK	88	17	19,3	3,8
BAK	1429	122	8,5	61,1
DRK–Landesverband Baden-Württemberg e.V.	48	3	6,3	2,1
GKV-Spitzenverband	356	16	4,5	15,2
KCQ	7	5	71,4	0,3
KZBV	120	4	3,3	5,1
MDS e.V.	70	36	51,4	3,0
WidO	89	14	15,7	3,8
ZQ	11	6	54,5	0,5
Overall	**2,338**	**291**	**Total product: 12,4%**	

2. **Analysis of cluster: Companies**

Analogous to the cluster "Public Corporations" with the corresponding calculation basis, the "slope" was obtained from the linear regression.

The results of the cluster can be presented as follows:

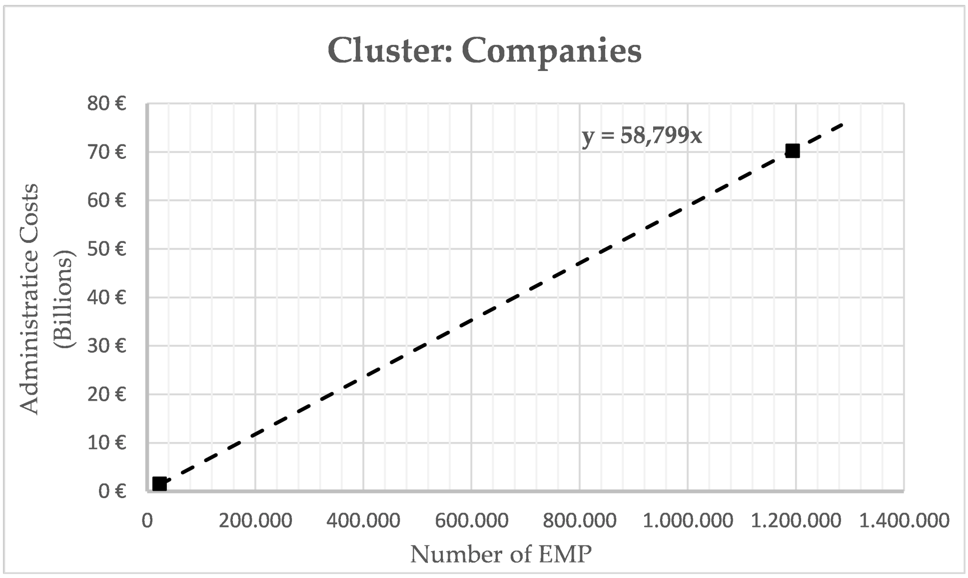
- The slope yields €58,799.00 "cost per employee"

TÜV NORD had to be excluded from this calculation because the administrative costs per employee amount to approx. €165,000 and would thus have distorted the calculation.

In this cluster, as well, the necessary data sets are missing in order to calculate the standard deviation and the coefficient of variation.

The percentage share of EMP-QA, calculated according to percentage weighting, is:

- EMP-QA share: 9,3%
- Accordingly, according to preliminary extrapolation: 462,411 EMP-QA
- The total volume for QA administration costs is: €20,692,314,246.10



Cluster: CompaniesW7933				
Organizations	**#EMP**	**#EMP in QA**	**% of EMP in QA**	**Weighting [%]**
BAR e.V.	32	5	15,6	0,8
BAV Institut	65	2	3,1	1,7
Cert iQ Zertifizierungsdienstleistungen e.V.	11	9	81,8	0,3
CertEuropA GmbH	10	6	60,0	0,3
ENPP-Boehm GmbH	13	13	100	0,3
i-med-cert GmbH	4	4	100	0,1
IQD	12	12	100	0,3
IQM e.V.	7	3	42,9	0,2
Kneipp-Bund e.V.	66	5	7,6	1,7
KTQ	6	3	50,0	0,2
T_V Nord Cert GmbH	3653	300	8,2	93,8
VDBD e.V.	14	1	7,1	0,4
Overall	**3,893**	**363**	**Total product: 9,3%**	

3. **Analysis of cluster: Liberal Professions**

   In this cluster, important data sets are missing that would have been necessary for the calculation of the key figures of the linear regression and the coefficient of variation. For the calculation and preliminary extrapolation based on the percentage weighting, the data sets of eight companies in the freelance professions could be used.

   The results are as follows:

   - The calculated percentage share of QA employees in this cluster is: 11.6%.

   To be able to determine the QS administrative costs, the average of the other three clusters was applied.

   - Cluster 1: €51,121.00
   - Cluster 2: €62,364.00
   - Cluster 3: €58,799.00

Cluster: Liberal Professions				
Organizations	**#EMP**	**#EMP in QA**	**% of EMP in QA**	**Weighting [%]**
DAKJ e.V.	3	3	100	1,2
DDG	10	2	20,0	4,1
DEGAM	19	2	10,5	7,9
Deutscher Paritktischer	140	4	2,9	57,9
DGHO e.V.	12	7	58,3	5,0
DGIM e.V.	10	5	50,0	4,1
DGK e.V.	34	4	11,8	14,0
VDBD e.V.	14	1	7,1	5,8
Overall	**242**	**28**	**Total product: 11,6%**	

4. **Analysis of cluster: Political Bodies**

   Analogous to the previous two clusters, the results for this cluster can be presented as follows:

   - "Cost per employee" based on linear regression: €51,121.00
   - Standard deviation: €12,889.98
   - Coefficient of variation: 25.3%

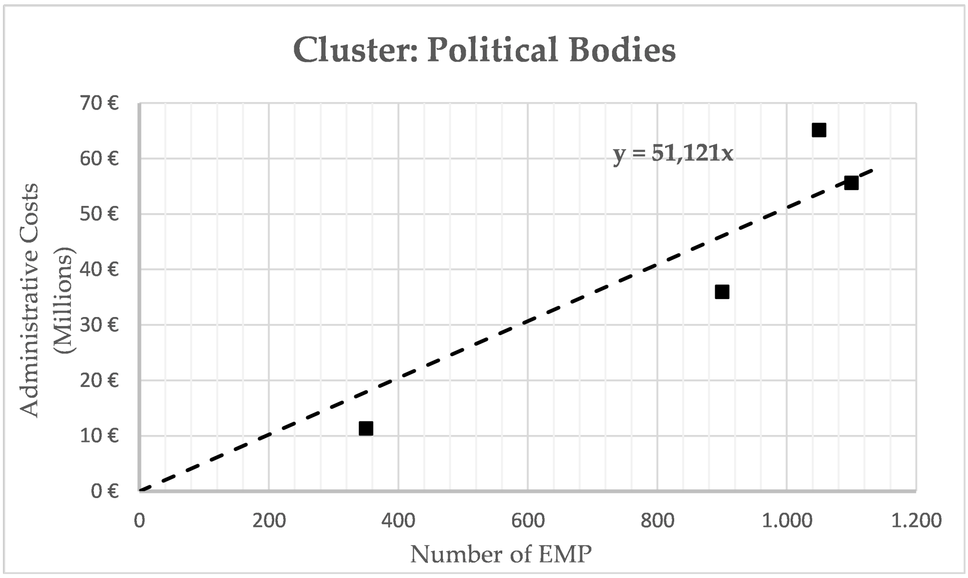
   Based on administrative costs and administrative costs QA of the RKI, the percentage share of employees in QA can be calculated.

   This percentage is: 4.5%.

For the calculation of the number of employees in QA, the calculated value of 4.5% can be applied to obtain the preliminary extrapolation of total administrative costs QA.

The extrapolated QA administrative costs are: €49,935,088.00 with an extrapolated share of employees in QA of 976.

Cluster: Political Bodies				
Organizations	**#EMP**	**#EMP in QA**	**AC**	**AC in QA**
BfArM	1,050	n/a (47)	€65,117,000.00	n/a (€2,402,687.00)
BMG	n/a (1,193)	n/a (54)	€60,979,000.00	n/a (€2,760,534.00)
BZgA	350	n/a (16)	€11,336,000.00	n/a (€817,936.00)
DIMDI	n/a (164)	n/a (7)	€8,403,000.00	n/a (€357,847.00)
Gesundheitskmter	17,000	n/a (763)	n/a (€869,057,000.00)	n/a (€39,005,323.00)
Paul-Ehrlich-Institut	900	n/a (41)	€35,894,000.00	n/a (€2,095,961.00)
Robert-Koch-Institut	1,100	n/a (48)	€55,601,000.00	€2,494,800.00
SVR	n/a	n/a	n/a	n/a
Overall	**21,757**	**976**	**€1,106,387,000.00**	**€49,935,088.00**



Aggregation of QA administrative costs and the final extrapolation

A total of 87 datasets were available across four clusters, which acted as the basis for calculating QA administrative costs.

Explanation of the extrapolation with missing data sets based on the second cluster:

For this cluster, 66 out of 34 data sets were available to be used for the calculation of QA administrative costs.

On the premise that the 34 records account for $\frac{34}{66}$, the QA administrative costs can be extrapolated to 100%, namely from €20,692,314,246.10 to €40,167,433,536.55.

In this way, clusters one and four can also be extrapolated to 100%, so that the final result of the four clusters for QA administration costs is €110,698,889,404.86.

Discussion

Quality assurance represents an important building block in the German healthcare system. Nevertheless, there is little to no transparency and publicly accessible sources of information dealing primarily with the cost structure in this subject area.

The lack of transparency and the quiet suspicion that installed quality assurance cannot necessarily contribute to increasing and improving quality, reinforces the impression that the identified companies would like to provide little to no information if possible. Even the state-run institutions, which have made QA their priority, were "consciously" or "unconsciously" unable to name the specific costs requested and referred to the legally standardized annual reports.

The research and analysis of public documents, such as annual reports or balance sheets, conducted here has shown that the costs of QA are not declared as a separate, important part of the company's operations. Rather, the costs of QA, and thus the administrative costs for it, are considered and presented as part of the general administrative costs. Controlling and transparent presentation of the costs is therefore not possible.

According to the Federal Statistical Office, spending on health - i.e., prevention, treatment, rehabilitation and care - rose steadily between 1992 and 2019 from 159.5 to 413.8 billion Euros. This corresponds to an increase of 3.6 percent per year. Due to the Corona pandemic, spending increased at an above-average rate from 2019 to 2020 - by almost 6.5 percent to 440.6 billion Euros.

The estimated total QA costs of over 110 billion Euros accordingly account for around a quarter of total costs.

According to a study by management consultants A. T. Kearney, administrative costs in the German healthcare system are apparently much higher than previously assumed. According to the study, bureaucracy accounted for 23 percent of the GKV system's total expenditure of 176 billion Euros in 2011.

The 23 percent administrative cost ratio of the healthcare system is - the study notes - 3.8 times higher than the average value in German industrial companies, which is 6.1 percent. In addition, the study concludes that 68 percent of the total administrative costs, or 27.5 billion Euros, are caused precisely by the GKV system. This corresponds to an actual administrative cost of 15.6 percent in relation to the 176 billion Euros in total expenditure. This share of administrative costs is 2.9 times higher than the 5.4 percent or 9.5 billion Euros of administrative costs officially reported by the GKV.

One of the main problems identified by the authors of the study is the large number of inspection bodies that have been created over the years. Legislators are constantly imposing new bureaucratic auditing tasks on the health insurers. No one has yet demonstrated so clearly that the obsession with saving money, which is focused on efficiency gains, could actually be a cause of the cost-driving.

According to calculations made in 2018 by Friedrich Breyer, holder of the Chair of Economic and Social Policy at the University of Konstanz, technical progress ensures an annual cost increase of two percent. In general, it can be stated that when health care spending increases, quality assurance costs can also be assumed to increase proportionally.

The estimated administrative burden of over 110 billion euros can thus essentially be attributed to increasing lack of transparency and complexity in the German healthcare system. As already outlined, a number of complexity drivers contribute to this, such as the large number of different players, an oversupply of products and services, a wide variety of IT systems, frequently changing reforms and laws, and interface problems due to processes that are not coordinated (Breyer 2018).

The parties involved are organized in silo-like structures with lone wolf interests and are represented by more than 300 lobby and interest groups in order to maintain or make their own services billable at the expense of the community and at the highest possible prices.

The system would benefit, for example, from an overall coordinated reduction in the number of statutory health insurance funds and optimization of the administrative apparatus of the associations of statutory health insurance physicians.

In addition, the complex portfolio of products and services and the associated administrative burden should be reduced to a level that makes sense for the overall system. It is also important to have lean, direct and continuous information flows that save costs and reduce interfaces.

References

[1] C. Duller, *Einführung in die nichtparametrische Statistik mit SAS, R und SPSS*, Ein anwendungsorientiertes Lehr- und Arbeitsbuch, Springer Gabler Berlin, Heidelberg, 2018.

[2] C. Grbe, *Patientenorientierung im Qualitätsmanagement im Gesundheitswesen*, Theoretische Grundlagen, gesetzliche Regelungen und eine sektorenübergreifende qualitative Studiem Springer Fachmedien Wiesbaden GmbH, 2021.

[3] J.H. Haeske-Seeberg, *Handbuch Qualitätsmanagement im Krankenhaus*, Kohlhammer, Stuttgart, 2001.

[4] A. Helou, F.W. Schwartz and G. Ollenschläger, *Qualitktsmanagement und Qualitätssicherung in Deutschland*, Bundesgesund. Gesundheits. Gesundheitsschutz **45** (2002), 205–214.

[5] P. Hensen, *Methoden und Instrumente des Qualitktsmanagements. In: Qualitätsmanagement im Gesundheitswesen*, Springer Gabler, Wiesbaden, 2019.

[6] R. Klakow-Frank, *Gesundheitswesen aktuell 2020*, Qualitätsent. Gesund. Defizite Lösungsvor. **2020** (2020), 60–77.

[7] M. Mörch, *Qualitktstransparenz für ambulante und stationkre Gesundheitsleitungen*, Peter Lang, Frankfurt, 2005.

[8] M. Nothacker, C. Muche-Borowski, and I.B. Kopp, *20 Jahre ärztliche Leitlinien in Deutschland – was haben sie bewirkt?*, Zeitsch. Evid. Fortb. Qual. Gesund. **108** (2014), no. 10, 550–559.

[9] O. Obst and E. Jandl, *Theorie und Strategie der Literatursuche*, Chirurgische Forschung, Krukemeyer, 2005.

[10] S. Poll, *Bedeutung von Krankenhausdatenbanken im Internet aus Nutzersicht - Eine informationsökonomische Analyse*, Dissertation, Trier, 2008.

[11] E. Schupeta and H. Hildebrandt, *Patientenzufriedenheit messen und steigern*, Sankt Augustin, 1999.

[12] C. Thielscher, *Medizinökonomie. Band 2: Unternehmerische Praxis und Methodik*, Springer-Verlag, 2012.