IT-structures and Algorithms for Quality Assurance in the Health Insurance Medical Advisory Service Institutions in Germany

Vera Ries[1], Klaus-Peter Thiele[2], Martin Schuster[3] and Reinhard Schuster[4]
[1]*Project Coordinator and Assistant to the Medical Director of the Medical Advisory Service Institution of the Statutory Health Insurance in North Rhine (MDK Nordrhein), 40212 Düsseldorf, Germany*
[2]*Medical Director of the Medical Advisory Service Institution of the Statutory Health Insurance in North Rhine (MDK Nordrhein), 40212 Düsseldorf, Germany*
[3]*Faculty of Epidemiology, Christian-Albrechts University Kiel, 24105 Kiel, Germany*
[4]*Chair of Department of Health Economics, Epidemiology and Medical Informatics, Medical Advisory Service Institution of the Statutory Health Insurance in Northern Germany (MDK Nord), 23554 Lübeck, Germany*

Keywords: Quality Assurance, Statutory Health Insurance, Medical Advisory Service Institution, Communication Structures between Different IT-systems, Server Data Structures, Data Protection, Script Programming, Client Office Answers using Perl Modules, Integer Linear Programming.

Abstract: The 15 Regional Medical Advisory Service Institutions of the Statutory Health Insurance in Germany (*MDK*) create medical expertises (*sozialmedizinische Gutachten*) on behalf of the Statutory Health Insurance Funds in the fields of inpatient and outpatient treatment, incapacity of work and other fields. The process of internal quality assurance within the local advisory service institutions as well as between them plays an increasing role and got new impulses by actual national legislation. The assessment process was established in 2004 and covers only one single indication: long-term care. It consists in a paper-based procedure focusing on a manual distribution process performed by staff of the central quality assurance bureau. We will present organizational concepts of the new standardized regional and nationwide peer review process that will cover the multitude of all existing indications of health care. It is completely digitalized using mathematical IT-based procedures not only for randomized sampling, but as well to achieve an equal distribution of the medical expertises to be assessed by the peer Medical Advisory Service Institutions. These peer reviews are supposed to be distributed among the institutions depending on occasion groups and further subtypes of medical expertises, posing a constraint satisfaction problem that needs to be solved. Therefore we discuss models that address this kind of problem type and present possible solutions for the concrete distribution problem mentioned above. We further present our technical framework that will support the workflow needed for peer review distribution, data collection and final result analysis. The 15 regional medical service institutions with different IT-system have to be connected, while data protection concerns have to be taken into account. Finally, the statistical distribution of the review results is analyzed.

1 INTRODUCTION

In Germany, 15 Regional Medical Advisory Service Institutions are mandated by the German Social Code to assist the Statutory Health Insurance (*Krankenversicherung*) as well as the Statutory Long-term Care Insurance (*Pflegeversicherung*). Overall, nearly 10.000 employees are working within the Medical Advisory Service Institutions, providing medical knowledge by case management consultancy and by medical expertise recommendation services in several fields of long-term care and of healthcare.

1.1 Preexisting Quality Assurance Plan

For long-term care there was a preexisting quality assurance plan for medical expertises. About 3.500 specialized nurses are employed by the Medical Advisory Service Institutions in those both assessment fields about long-term care. A national quality assurance plan was mandated by law since 2004 (Schmacke, 2016). It consists of the exchange of paper printed medical expertises that are distributed manually by staff of the central quality assurance bureau. For two years now, the assessment forms are no

Ries, V., Thiele, K., Schuster, M. and Schuster, R.
IT-structures and Algorithms for Quality Assurance in the Health Insurance Medical Advisory Service Institutions in Germany.
DOI: 10.5220/0008912303530360
In *Proceedings of the 13th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2020) - Volume 5: HEALTHINF*, pages 353-360
ISBN: 978-989-758-398-8; ISSN: 2184-4305

longer paper-based assessment sheets but can be filled out using a web-based data platform as well.

1.2 Starting a New Quality Assurance Plan in 2020

In 2020 a new quality assurance plan for medical expertises covering the multitude of all health care service fields provided by the Statutory Health Insurance Funds in 2020 will be started. For health care, the topics covered by the Medical Advisory Service Institutions' consultancy and expertise services are diverse. About 2.500 specialized physicians, 90 other health care professionals and 400 medical coders are working in the health care division of the Medical Advisory Service Institutions. The main tops are

- hospital quality and billing control on behalf of the health care insurance funds (Thiele K.-P., 2018),
- out-patient case management consultancy and medical expert opinion service (Nüchtern, 2008) in multiple fields of medical treatment.
- To learn more about the German social system of health care in detail, see (Altenstetter and Busse, 2005), (Cortina, 1993), (Machnik, 2009), (Sawicki and Bastian, 2008).

So, there are many different health care topics covered by the Medical Advisory Service Institutions' consultancy and medical expertises. For a long time, however, there was no monolithic quality assurance plan established for the health care segment of consultancy and medical expertises. The Regional Medical Service Institutions performed a kaleidoscope of singular quality assurance measures lacking a nationwide perspective and any public reporting, instead. In 2017, a nationwide task force was established by the medical directors' conference of all regional Medical Advisory Service Institutions to design, to initiate and to implement a total quality assurance plan.

The preexisting regional quality assurance initiatives were used as a draft to establish a unified continuous quality circle (Deming, 1982), (Shewhart, 1931) suitable to the diversity of medical topics, performing annual quality measurements and validations. The total quality assurance organizational framework established addresses both, the internal quality assurance of every Regional Medical Advisory Service Institution as well as a nationwide quality assurance. The quality assurance method is based on peer reviews, a common method in statutory social insurance frameworks in Germany (DRV, 2018), (Polak, 2018), (Strahl et al., 2016), (Strahl et al., 2018). By facilitating mutual learning, thereby enhancing continuous quality improvement, both regionally and nationwide. Crucial to this nationwide initiative was a commitment by all Regional Medical Advisory Service Institutions of a joint and uniform conception of quality. A prime nationwide reporting of the quality assurance results of all Medical Advisory Service Institutions starting in 2020 will contribute to a high degree of transparency between each other and fosters mutual learning. What is more, it inspires reliability and trustworthiness towards their clients, the Statutory Health Insurance Funds, and towards society.

All nationally and internationally available information on quality assurance in general and especially in health systems and health and pension insurance have been adapted to the newly established system with the conceptional time ending 2019 and the role out starting in 2020.

2 MATERIAL AND METHODS, PROCESS ORGANIZATION

The Medical Advisory Service Institutions yearly produce about 2.5 million medical expertises about inpatient hospital treatments and documentations for on behalf of the statutory health insurance funds. A random sample of 0.5% will be used for regional quality assurance using peer reviews. 10 % of those medical expertises randomly chosen get a second external peer review by another advisory service institution generating the nationwide perspective. Thereby we get 12,500 peer reviews for regional quality assurance and 1,250 nationwide peer reviews.

Additionally there are about 650 thousand medical expertises in other fields classifies in eight occasion groups (inability of work, outpatient care, new and unconventional treatment methods and drug treatment, prevention and rehabilitation, remedies, dentistry, factual or putative medical treatment, claims against third, other) with further subgroups. In analogy to industrial quality standards for production, quality is defined as a measure that describes the degree of correspondence between the service provided by an deliverer and the service expected by the customer (Masaaki, 1986), (Gerlach, 2001), (Kamiske and Brauer, 2011), (Institute of Medicine; Committee to Design a Strategy for Quality Review and Assurance in Medicare, 1990), (Institute of Medicine, 2001), (Internationale Organisation für Normung, 2015). The concept of nationwide and regional quality assessment is worked out and will be further developed by the permanent commission of senior physicians and quality managers. The questionnaire

was developed by a task force based on an information retrieval about preexisting regional quality assurance concepts for the common set of quality criteria followed by a modified Delphi method to shape the different sets of indication-specific quality criteria (Normand et al., 1998). Decisions are made by the conference of medical directors conference (*Leitende Ärzte*) and the conference of the managing directors (*Geschäftsführer*).

In Germany the regions differ considerably in size and population density. So, questions of equity are a strong point in any discussion. Therefore the reviews start with a minimal size setting feasible for any participant.

The peer review process is organized quarterly. If at least one medical expertise is transferred to every other local advisory service institution, 14 expertises have to be included per quarter or 56 per year. If we demand, that at least the double size have to be included for local quality assurance, we come to ask for 112 expertises in each field. This leads to 11,660 medical expertises in total on the regional level and 5,300 medical expertises on the nationwide level in peer review evaluations for all smaller occasion fields. This number of medical expertises leads to a personal and financial effort which should not be exceeded. For reasons of practicability, we reduce the number of medical expertises given to an external review process for the smaller five medical service institutions to a half (7 instead of 14 medical expertises for the 15 existing regional medical service institutions).

The internal number of peer reviews for hospital expertises nearly equals the total number of all other occasion groups, thereby is no need to enlarge the number of hospital peer reviews.

The peer reviews for medical expertises are based on pdf files and related meta data. Additional medical treatment information is not used by reasons of practicability (for instance large hospital treatment records). The use of additional medical data in regional quality assurance would create a bias towards the nationwide quality assurance because the additional data like clinical case logs are unable to be transfered for data protection and practicability reasons. What is more, the medical expertises aims to be perspicuous and understood without additional medical documentation data available.

Meta data contain the type (occasion) and date of release of medical expertise, subregion of the local advisory service institution and possibly additional process parameters. Pdf files and meta data are related by a regional unique identity key.

With respect to m=9 occasion types (starting with m=2) and further subtypes of medical expertises the pear review process yields an n-dimensional (n around 30 in different groups) quality vector with components of k outcome types (actually used: ”correct”, ”improvement potential with no result relevance”, ”incorrect” and ”not relevant”), cf. Figure 2. Only a minority of questions depend on the occasion type. The questions for the peer reviews have been tested with positive results in a reliability analysis which will be repeated periodically, cf. (Bühn et al., 2017), (Wirtz and Caspar, 2002), (Eisinga et al., 2012).

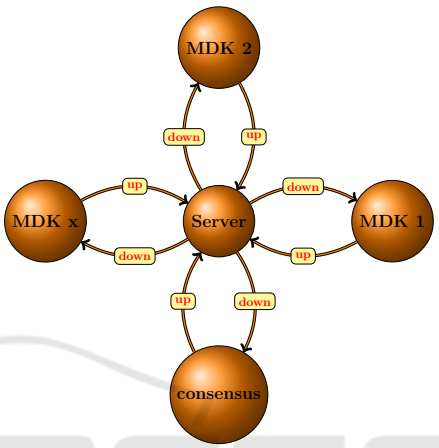

Figure 1: Communication between Medical Advisory Service Institutions (*MDK*) and the QA-server.

The central IT-system is a linux sever (Debian) which is located in an external data center which we will call QA-server (**q**uality **a**ssurance). It uses MySQL as database system, perl as the main script language including the perl module writeexcel in order to generate excel files to give overviews about data and aggregate informations for medical review service institutions. The communication uses a role concept for administrative persons, medical experts, peer reviewers and members of consensus conferences. The server management uses the ssh protocol with key pairs and a passphrase. The communication between the server and the regional advisory service institution use the https protocol. Every user needs a personal password. All transfer actions are logged. Data backups are made daily.

Every communication step is a pair of upload and download exchange streams between advisory service institutions with individual IT-structure and the QA-server or between other participants as the consensus conference and the server. It is a star topology with the QA-server in the center, cf. Figure 1.

In order to start the quarterly local quality assurance process, a matrix which contains all local medical expertise identities (pdf file names) with the men-

tioned meta data concerning a period of a month or quarter has to be uploaded. The respective download answer gives the random sample for regional quality assurance. A second random sample for external quality assurance is determined, but not transmitted to the local advisory service institution. As a next step the pdf files related to the sample have to be uploaded to the QA-sever as single pdf files or as zipped files in individual parts or totally. The server answer consists in a first file structure analysis. The regional identity is converted into an external unique file identity, the transformation is stored in a MySQL table.

As a next step one or several peer review managers distribute the medical expertise files e_i with meta data $m(e_i)$ to peer review experts r_j as a function $f : e_i \rightarrow r_j$. This may be done using the detailed occasion reasons or regional features due to local configurations. It also can be done completely automatic. The server provides a radio buttom formula and the answer arranges the combination of review experts and review documents.

In order to carry out the review, the review expert first has to download the pdf file of the related medical expertise. After that there is the possibility to use a QA-server form online or to use a pdf form for offline review. In both cases the server will give the result of a formal answer check (all questions are answered and for all incorrect points reasons are stated).

The result of the peer review is transfered to the author of the medical expertise for feedback. The medical expert gives notice if she/he agrees or disagrees with the peer review. This happens primarily anonymously. However, the medical expert may ask for a personal contact to the peer reviewer. Further quality management actions are up to regional advisory service institution decisions.

Alternatively, the regional advisory service institution may carry out the regional peer review procedure in his own IT structure. In this case, the service institution quality manager has to upload the results in csv or xls-format again with a related answer check procedure.

	review question								
review pdf	#01	#02	#03	#04	#05	#06	#07	#08	#09
1	3	3	1	1	3	3	2	2	3
2	1	1	1	1	2	1	1	1	1
3	1	1	1	1	3	1	3	1	1
4	1	1	1	1	3	3	3	1	1
5	1	1	1	1	1	1	1	1	2
6	1	1	3	2	3	3	3	1	1
7	1	1	1	2	3	1	3	2	1
8	1	1	1	1	2	3	2	3	1
9	1	2	3	1	3	2	2	2	1
10	1	1	1	1	3	3	3	1	1
11	1	1	1	1	3	1	1	1	1
12	1	1	2	2	2	1	3	1	1

Figure 2: Upload regional review results with meaning green: correct, yellow: improvement potential with no result relevance, red: incorrect.

Using the upload, a random sample is chosen for the nationwide peer review. The related identities are given as the server answer in order to select the related pdf file for an upload. All identities are transformed to QA-server identities. One regional extension of the quality assurance is consisting in the demand, that every physician who is included in the creation of medical expertises shall be included in the peer review process at least two times a year in each relevant occasion group. If this is not reached by the random process, additional expertises have to be included in the regional peer review process. If this enlargement is done in the last quarter, the number of added expertises remain as low as possible, but it results in an unequal workload. The expertises of the enlargement shall not be included in the random selection for external peer reviews, because of the fact that it would increase the possibility of physicians with a small number of expertises to be in the external peer review process.

The distribution of the regional medical expertises to external medical service institutions shall satisfy the following conditions:

- Every advisory service institution shall do as much peer reviews as it has given medical expertises to the nationwide quality assurance using the case numbers stated above
- Every other advisory service institution shall be included in order to get large range of opinions
- Boards with more medical expertises shall get more peer reviews (with possibly ”smaller” violations of this rule)
- The service institutions with the largest number of medical expertises shall get as few peer reviews as possible in order to avoid concentration of special opinions in one spot

We will discuss the ILP based method to solve this problem later on. Only for the hospital treatment the distribution differs essentially from an equal distribution.

Before the medical expertises are forwarded to a the external peer reviewers is realized, a further random process is performed. The distribution of the medical expertises to the peer reviewers is done by an analogous procedure described for internal peer reviewers. Every peer reviewer may propose medical expertises for a discussion in the consensus conference. For each occasion group there shall be two consensus conferences each year. The external peer reviewer does not know the results of the internal peer review. If the regional valuation differs from the nationwide valuation, the internal reviewer is asked, if he accepts the result of the external peer review. If the

difference persists, the expertise has to be discussed in the consensus conference. In order to organize this comparison, every internal peer reviewer has to consider differing results to his own peer review results after peer reviews are completed. The comparison is generated automatically by the QA-server.

Every participant of the consensus conference will be provided with detailed information about all expertises proposed to the consensus conference of the respective occasion group. Every advisory service institution sends out a member into this conference. The conference members only get labels of the peer reviews done by members of there home advisory service institution in order to remain information as anonymous as possible at the one hand and to get additional information about reasons of review decisions as deep as possible at the other hand. If the consensus conference does not reach an agreement with respect to a special expertise, the simple majority will decide about the final result. The responsible manager of the consensus conference has to upload the final decisions to the QA-server, this can be done during the conference.

The amount of uploaded data in relation to the stated case numbers and the fraction of realized peer reviews may be requested by the quality managers with the related roles (regional or nationwide level). Additionally frequency results of all parts of the review process can be downloaded at every time. The results are given as files in xls or xlsx format with choosable evaluation contents. Every advisory service institution can download all details about the expertises of his home service institution including the external peer review results, but there are no labels which external service institution gave the nationwide evaluation. The overall amount of analysis is done via background processing.

Aggregated results of the regional and the nationwide peer review process will be published. In order to get information about the security of the results confidence levels are calculated. Tables and figures of the results in the decided amount are produced by the QA-server.

3 RESULTS, MATHEMATICAL MODELS

A first pretest in the field of incapacity of work and a second pretest outpatient care with the subgroups DRG (diagnoses related groups) and PEPP (pay groups in psychiatry and psychosomatic medicine) has been used to reduce the variance and to increase the kappa values between the reviewers with respect to the questions and with respect to the medical opinions. After improvement and clarifying of the questionnaire we got an intraclass correlation coefficients (average measures) of 0.923, 0.903 and 0.868 in the mentioned fields and thereby a good reliability was reached, cf. (Cortina, 1993), (Eisinga et al., 2012), (Wirtz and Caspar, 2002). Let $w(i,k)$ be the result with value 1 (resp. 2 and 3) for correct (resp. improvement potential with no result relevance and incorrect) medical opinions of the expert from the evaluation of the reviewer for medical opinion i and question k. Then

$$d(x,y) = \sum_i |w(i,x) - w(i,y)|$$

is the Manhattan distance of the opinion vector with respect to the questions x and y. For each question we look for the other questions with the smallest two distances. We use the questions as the vertices of a graph and the edges we get by the shortest distances. Using the modularity method we get graph communities. For the calculations we have used Mathematica by Wolfram Research. An example for the pretest is shown in Figure 3. The results after a year of quality assurance will give further hints for improvement of questions and further training opportunities for the medical experts.

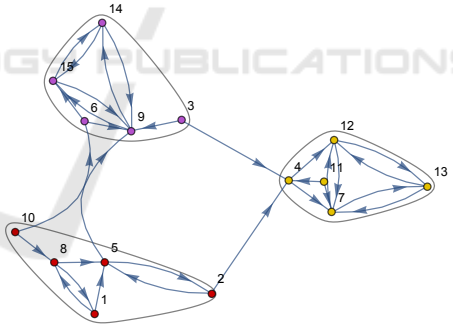

Figure 3: Graph Communities of questions Selection.

Let $n(i,j)$ be the number of medical expertises which are transfered from medical service institution i the medical expert to the medical service institution j of the reviewer and $w(i)$ the number of expertises from service institutions i given in the external review process. The order of medical service institutions are chosen such that $w(i) \geq w(j)$ for $i > j$. No expertise remains in the same medical service institutions, so the the matrix $n(i,j)$ has an empty diagonal. If one would use the proportional approach

$$n(i,j) = int(w(i)\frac{w(j)}{\sum_{k \neq j} w(k)})$$

one would not get a balance between ingoing and outgoing expertises (rounding problems are only a minor effect). One could implement a random process with certain rules, but it would result in under- and over-representation of certain medical service institutions in most quarters.

We will look for solutions of equations and inequalities with integer solutions with a certain optimal solution. As a balance demand and the demand the no expertise stay in the same medical service institutions we state

$$\begin{aligned}\sum_j n(i,j) &= w(i)\\ \sum_i n(i,j) &= w(j)\\ \sum_i n(i,i) &= 0.\end{aligned}$$

If we require the symmetry of the matrix $n(i,j)$ it will reduce the number of variables which may reduce the machine time, but it may lead to insolvability of the problem. Let $m(i,j)$ the matrix we get from $n(i,j$ by deleting the main diagonal (delete in columns). We could demand monotony in the columns and rows without the main diagonal:

$$\begin{aligned}m(i,j) &\geq m(k,j) \text{ for } i > j\\ m(i,j) &\geq m(i,k) \text{ for } i > j\end{aligned}$$

Already with $w(l) < l$ (l number of medical service institutions) the problem is unsolvable. So we will take a weaker demand:

$$\begin{aligned}m(i,j)+d &\geq m(k,j) \text{ for } i > j\\ m(i,j)+d &\geq m(i,k) \text{ for } i > j\end{aligned}$$

with d as small as possible for solvability.

The reviews should concentrate as little as possible to large service institutions and be as uniformly as possible between the medical service institutions. Together with the weak monotony demand we will use

$$n(1,2)+n(2,1) \longrightarrow Min!$$

as an optimization condition.

In the case $w(i) = 14$ for $1 \leq i \leq 10$ (large medical service institutions) and $w(i) = 7$ for $11 \leq i \leq 15$ (small medical service institutions) as an example for two equal groups which will be applied for all occasion groups except of hospital care a balance between and within the groups can be reached by additional constraints as

$$\begin{aligned}\sum_{i=11}^{15} n(i,j) &= 2 \text{ for } j = 11,...,15\\ \sum_{j=11}^{15} n(i,j) &= 2 \text{ for } i = 11,...,15.\end{aligned}$$

Thereby we have got an integer linear program (ILP). This can by solved (if there is a solution at all) e.g. a Python program and using CPLEX bindings. We also could first solve a linear program without the restriction to integers and then do a post processing step to obtain integer solutions by local changes. This can be done with Mathematica by Wolfram Research or also using Python and CPLEX, cf. (Shinano et al., 2003). For finding a suitable distribution matrix for distributing the reviews over the medical advisory we use an integer linear programming approach. Integer linear programs (ILP) allow to formulate an optimization problem with an linear objective function and constraints given as a series of linear inequalities. Then, an integer valued assignment of the variables is wanted that fulfills all constrains. Although solving ILP problems is computationally hard (NP-hard to be more specific), there are powerful ILP solvers available that perform well on our concrete problem formulations, see (Newman, 2006).

A solution with the smallest value $d = 1$ is shown in Figure 4. In each quarter one should apply a random permutation within the two groups and/or use an other solution of the ILP system. One solution of the ILP is shown in Figure 5.

	quarterly permutation										quarterly permutation					
	BY	NR	BW	HE	WL	NI	NO	RP	BB	SN	ST	TH	MV	SL	HB	sum
BY		1	2	1	1	1	2	1	1	2			1	1		14
NR	1		2	2	1	1	1	1	1	2				1	1	14
BW	1	2		2	1	1	1	1	1	1	1			1	1	14
HE	1	2	1		2	1	1	2	1	1	1				1	14
WL	1	1	1	2		2	1	1	1	1	1	1			1	14
NI	1	1	1	1	2		2	1	1	1	1	1			1	14
NO	2	1	1	1	1	2		1	1	1	1	1	1			14
RP	2	2	1	1	1	1	1		2	1		1	1			14
BB	1	1	1	1	1	1	1	2		2		1	1	1		14
SN	2	1	1	1	1	1	1	2	2				1	1		14
ST			1	1	1	1	1				*1*			*1*	*1*	7
TH					1	1	1	1	1		*1*				*1*	7
MV	1						1	1	1	1	*1*	*1*				7
SL	1	1	1						1	1		*1*	*1*			7
HB		1	1	1	1	1							*1*	*1*		7
sum	14	14	14	14	14	14	14	14	14	14	7	7	7	7	7	

Figure 4: Distribution matrix of medical expertises to external peer reviews for occasion group "incapacity of work" of medical service institutions.

The result of the review (i.e. probability of improvement potential) may be considered as binomial distribution and one can calculate confidence intervals for yearly reports. Such confidence intervals are much more suitable values which contain much noise. The review results may depend on further parameters which can be used to optimize the quality assurance process. An example of the comparison of the results of the considered medical service institutions is shown in Figure 6, real results can be given after for quarters of quality assurance, which will be started in 2020.

quarterly permutation

	BY	NR	BW	HE	WL	NI	NO	RP	BB	SN	ST	TH	MV	SL	HB	sum
BY		7	6	5	6	5	5	5	5	5			1	1		51
NR	6		6	6	4	5	5	1	1	2			1	1	1	39
BW	6	6		6	5	4	2	2	1	2	1			1	1	37
HE	5	5	6		4	2	1	1	2	1	1			1		29
WL	6	4	5	3		3	1	1	1	1	1	1			1	28
NI	6	4	3	2	2		2	1	1	1	1	1			1	25
NO	5	5	3	1	1	1		2	1	1	1	1	1			23
RP	5	2	2	1	1	1	2		2	1		1				18
BB	5	1	1	1	1	1	1	2		1		1	1	1		17
SN	5	2	2	1	1	1	1	1	1				1		1	17
ST			1	1	1	1	1				*1*			*1*	*1*	7
TH					1	1	1	1		1	*1*				*1*	7
MV	1	1					1	1	1		*1*	*1*				7
SL	1	1	1	1					1			*1*	*1*			7
HB		1	1	1	1					1			*1*	*1*		7
sum	51	39	37	29	28	25	23	18	17	17	7	7	7	7	7	

Figure 5: Distribution matrix of medical expertises to external peer reviews for for occasion group ”hospital care” among medical service institutions.

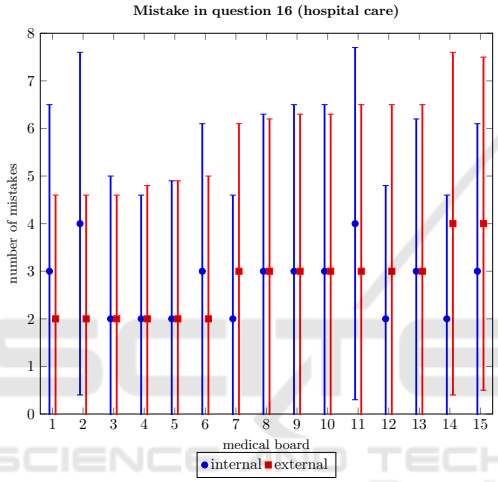

Figure 6: Results and confidence intervals for external and nationwide failure medical expertises.

4 CONCLUSIONS

The quality assurance process in the sense of the considered PDCA cycle will be especially successful if acceptance, practicability and transparency support each other. In each Medical Advisory Service Institution has to reach each medical expert in order to reach a continuous development of a common understanding ans progression of quality. The external peer reviews shall reach a common understanding of quality in each region with a transparent presentation of the results to the public. In Germany there is a considerable heterogeneity in size and number of employees among the medical service institutions involved. Questions of equity are a constant concern. The discussed methods provide the necessary framework that enables all medical advisory service institutions to take part and to shape trustful cooperation among each other.

The distribution of the regional expertises to nationwide reviews are subject to distributional fairness. All regions have to be appropriately included on the one hand and larger resources of larger services on the other hand must be taken into account. There shall be no avoidable regional concentrations of peer reviews as well as no no overemphasis on certain coupling patterns. It is realized as a solution of an integer linear program which shall not only change slightly from quarter to quarter. An important point is the reliability and validity of the valuation results. It is essential to communicate confidence intervals for the results instead of single values so that the public is in a position to distinguish between random and secured differences. In the practical implementation the interaction of different IT systems used by the medical experts in terms of good manageability and great security is a big challenge.

The implementation uses a Linux based MySQL server which is reached by the regional medical advisory institutions as well as for coordinating working groups by the https protocol. In the working process are used predominantly perl scripts. Overviews are generated on the server as excel tables. There are flexible regional design options through configuration tables. The described concept should enable a fast nationwide development of a common understanding and improvement of quality. Quality conferences in each occasion group twice a year will discuss different views and will lead to a final decision process which has to be communicated in each region by the quality managers.

REFERENCES

Altenstetter, C. and Busse, R. (2005). Health Care Reform in Germany: Patchwork Change Within Established Governance Structures. *Journal of Health Politics, Policy and Law*, 30(1-2):121–42.

Bühn, S., Mathes, T., Prengel, P., Wegewitz, U., Ostermann, T., Robens, S., and Pieper, D. (2017). The risk of bias in systematic reviews (ROBIS) tool showed fair reliability and good construct validity to assess the risk of bias in systematic reviews. *Journal of Clinical Epidemiology*, pages 121–128.

Cortina, J. (1993). What Is Coefficient Alpha? An Examination of Theory and Applications. *Journal of Applied Psychology*, 78 (1):98–104.

Deming, W. E. (1982). *Out of the crisis.* Cambridge: Massachusetts Institute of Technology.

DRV (2018). *Qualitätssicherung der sozialmedizinischen Begutachtung. Manual zum Peer Review-Verfahren*. Berlin: Deutsche Rentenversicherung Bund. Überarbeitete Fassung August 2018.

Eisinga, R., Grotenhuis, M., and Pelzer, B. (2012). The reliability of a two-item scale: Pearson, Cronbach or Spearman-Brown? *International Journal of Public Health*, 58 (4):637–642.

Gerlach, F. M. (2001). *Qualitätsförderung in Praxis und Klinik: eine Chance für die Medizin.* Stuttgart: Thieme.

Institute of Medicine (2001). *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC, National Academy Press.

Institute of Medicine; Committee to Design a Strategy for Quality Review and Assurance in Medicare (1990). *Medicare: A Strategy for Quality Assurance. Volume II: Sources and Methods.*, volume 2. Washington (DC): National Academies Press (US).

Internationale Organisation für Normung (2015). *DIN EN ISO 9001:2015-11. Qualitätsmanagementsysteme - Anforderungen*. Beuth Verlag.

Kamiske, G. F. and Brauer, J.-P. (2011). *Qualitätsmanagement von A bis Z: Wichtige Begriffe des Qualitätsmanagements und ihre Bedeutung*. Carl Hanser Verlag.

Machnik, W. (2009). *20 Jahre MDK Nordrhein. Jubiläumsschrift*. MDK Nordrhein.

Masaaki, I. (1986). *Kaizen: The Key To Japan's Competitive Success*. McGraw-Hill Education.

Newman, M. (2006). Finding community structure in networks using the eigenvectors of matrices. *Physical Review E*, 74.

Normand, S.-L., Mcneil, B., Peterson, L., and Palmer, R. (1998). Eliciting expert opinion using the Delphi technique: identifying performance indicators for cardiovascular disease. *International Journal for Quality in Health Care*, 10:247 – 260.

Nüchtern, E. (2008). Das professionelle Gutachten. Besonderheiten in der gesetzlichen Krankenversicherung. *MedSach*, 105 (3):96–98.

Polak, U. W. (2018). Evaluation von Durchgangsarztberichten mithilfe eines Peer-Review-Verfahrens. *Trauma Berufskrankh*, 20 (Suppl. 4):237–240.

Sawicki, P. and Bastian, H. (2008). German Health Care: A Bit of Bismarck Plus More Science. *BMJ*, 337:a1997.

Schmacke, N. (2016). Zur Positionierung des MDK in der gesetzlichen Kranken- und Pflegeversicherung, Gutachten im Auftrag des AOK-Bundesverbandes. In *IPP-Schriften, Eds. Bolte, G, Görres, S. and Gerhardus, A.*, page Bd. 14. Institut für Public Health und Pflegeforschung (IPP) der Universität Bremen.

Shewhart, W. A. (1931). *Statistical Method from the viewpoint of quality control. (Reprint im Jahr 1986 einer Veröffentlichung des Department of Agricultuire Milwaukee 1931).* New York: Dover Publication.

Shinano, Y., Fujie, T., and Kounoike, Y. (2003). Effectiveness of Parallelizing the ILOG-CPLEX Mixed Integer Optimizer in the PUBB2 Framework. In *Euro-Par 2003 Parallel Processing*, pages 451–460. Springer Berlin Heidelberg.

Strahl, A., Gerlich, C., Alpers, G., Ehrmann, K., Gehrke, J., Müller-Garnn, A., and Vogel, H. (2018). Development and evaluation of a standardized peer-training in the context of peer review for quality assurance in work capacity evaluation. *BMC Medical Education*, 18:135–145.

Strahl, A., Gerlich, C., Wolf, H.-D., Gehrke, J., Müller-Garnn, A., and Vogel, H. (2016). Qualitätssicherung in der Sozialmedizinischen Begutachtung durch Peer Review - ein Pilotprojekt der Deutschen Rentenversicherung. *Gesundheitswesen*, 78:156–160.

Thiele K.-P., Kreuzer C., M. R. (2018). *Krankenhauslandschaft in Deutschland: Zukunftsperspektiven, Entwicklungstendenzen, Handlungsstrategien*, chapter Exkurs 3: MDK-Prüfung - Fluch oder Segen?, pages 89–94. Stuttgart: Kohlhammer.

Wirtz, M. and Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität: Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen*. Göttingen: Hogrefe.