

ORIGINAL ARTICLES

Development and evaluation of a three-dimensional multi-level model for visualising the workflow composite score in a health IT benchmark

Johannes Thye, Ursula Hübner, Matthias-Christian Straede, Jan-David Liebe

Health Informatics Research Group, University of Applied Sciences Osnabrück, Department of Business Management and Social Sciences, Osnabrück, Germany.

Correspondence: Ursula Hübner. Address: Health Informatics Research Group, University of Applied Sciences Osnabrück, Department of Business Management and Social Sciences, Caprivistr, 30A, D-49076 Osnabrück, Germany. Email: u.huebner@hs-osnabrueck.de

Received: January 8, 2016
DOI: 10.5430/jbei.v2n2p83

Accepted: January 29, 2016
URL: <http://dx.doi.org/10.5430/jbei.v2n2p83>

Online Published: February 18, 2016

Abstract

Background: Clinical information logistics is the backbone of care workflows inside and outside of hospitals. Due to the great potential of health IT to support clinical processes its contribution needs to be regularly monitored and governed. IT benchmarks are a well-known instrument to optimise the availability and use of IT by guiding the decision making process. The aim of this study was to translate IT benchmarking results that were grounded on a hierarchical workflow scoring system into an appropriate visualisation concept.

Methods: To this end, a three-dimensional multi-level model was developed, which allowed the decomposition of the highly aggregated workflow composite score into score views for the individual clinical workflows concerned and for the descriptors of these workflows. Furthermore this multi-level model helped to break down the score views into single and multiple indicator views.

Results: The results could be visualised per hospital in comparison to the results of organisations of similar size and ownership (peer reference groups) and in comparison to different types of innovation adopters. The multi-level model was implemented in a benchmark of 199 hospitals and evaluated by the chief information officers. The evaluation resulted in high ratings for the comprehensibility of the different types of views of the scores and indicators.

Conclusions: The implementation of the multi-level model in a large benchmark of hospitals proved to be feasible and useful in terms of the overall structure and the different indicator views. There seems to be a preference for less complex and familiar views.

Key words

IT benchmarking, Innovation benchmarking, Information visualisation, Data visualisation, Hierarchical visualisation, Multi-level model, Indicator views, Workflow composite score

1 Introduction

Complex and advanced health care processes, which require good communication among the health care professionals, on time exchange of the relevant patient data and their thorough analyses, cannot be performed without appropriate health IT systems^[1-5]. The construct information logistics helps to better understand the nature of information flows between the

different actors and hereby to ensure information continuity^[6-9]. However, health information technology (HIT), *i.e.* the carrier of the information and enabler of information flows, is often regarded as a black box by executives. This happens because the mechanisms behind HIT systems frequently lack transparency and systematic approaches. In addition, valid models to regularly measure and evaluate IT performance and IT services are missing as well^[5, 10]. Only when strategic methods for planning, monitoring and governing are implemented is IT optimisation possible^[5, 11]. Benchmarking of HIT systems is regarded as a suitable method for managing IT in health care organisations on a strategic level^[6, 12]. Benchmarking procedures can be classified into two groups: statistical benchmarks with a large number of participants and in-depth benchmarks in smaller and often closed groups of participants^[13]. Benchmarking allows comparing equivalent structures, processes and methods within a single enterprise or a group of enterprises and according to the pure doctrine between an enterprise and the market leader in this or another industry^[14-16]. In order for benchmarks to grasp the essence of HIT supported processes, the construct clinical information logistics has been proposed and targeted by benchmarks^[17]. Clinical information logistics was measured by the workflow composite score (WCS)^[18], which represented the most aggregated level within a hierarchical scoring system of single IT features (raw indicators), sub scores and finally the WCS. Due to its novelty, the WCS system has not been presented in an appropriate graphical form.

Besides measuring IT structure and IT process support, benchmarking can be employed to compare enterprises on the basis of external criterion ability and power to be innovative^[19-21]. HIT can appear as an innovative product itself but can also enable process innovations, *i.e.* leveraging new processes or increasing the efficacy and efficiency of workflows^[5]. Benchmarking organisations in terms of HIT innovations can thus investigate IT structures as well as IT support of clinical processes^[22] and can classify the organisation according to well established groups of innovation adopters^[23]. The innovation perspective broadens the scope towards developing the organisation by realising new opportunities through IT^[5]. Still, classic benchmarks rather focus on costs, efficiency and efficacy less on innovation.

Due to its high potential, HIT benchmarks have become an area of interest in the recent years^[24-28]. Despite this high interest, there is no explicit study – neither within nor outside healthcare – that proposes and evaluates a coherently structured visualisation system for indicators and scores. Therefore, the challenge was to implement the presentation of the benchmark results in a logical and understandable structure and thereby reflect a new highly aggregated indicator of clinical information logistics. Such system would help to meet the big challenges of communicating the benchmark results to members of the executive level. According to our understanding these challenges are:

- To select the most important and significant results flexibly from a wealth of results
- To present these results in a manner that is understandable for non IT experts
- To present the important results so that their messages can be captured within a short time span

Against this background, this study aims at developing and evaluating a visualisation system for benchmark results that could be used as an information base for discussions between chief information officers and (other) members of the executive level, *e.g.* chief executive officers, medical and/or nursing directors. This visualisation system should allow the analysis of large data sets and the processing of these data so that they become accessible, comprehensible and usable for the users of these data^[29, 30]. In this sense, it also should visually establish links between the different scores and raw indicators, *i.e.* the IT features, among each other and towards the overall goal to represent clinical information logistics. The following research questions guided this study:

- 1) Which visualisation system allows the users to obtain an overview of the IT benchmarking results and at the same time provides enough details?
- 2) How can the scores and the raw indicators, *i.e.* the IT features, be represented in a graphical form so that individual results can be compared with the results of the peers in the reference groups?
- 3) How can measures of the innovative strength of an organisation be displayed graphically?
- 4) How do benchmark participants evaluate the visualisation of the raw indicators, *i.e.* the IT features, and the different scores?

2 Methods

2.1 Data collection and aggregation

The visualisation concept to be developed should be implemented in a benchmarking environment with real data. The data basis utilised was provided by a survey, which is conducted regularly in German hospitals, the IT Report Healthcare (www.it-report-healthcare.info) and which was used in the 2014 version [31]. The data basis consisted of 92 IT features [18], the so-called raw indicators, which were grouped in the categories IT functions, IT structure and IT support of clinical processes. These 92 IT features were allocated to four sample benchmark processes, *i.e.* ward rounds, pre-surgery processes, post-surgery processes and discharge and to the four descriptors data and information, IT functions, integration and distribution. Scores were computed on the level of the four processes and on the level of the four descriptors. The sum of these scores yielded the WCS. All data were statistically analysed using IBM SPSS V21 and Microsoft Excel. Thus a total of 92 features (raw data) per hospital had to be structured and aggregated before they could be visualised. We largely followed the process of data aggregation proposed by Card [32] (see Figure 1). To this end, the raw data were first internally organised in tables, which contained individual values and corresponding statistical parameters of the reference group. Then an overall visual access to the benchmark results, the visualisation system was devised from which particular views, the diagrams were deduced.

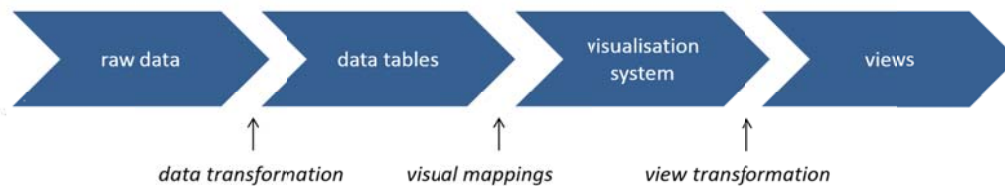


Figure 1. Data aggregation and transformation (following 32)

2.2 Reference groups

A total of 259 hospitals represented by their chief information officer took part in the survey [31]. A workflow composite score could be computed for a group of 183 hospitals, which provided enough details for the computation. Out of the 259 hospitals 199 participated in the IT benchmark and of which 156 hospitals were benchmarked based on the workflow composite score, which could be calculated for them. The other 43 hospitals out of the 199 were benchmarked on the basis of the raw indicators. In order to allow the comparisons, organisations of similar size and ownership were grouped [28] and statistical parameters within these groups were calculated. They formed the so-called peer reference groups. Table 1 shows the characteristics and numbers of the respective groups.

Table 1. Participants’ structure of the IT benchmarking

Reference groups		Number of benchmark participants (WCS)	Size of peer reference group (WCS)	Number of all benchmark participants (raw indicators)	Maximum size of peer reference group (raw indicators)
Ownership	public	130	150	161	202
	private	26	33	38	57
Size	< 200 beds	33	38	59	81
	200 – 399 beds	48	60	55	77
	400 – 599 beds	33	39	36	43
	600 – 799 beds	16	16	20	22
	≥ 800 beds	26	30	29	36

In addition to these peer reference groups, innovation reference groups were formed according to the adoption categories of Rogers [23] and the organisations were classified as innovators, early adopters, early majority, late majority and laggards (see Figure 2) with regard to the WCS and individual raw indicators (IT features).

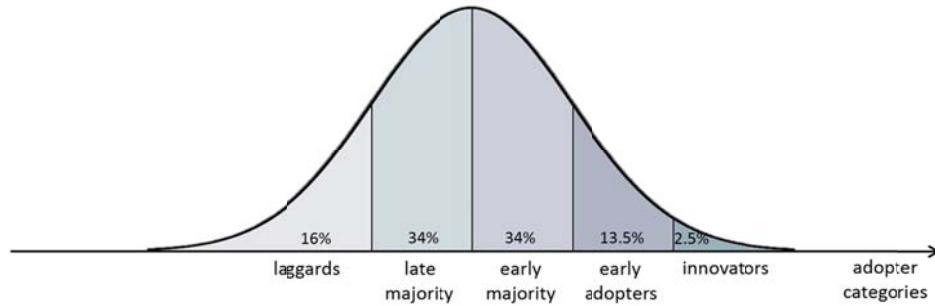


Figure 2. Adoption categories according to the Diffusion of Innovations model ^[23]

2.3 Information to be visualised

Based on the hierarchical scoring model. All scores and related IT features (raw indicators) composing the WCS had to be presented in an appropriate form. Table 2 gives an overview of the indicators to be visualised.

Table 2. Indicators to be visualised

Type of indicator	Content of indicator	Name of indicator
Composite score	composite score	workflow composite score
Sub scores	workflow scores	workflow score <i>ward round</i>
		workflow score <i>pre-surgery</i>
Sub scores	workflow descriptor scores	workflow score <i>post-surgery</i>
		workflow score <i>discharge</i>
Sub scores	IT structures	workflow descriptor score <i>data and information</i>
		workflow descriptor score <i>functions</i>
Sub scores	IT functions	workflow descriptor score <i>integration</i>
		workflow descriptor score <i>distribution</i>
Raw indicator/ IT features	IT processes	architecture of health information system
		WiFi implementation
Raw indicator/ IT features	IT processes	workflow management system
		electronic devices for processing patient data
Raw indicator/ IT features	IT processes	electronic health record system
		clinical documentation
Raw indicator/ IT features	IT processes	order entry and observation reporting
		clinical decision support
Raw indicator/ IT features	IT processes	patient safety
		supply chain functions
Raw indicator/ IT features	IT processes	interface functions
		ward rounds
Raw indicator/ IT features	IT processes	<ul style="list-style-type: none"> list of patient data available on desktop PC or on mobile device access to patient data via desktop PC or mobile device
		pre-surgery processes
Raw indicator/ IT features	IT processes	<ul style="list-style-type: none"> surgery planning list of patient data available prior to surgery
		post-surgery processes
Raw indicator/ IT features	IT processes	<ul style="list-style-type: none"> list of patient data available after surgery on normal ward or on intensive care unit format of patient data to be transmitted to normal ward or intensive care unit
		discharge
Raw indicator/ IT features	IT processes	<ul style="list-style-type: none"> reminders for pending activities before discharge electronic systems for clinical pathways and medical guidelines list of patient data available for medical summaries electronic medical summaries electronic nursing summaries

2.4 Development of the visualisation system

In order to answer the first research question, an overall framework had to be developed that could encompass the entire hierarchical scoring model and make its structure transparent and understandable and at the same time allow individual comparisons within the reference group and with the best of the group. The challenge for constructing the visualisation system was less to find the best types of diagrams but rather to build a framework within which the users could navigate in a large dataset and could interact with the data^[33]. Since the scoring system was hierarchical the visualisation system had to be constructed accordingly and had to follow a tree-structure^[30, 34]. Hereby, the hierarchy defined the sequence of the presentations and the categories (clinical workflows and workflow descriptors) determined the content at each hierarchical level^[34]. This drill-down procedure should allow the user to navigate within the dataset along a predefined scheme without getting drowned in a sea of information^[35].

2.5 Principles of data visualisation applied in this study

The development of the visualisation system as such and of the different types of diagrams was based on the principles of information visualisation^[29, 30, 33-39] and was guided by the goal “[...] to amplify cognition by the [...] visual representation of abstract data”^[32]. The display of the data rested on general issues of information perception and interpretation as well as on issues of acceptance^[30, 35, 36]. Against this background, principles of cognition and information representation (see Table 3) could be applied to developing the visualisation system (*e.g.* correct perception, compactness, comprehensibility) as well as to the different types of diagrams (*e.g.* saliency, compactness, differentiation).

Table 3. Principles of cognition and information representation^[34, 40]

Principles of cognition / information representation	Explanation and application in this study
Correct perception	to support an appropriate understanding of the results.
Saliency	to direct the attention to the most important information.
Compactness	to present pieces of information together, which belong together.
Similarity	to visualise similar information similarly.
Differentiation	to present information depending on the context, <i>e.g.</i> distinguishing between results of individual benchmark participant and these of the peers (reference group).
Comprehensibility	to ease the understanding of the results, <i>e.g.</i> by providing descriptions.

In order to provide enough flexibility and to allow different perspectives, identical information could be displayed in different types of diagrams^[30, 33, 39, 41, 42] to enable the users to deliberately change their point of view. In addition to these principles of cognition and information representation, well-known design principles could be utilised for constructing the layout of the diagrams, *i.e.* colours in general^[43], font^[34, 43], background colours^[34, 37, 43], highlights^[29, 30, 33-35, 43], 2D visualisation^[33, 43] and presentation of digits^[30]. For example, the use of highlights could ease not only the perception but also facilitate the analysis of the data and help leading to appropriate actions^[35].

2.6 Development of the views

The views on the benchmarking results should have a graphical format first and foremost but should also display briefings or profiles in tabular format^[16]. Graphical formats should make use of well-known types of diagrams—if possible—in order to increase the recognition value and minimise the need to learn how to interpret the data^[36]. Diagram styles should vary in order to maintain a high level of attention, however similar information should be coded in a similar style. Raw indicators should be presented either in groups of similar indicators (multiple indicator view), if they could be summarised by a topic (*e.g.* clinical documentation) otherwise they should be presented as single indicators. If multiple indicators were grouped different compositions of diagrams were built to allow different perspectives^[39, 41, 42]. These perspectives were related either to the peer reference group, to the best of the group within the peer reference group or to the innovation adoption group.

2.7 Implementation of the visualisation system and the views

The visualisation system and the views were implemented in Microsoft Excel after importing the data and statistical parameters from a separate Microsoft Excel reference document (calculation results, individual results, statistical parameters). Statistical calculations and parameters were made in SPSS and Microsoft Excel. The output was realised in three ways: in terms of printouts of a benchmark brochure (short version) and electronic pdf prints (long version), which were semi automatically bundled (one per participant), as well as in terms of an interactive dashboard. Dashboards are well-known instruments to optimise processes, plans and strategies to present data in the best manner for decision-making^[44]. Content and design elements of different types of dashboards^[45] were implemented. By being able to represent different views side by side, the users could compare them and choose the one most appropriate for their purpose^[39]. All views were annotated and had comprehensive legends.

2.8 Evaluation

In order to optimise the visualisation in future IT benchmarks, ratings of the comprehensibility of the views should be obtained from the benchmark participants. In addition, the utility of the workflow composite score was measured and correlated with the comprehensibility. This evaluation should be conducted in a comparable way as the evaluation of a previous IT benchmarking^[28] to show the progress made. Comprehensibility was captured in four main questions with three corresponding sub questions related to the types of visualisation and the scores. A four point Likert scale was used with a codomain from very (comprehensible, helpful) to not (comprehensible, helpful) at all. The online questionnaire was sent to all benchmark participants.

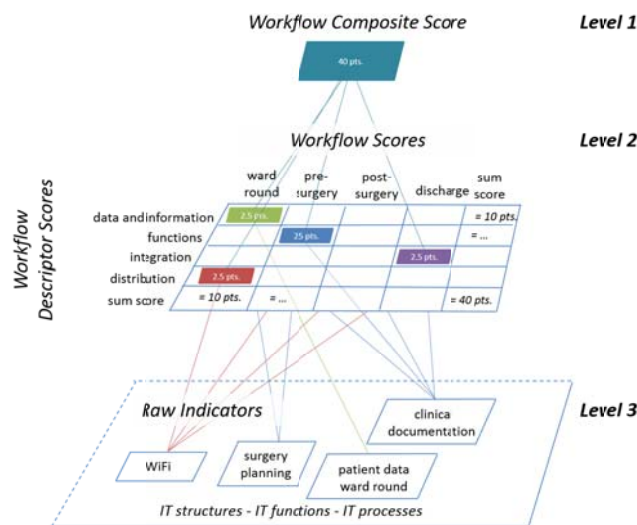


Figure 3. Schematic illustration of the three-dimensional multi-level model and its relationships

3 Results

3.1 Three-dimensional multi-level model as the visualisation system

In order to correspond with the hierarchical structure of the scoring system and to accommodate its scores and indicators we decided to construct a three-dimensional multi-level model (see Figure 3). The workflow composite score, the sum of all sub scores, resided at the top level (level 1). The workflow descriptor scores and the workflow scores constituted the next level (level 2). This level was defined by the four workflow descriptor scores (data and information, functions, integration and distribution) as the one axis and the four workflow scores (ward rounds, pre- and post-surgery processes, discharge) as the other. This scheme divided level 2 into a grid of 16 independent cells with each cell containing a

maximum value of 2.5 points. The sums per workflow (maximum of 10 points) or per workflow descriptor (maximum of 10 points) appeared at the rim of this level in eight additional cells so that the total of cells was 24. The sum of these cell values constituted the WCS, which could reach 40 points at maximum. All 92 raw indicators, the IT features describing the IT structure, the IT functions and the IT processes, resided at level 3. The multi-level model made the relationship between level 2 and level 3 transparent and showed which raw indicators contributed to which cells at level 2 (see Figure 3).

The three-dimensional multi-level model was presented to the benchmark participants in a hierarchical way as depicted in Figure 4.

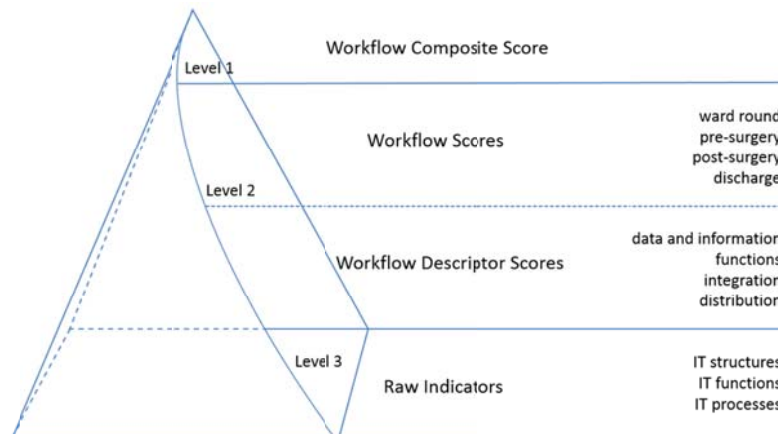


Figure 4. Visualisation of the multi-level model

3.2 Views on all levels of the three-dimensional multi-level model

3.2.1 Level 1

The WCS was graphically depicted as a composite score innovation view (see Figure 5), which was developed for the purpose of this benchmark and had not been used before. This view combined information about the peer reference group and about the categories of innovation adopters according to Rogers’ Diffusion of Innovations model, which segments the distribution of a certain score into the five categories: innovators (2.5%), early adopters (13.5%), early majority (34%), late majority (34%) and laggards (16%) [23]. The distribution of the WCS values was related to the peer reference group. As there were two criteria to build the reference groups, *i.e.* number of beds and ownership, two composite score innovation views per WCS could be produced. Innovation categories were coded by colour and the individual value of the benchmark participant was shown as a vertical bar with the exact WCS value annotated.

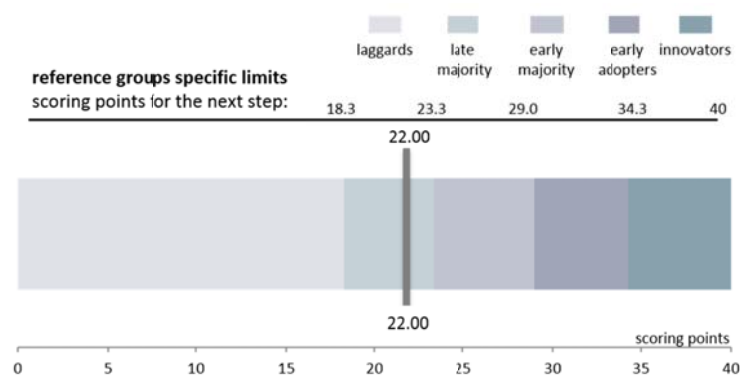


Figure 5. Composite score innovation view – workflow composite score (WCS)–public hospitals (n = 150)

The composite score innovation view required the score values to be distributed symmetrically, *i.e.* approximately normally distributed, so that the Rogers adopter categories were appropriate. As the WCS met these requirements^[18], this type of view could be utilised.

3.2.2 Level 2

At the level of the sub scores, the values were not normally distributed so that this innovation view was not appropriate. We therefore developed a view for these sub scores that segmented the distribution of the score values into two areas within each peer reference group. These two areas corresponded with the combined three categories innovators (2.5%), early adopters (13.5%), and early majority (34%) on the one hand and with the two categories late majority (34%) and laggards (16%) on the other. Statistically speaking, the distribution was split into two halves by the median. We called this view simple score innovation view (see Figure 6). A vertical bar denoted the individual value of the hospital. It was applied for all workflow scores and all workflow descriptor scores. In step one each of the workflow scores diagram could reach a maximum of 10 points (sum per workflow).

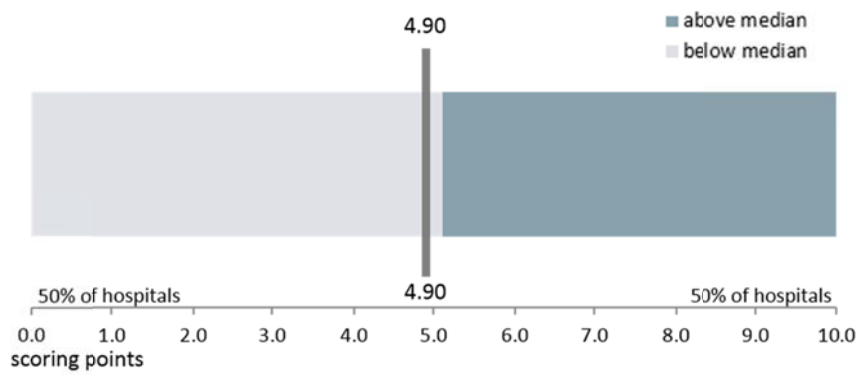


Figure 6. Simple score innovation view-example workflow score discharge-public hospitals (n=150)

To display the individual composition of a workflow score it was broken down to its workflow descriptor scores and displayed accordingly in a second step. Thus, there was a view for each of the descriptors data and information, functions, integration and distribution per workflow score. Here, the maximum value of scoring points was 2.5 points. Figure 7 shows an example of the specific combination of the workflow ward rounds with the workflow descriptor data and information. In other words, this score described how well ward rounds were supported by information technology with regard to the amount and appropriateness of the data and information provided for this workflow.

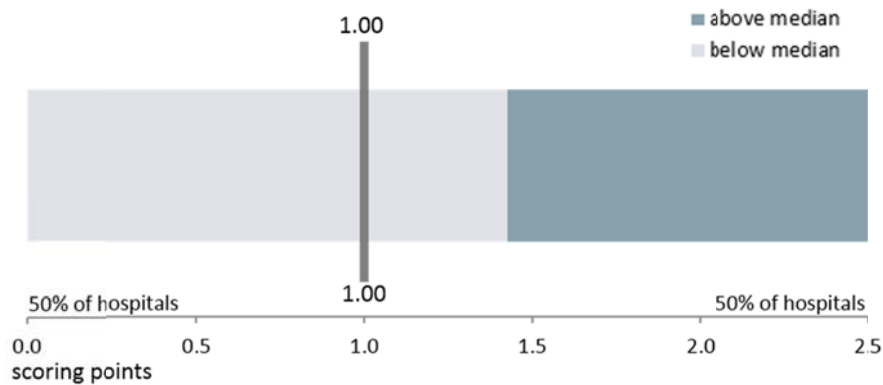


Figure 7. Simple score innovation view-example score combining the workflow ward rounds and the workflow descriptor data and information-public hospitals (n=150)

3.2.3 Score profiles at level 1 and 2

The meaning of all scores and score views was explained in so-called score profiles, which were attached to the views. They contained a definition, the purpose, the type, the aggregation level, related concepts, the codomain, the formula and the target group (see Table 4) [16].

Table 4. Indicator profile-workflow composite score (WCS)

Workflow composite score	
Definition and delimitation	The workflow composite score constitutes the highest level of all scores and measures the IT support of clinical processes per benchmark participant. It is composed of the sum of scores measuring the specific support for the processes ward rounds, pre-surgery process, post-surgery process and discharge.
Purpose	Measurement of overall clinical information logistics in an organisation
Indicator type	Composite score
Aggregation level	1
Related concepts	Diffusion of Innovations model by Rogers [23]
Codomain	0-40 points
Computation	Sum of all workflow scores of level 2 or sum of all workflow descriptor scores of level 2
Target group	Chief information officer, chief medical / nursing officer, chief executive officer

3.2.4 Level 3

The raw indicators used to build the WCS were depicted at level 3 in form of single indicator views and in form of multiple indicator views. Single indicator views represented—as the name reflects—one indicator and embraced the view types frequency view (see Figure 8), a view well-known in statistics, and performance bar (see Figure 9), which had been proposed by Few [43]. In both types of views, the category the individual hospital belonged to was marked: either in red as in case of the frequency views or in black as in case of the performance bars. Frequency views were employed for nominal data and ordinal data, performance bars only for ordinal data. Both types of single indicator views showed the distribution of values in the peer reference group.

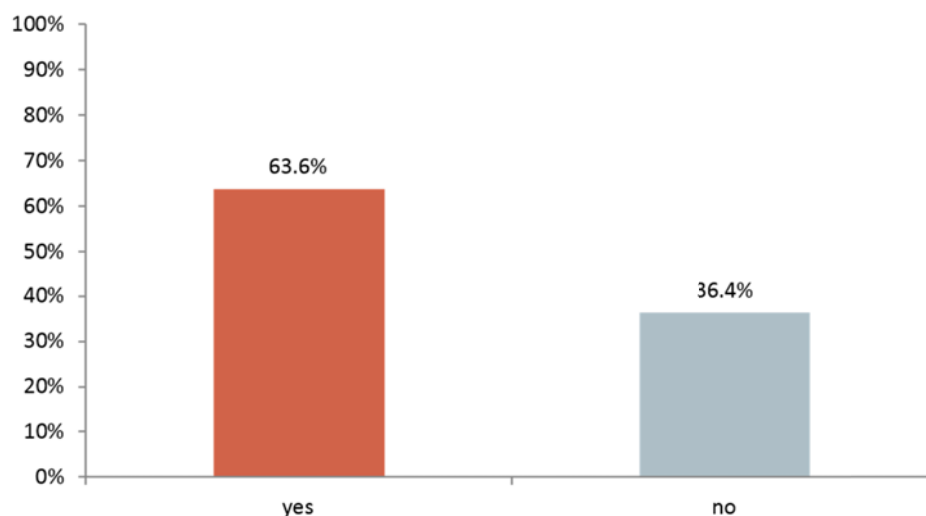


Figure 8. Single indicator view-frequency view-example integration of a workflow management system in the health information system—public hospitals (n=162)

Multiple indicator views encompassed the view types multiple distance view, multiple innovation view and multiple performance bar. The multiple indicator views reflected groups of raw indicators, e.g. measuring the degree of adoption of

different types of clinical documentation. Multiple performance bars resembled single indicator performance bars^[43] except that the bars of several indicators were depicted one below the other. Again, the individual value was presented within the distribution of values from all members of the peer reference group. In contrast, multiple distance views showed the distance between the individual value and the best of the peer reference group (see Figure 10) for the group of similar raw indicators. Technically speaking, multiple distance views were net diagrams, which are well established in statistics.

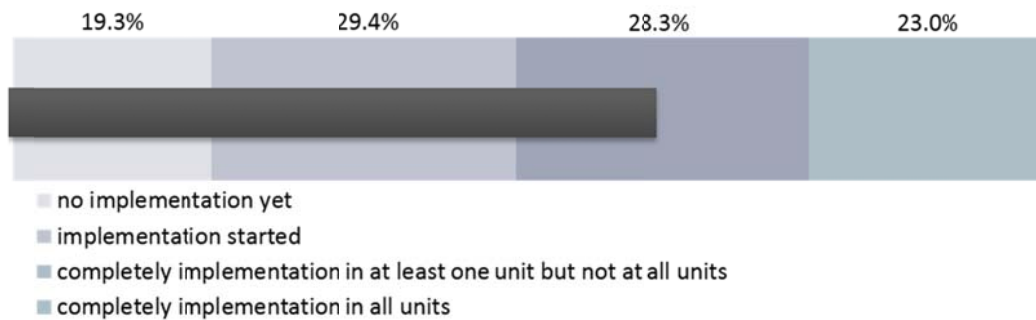


Figure 9. Single indicator view-performance bar-example degree of WiFi implementation-public hospitals (n=187)

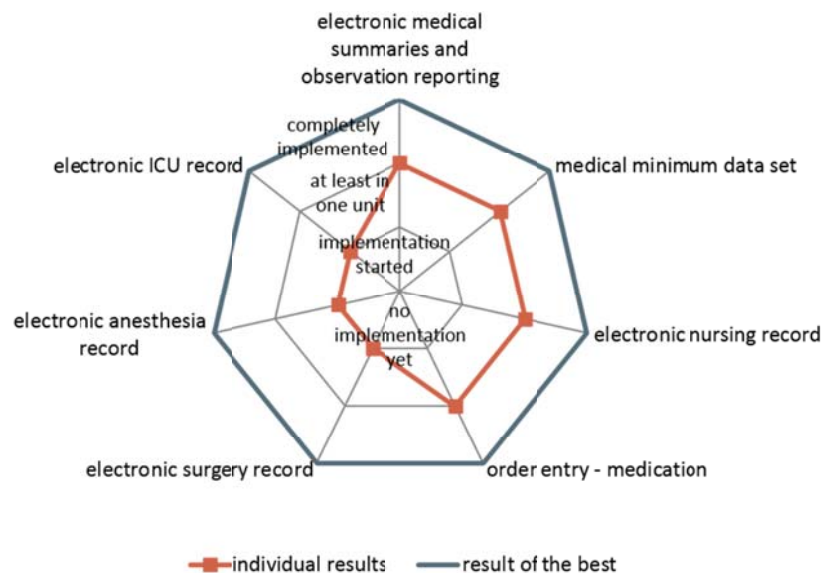


Figure 10. Multiple indicator view-multiple distance view-example existence of computerised systems for clinical documentation-public hospitals

The multiple innovation view, finally, expressed the implementation status of a group of raw indicators of the individual organisation (x-axis) in comparison to the frequency of this implementation status within the entire representative sample of hospitals the (y-axis). The resulting two-dimensional diagram was then divided into quadrants separating the entire field into a 2 × 2 matrix (see Figure 11). The two quadrants above the 50% line of the y-axis, *i.e.* the median, were called majority with the distinction of an early and a late majority. Organisations were called an early majority with regard to this IT function when they had implemented it in at least one unit (but not in all) or in all units. Organisations of the late majority had not implemented the IT function or were those in which the implementation started. If less than 50% of the hospitals showed this particular implementation status of the individual organisation, the organisation could be either an innovator or a laggard. Innovators had implemented the IT function (in at least one unit (but not in all) or in all units), laggards had not implemented it or the implementation (had just) started. The dot represented the category of the organisation and the frequency of this category (not the other potential categories) within the reference group. The

multiple indicator views were implemented as proposed by [28] (In this view, the understanding of innovators, early majority, late majority and laggards deviated from the Rogers definition [23] who classified adoption units according to the percentage of availability not the implementation status).

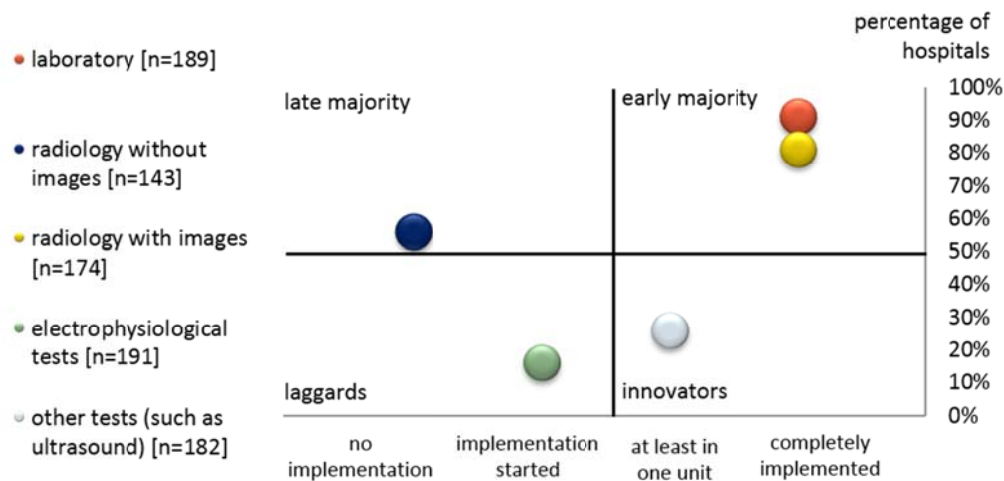


Figure 11. Multiple indicator view-multiple innovation view-example existence of computerised systems for order entry system for various tests-public hospitals

3.2.5 Toolbox of views

The views explained above in combination with tables constitute a toolbox of views that could be employed for the benchmarks of different types. Table 5 summarises these views in light of what they intended to measure, what benchmark type they supported and for which level they were most appropriate. Not all potential types of views were implemented at all levels. The ones actually used were selected followed the strategy:

- the benchmark should be the peer reference group in general
- the views should be most elaborate at level 3

Table 5. Toolbox of views (views in italics were implemented in the benchmark)

Measurement	Degree of innovation	Adoption rate	Status
Benchmark	Peer reference group	Peer reference group	Best within peer reference group
Level 1	<i>composite score innovation view</i>	table	table
Level 2	<i>simple score innovation view</i>	table	table
Level 3	<i>multiple innovation view</i>	<i>multiple performance bar, performance bar, frequency view</i>	<i>multiple distance view</i>

3.3 Publication of the benchmarks

The results of the benchmark were published stepwise and semi-automated using Microsoft Excel. A minimal version of the benchmark showing the results of three raw indicators of their choice were issued within one month after the end of the survey. In a second step, a short version with 16 tables and diagrams (e.g. WCS and its composition, workflow scores and their composition, clinical documentation functions, architecture of the information system) were produced, printed and sent as a paper brochure of 12 pages to the participants three months after the end of the survey. We chose a high quality paper version to provide a real (as opposed to a virtual) medium to the benchmark participants, which they could immediately show to their colleagues and the board members. Finally, they received the full report of 106 pages with 169 tables and diagrams as a pdf document three and a half months after closing the survey by eMail. In addition to the paper and pdf version, a fully dynamic dashboard was developed in Microsoft Excel (see Figure 12) and tested internally. This

version was not meant to be used by the participants because identification data of the other benchmark participants were included. The dashboard consisted of a navigation area that reflected the three-dimensional multi-level model and a display area to present the different views. Whenever different types of views were available at one level the user could choose in a drop-down menu. The dashboard prototype demonstrated the feasibility of the approach but also its limitations. The computation and display of the views in Microsoft Excel took about 5 to 10 minutes per view or had other performance problems due to the large dataset.

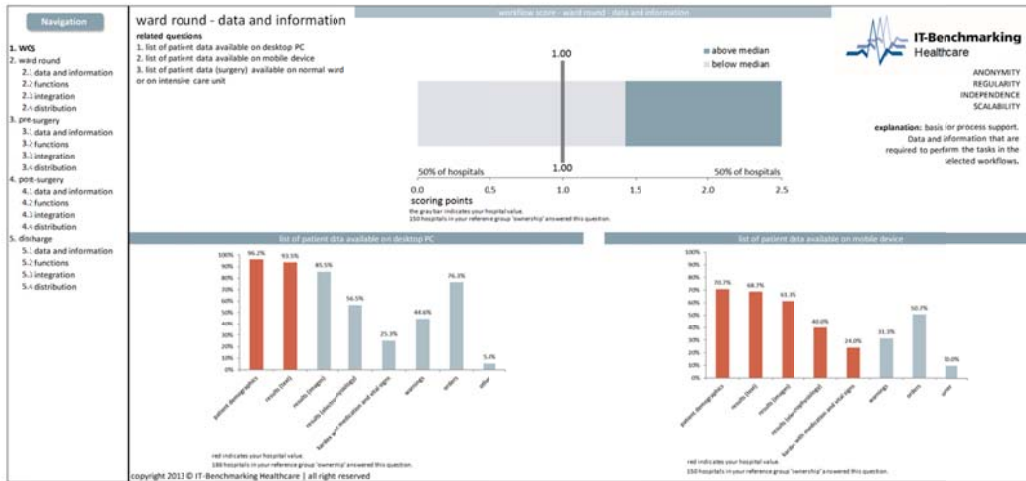


Figure 12. Screenshot of the dashboard-example workflow ward round-workflow descriptor score data and information

3.4 Evaluation

The evaluation questionnaire was returned by 67 out of the 199 benchmark participants, which corresponded with a response rate of 34%. The analysis showed that more than two third of the participants rated all scores as very comprehensible or comprehensible. However, there were differences with regard to the preference of the views in terms of their comprehensibility. While detailed and very well-known views, *i.e.* performance bars, frequency views, and tabular indicator profiles received positive ratings from more than 90% of the participants, combined and rather new views (the composite score innovation view, multiple distance view and multiple innovation view) were rated positively by only about 68% to 82% of the participants (see Table 6).

Table 6. Comprehensibility of the different types of views

Views	Very com-prehensible	Com-prehensible	Less com-prehensible	Not com-prehensible
Composite score innovation view (n=67)	23.9%	58.2%	14.9%	3.0%
Performance bars (n=66)*	37.9%	53.0%	7.6%	1.5%
Frequency views (n=65)	35.4%	58.5%	4.6%	1.5%
Multiple distance views (n=67)	22.4%	55.2%	16.4%	6.0%
Multiple innovation views (n=66)	22.7%	45.5%	28.8%	3.0%
Tabular indicator profiles (n=63)	27.0%	65.1%	6.3%	1.6%

Note. *Including multiple indicator view performance bars.

The Spearman rank correlation coefficient between the comprehensibility of the composite score innovation view and the utility of the workflow composite score was 0.59 (n = 64). This underlines the comprehensibility and utility of the workflow composite score.

4 Discussions

Performing IT benchmarks and presenting their results in an appropriate manner gains importance against the background of the increasing relevance of information management in organisations^[46], the involvement of the chief information officers into the decision making processes at board level and their role as link between the board and IT operations^[46,47].

A good balance between highly aggregated facts and details that illustrate these compact facts seems desirable. The hierarchical system of the WCS to measure the performance of the organisation in terms of its clinical information logistics capacity^[18] aims at providing this information balance. This study translated the WCS system into a three-dimensional multi-level model to visualise the benchmarking results at all levels. To this end, a set of views was developed (level 1 and 2) or taken from previous publications (level 3 and score profiles) and arranged to allow the comparisons with the peer reference groups of hospitals and the best of the groups. These views also enabled the benchmark participants to find out how the own organisation was classified into segments of innovation. The three-dimensional multi-level model seemed to be appropriate to present any highly aggregated lead indicator, here the WCS, and decompose it into its constituents. This approach could serve as a general principle of structuring benchmarks and guiding the benchmark participants through a jungle of data.

The benchmark results were made available to the participants in various paper and digital formats. A spreadsheet version proved to be less performing and difficult to ensure anonymity of the data. However, it leads the way towards a fully dynamic manner to present the results. Web-based solutions, where the data can reside safely in one place and where the user can interact with the benchmarking system in a self-determined way, seem most promising. This requires the technical basis of the benchmarks to be changed.

Apart from technical issues, which are important but were not the focus of this study, the visualisation system proposed and evaluated here provided encouraging results in terms of acceptance. The large majority of the users rated the various views as at least comprehensible. This applied to newly developed views and to conventional views. A preponderance of ease of comprehension towards traditional views is understandable and leads to the conclusion that these presentation formats definitely have a place in the canon of views. They must not be sacrificed to other more experimental versions. New versions, however, need to be explored. We found a group of views that were regarded as positive but may need some additional support to be accepted as comprehensible by nearly all participants. Mere explanations do not seem to be sufficient as these views came along with indicator profiles, which exactly provide these explanations. Further investigations on the role of innovation benchmarking are necessary^[22] because both views that targeted innovation came off least positive in comparison to the other views. Very often, benchmarking aims at identifying cost and quality indicators^[16,48,49], innovation has been focussed only rarely^[19,21,50]. CIOs and the (other) board members need to be convinced that implementing IT innovation can entail competitive advantage in terms of better patient care and lower costs in the long run. Therefore, innovation benchmarks are a valuable instrument to govern the implementation of IT innovation in particular against the background of the strategic plan of the organisation. This is interesting in the view of the involvement of the CIO in strategic decision-making^[46,47].

It is interesting that multiple distance views also received positive rates from only about 75% of the participants despite the fact that net diagrams are well established. In this case, the term multiple distance view might have been misleading and some participants may not have recognised what this term denoted.

The main limitation is, in fact, the small sample of participants who evaluated the benchmark and its visualisation. Thus, the conclusions that can be drawn from these results are limited and conducting further benchmarks and presenting the results according to the three-dimensional multi-level model seems unavoidable. Simply presenting the different views to a group of experts is not advisable because a valid evaluation requires a valid context in which the evaluation takes place, *i.e.* a real benchmark with real chief information officers, who are interested in their benchmark outcome.

5 Conclusions

With the increasing importance of measuring IT adoption and utilisation ^[51-53] IT benchmarks will gain more and more importance. The practical procedures for conducting these benchmarks will have to undergo severe changes. In order to benchmark health organisations on a regular basis, the *time to display* needs to be shortened and this can only work with as much automation as possible. This study investigated important building blocks along this way, *i.e.* a visualisation system and appropriate views. It provided a solid mechanism for displaying the workflow composite score, a reliable and valid indicator of IT process quality.

Acknowledgments

We wish to thank the chief information officers of the hospitals who participated in the survey and the benchmark. This study was funded by a research grant of the European Union and the German State Lower Saxony within the European Regional Development Fund (ERDF).

Competing interests

Parts of this work were presented at the eHealth 2014 Conference in Vienna on 22 May 2014 and published in the conference proceedings.

References

- [1] Chaudhry B, Wang J, Wu S, *et al.* Systematic review: impact of health information technology on quality, efficiency, and costs of medical care. *Ann Intern Med.* 2006 May; 144(10): 742-52. PMID:16702590. <http://dx.doi.org/10.7326/0003-4819-144-10-200605160-00125>
- [2] Finn NB, Bria WF. *Digital Communication in Medical Practice.* London: Springer-Verlag. 2009. <http://dx.doi.org/10.1007/978-1-84882-355-6>
- [3] Kuperman G. Reflections on AMIA-looking to the future. *J Am Med Inform Assoc.* 2013 Dec; 20: e367. <http://dx.doi.org/10.1136/amiajnl-2013-002435>
- [4] Gansert UA. IT als Unterstützungsinstrument des Prozessmanagements in Krankenhäusern. Berlin: Springer-Verlag; 2009: 233-65. (Behrendt I, König HJ, Krystek U, editors. *Zukunftsorientierter Wandel im Krankenhausmanagement. IT-Nutzungspotenziale im zukunftsorientierten Krankenhausmanagement.*) http://dx.doi.org/10.1007/978-3-642-00935-8_11
- [5] Simon A. Die betriebswirtschaftliche Bewertung der IT-Performance im Krankenhaus am Beispiel eines Benchmarking-Projektes. Wiesbaden: Vieweg + Teubner Verlag. 2010: 73-90. (Schlegel H, editor. *Steuerung der IT im Klinikmanagement. Methoden und Verfahren.*)
- [6] Hübner-Bloder G, Ammenwerth E. Key Performance Indicators to Benchmark Hospital Information Systems-A Delphi Study. *Methods Inf Med.* 2009 Nov; 48(6): 508-18. PMID:19893854. <http://dx.doi.org/10.3414/ME09-01-0044>
- [7] Heuwinkel K, Deiters W, Königsmann T, *et al.* Information Logistic and Wearable Computing. In: Wu J, editor. *Proceedings of the 23rd International Conference on Distributed Computing Systems Workshops; 2003; Providence; RI, USA.* IEEE computer society. 2003: 283-8. <http://dx.doi.org/10.1109/icdcs.2003.1203568>
- [8] Bucher T, Dinter B. Process Orientation of Information Logistics-An Empirical Analysis to Assess Benefits, Design Factors, and Realization Approaches. In: Sprague RH Jr, editor. *Proceedings of the 41st Hawaii International Conference on System Sciences; 2008 Jan: 7-10; Waikoloa, Big Island, Hawaii, USA.* IEEE computer society. 2008: 392. <http://doi.ieeecomputersociety.org/10.1109/HICSS.2008.361>
- [9] Kremer H. *Informationsmanagement.* 6th revised ed. Berlin: Springer Gabler. 2015: 799p.
- [10] De Haes S, Grembergen WV. *Enterprise Governance of Information Technology. Achieving Alignment and Value, Featuring COBIT 5.* 2nd ed. Cham: Springer. 2015: 167p.
- [11] Winter A, Haux R, Ammenwerth E, *et al.* *Health Information Systems. Architecture and Strategies.* 2nd ed. New York: Springer-Verlag. 2011: 337p. <http://dx.doi.org/10.1007/978-1-84996-441-8>
- [12] Jahn F, Winter A. A KPI Framework for Process-based Benchmarking of Hospital Information Systems. *Stud Health Technol Inform.* 2011; 149: 542-6. <http://dx.doi.org/10.3233/978-1-60750-806-9-542>
- [13] Kütz M. *IT-Controlling für die Praxis. Konzeptionen und Methoden.* 2nd revised and expanded ed. Heidelberg: dpunkt.verlag. 2013: 329 p.

- [14] Zairi M, Leonard P. Practical Benchmarking: The Complete Guide. Dordrecht: Springer Science+Business Media.1996: 262p.
- [15] Camp RC. Benchmarking. The Search for Industry Best Practices that Lead to Superior Performance. Milwaukee: ASQC Quality Press.1989: 299p.
- [16] Kütz M. Kennzahlen in der IT. Werkzeuge für Controlling und Management. 4th revised and expanded ed. Heidelberg: dpunkt.verlag. 2011:350p.
- [17] Thye J, Straede MC, Liebe JD, *et al.* IT-Benchmarking of Clinical Workflows: Concept, Implementation, and Evaluation. Stud Health Technol Inform. 2014; 198: 116-24. <http://dx.doi.org/10.3233/978-1-61499-397-1-116>
- [18] Liebe JD, Hübner U, Straede MC, *et al.* Developing a Workflow Composite Score to Measure Clinical Information Logistics. A Top-down Approach. Methods Inf Med. 2015 Oct; 54(5): 424-33. PMID:26419492. <http://dx.doi.org/10.3414/ME14-02-0025>
- [19] Feeny S, Rogers M. Innovation and Performance: Benchmarking Australian Firms. The Australian Economic Review. 2003 Aug; 36(3): 253-64. <http://dx.doi.org/10.1111/1467-8462.00285>
- [20] Ahmed PK. Benchmarking innovation best practice. Management & Technology. 1998; 5(1): 45-58. <http://dx.doi.org/10.1108/14635779810206803>
- [21] OECD. OECD Science, Technology and Industry Outlook 2014. Paris: OECD Publishing; 2014: 476p.
- [22] Hübner U. What Are Complex eHealth Innovations and How Do You Measure Them? Methods Inf Med. 2015; 54: 319-27. PMID:25510406. <http://dx.doi.org/10.3414/ME14-05-0001>
- [23] Rogers EM. Diffusion of Innovations. 5th ed. New York: Free Press; 2003: 551p.
- [24] Dugas M, Eckholt M, Bunzemeier H. Benchmarking of hospital information systems: Monitoring of discharge letters and scheduling can reveal heterogeneities and time trends. BMC Med Inform Decis Mak. 2008 Apr; 8: 15. PMID:18423046. <http://dx.doi.org/10.1186/1472-6947-8-15>
- [25] Amarasingham R, Diener-West M, Plantinga L, *et al.* Hospital characteristics associated with highly automated and usable clinical information systems in Texas, United States. BMC Med Inform Decis Mak. 2008 Sep; 8: 39. PMID:18793426. <http://dx.doi.org/10.1186/1472-6947-8-39>
- [26] Otieno GO, Hinako T, Motohiro A, *et al.* Measuring effectiveness of electronic medical records systems: Towards building a composite index for benchmarking hospitals. Int J Med Inform. 2008 Oct; 77(10): 657-69. PMID:18313352. <http://dx.doi.org/10.1016/j.ijmedinf.2008.01.002>
- [27] Jahn F, Baltschukat K, Buddrus U, *et al.* Benchmarking of hospital information systems – a comparison of benchmarking clusters in German-speaking countries. GMS Med Inform Biom Epidemiol. 2015 Aug; 11(1): Doc04. <http://dx.doi.org/10.3205/mibe000160>
- [28] Liebe JD, Hübner U. Developing and Trialling an Independent, Scalable and Repeatable IT-benchmarking Procedure for Healthcare Organisations. Methods of Information in Medicine. 2013 Jul; 52(4): 360-9. PMID:23877646. <http://dx.doi.org/10.3414/ME12-02-0016>
- [29] Ware C. Information Visualization. Perception for Design. 3rd ed. Waltham: Morgan Kaufmann. 2013: 512p.
- [30] Dix A. Introduction to Information Visualisation. Berlin: Springer-Verlag. 2013: 1-27. (Agosti M, Ferro N, Forner P, *et al.* editors. Information Retrieval Meets Information Visualization). http://dx.doi.org/10.1007/978-3-642-36415-0_1
- [31] Hübner U, Liebe JD, Straede, MC, *et al.* IT-Report Gesundheitswesen. Schwerpunkt IT-Unterstützung klinischer Prozesse. Osnabrück: Schriftenreihe des Niedersächsischen Ministeriums für Wirtschaft, Arbeit und Verkehr. 2014: 87p.
- [32] Card SK, Mackinlay JD, Shneiderman B. Information Visualization. San Diego: Academic Press; 1999: 1-35. (Card SK, Mackinlay JD, Shneiderman B, editors. Readings in Information Visualization. Using Vision to Think).
- [33] Schulz HJ, Schumann H. Visualizing Graphs-A Generalized View. In: Banissi E, Aslak Burkhard R, Ursyn A, *et al*, editors. Proceedings of the Information Visualization; 2006 Jul: 5-7; London, England. IEEE computer society; 2006: 166-73. <http://doi.ieeecomputersociety.org/10.1109/IV.2006.130>
- [34] Stapelkamp T. Informationsvisualisierung Web-Print-Signaletik. Berlin: Springer-Verlag.2013: 454p.
- [35] Liang J, Huang ML. Highlighting in Information Visualization: a Survey. In: Banissi E, Bertschi S, Burkhard R, *et al*, editors. Proceedings of the Information Visualization. 2010 Jul: 26-9; London, England. IEEE computer society. 2010: 79-85. <http://doi.ieeecomputersociety.org/10.1109/IV.2010.21>
- [36] Reiterer H, Mann TM, Mußler G, *et al.* Visualisierung von entscheidungsrelevanten Daten für das Management. HMD-Praxis der Wirtschaftsinformatik. 2000; 212: 71-83.
- [37] Ware C. Visual Thinking for Design. Burlington: Morgan Kaufmann. 2008: 197p.
- [38] Card SK, Mackinlay JD, Shneiderman B. Readings in Information Visualization. Using Vision to Think. San Diego: Academic Press. 1999: 686 p.
- [39] Heer J, Shneiderman B. Interactive Dynamics for Visual Analysis. Communications of the ACM. 2012 Apr; 55(4): 45-54. <http://dx.doi.org/10.1145/2133806.2133821>

- [40] Preim B, Dachselt R. *Interaktive Systeme*. 2nd ed. Berlin: Springer-Verlag. 2010: 628p.
<http://dx.doi.org/10.1007/978-3-642-05402-0>
- [41] Elias M, Bezerianos A. Annotating BI Visualization Dashboards: Needs & Challenges. In: Konstan JA, Chi EH, Höök K, editors. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Austin, Texas, USA. ACM. 2012 May 5-10; 2012: 1641-50. <http://dx.doi.org/10.1145/2207676.2208288>
- [42] Chen C. Information visualization. *WIREs Comp Stat*. 2010 Jul/Aug; 2(4): 387-403. <http://dx.doi.org/10.1002/wics.89>
- [43] Few S. *Information Dashboard Design. The Effective Visual Communication of Data*. Sebastopol: O'Reilly Media Inc. 2006: 211p.
- [44] Eckerson WW. *Performance Dashboards. Measuring, Monitoring, and Managing Your Business*. 2nd ed. Hoboken, New Jersey: John Wiley & Sons Inc. 2011: 318p. PMID:22211179.
- [45] Rasmussen N, Chen CY, Bansal M. *Business Dashboards. A Visual Catalog for Design and Deployment*. Hoboken, New Jersey: John Wiley & Sons Inc. 2009: 286p.
- [46] Moghaddasi H, Sheikhtaheri A. CEO is a Vision of the Future Role and Position of CIO in Healthcare Organizations. *J Med Syst*. 2010 Dec; 34: 1121-8. PMID:20703595. <http://dx.doi.org/10.1007/s10916-009-9331-4>
- [47] Bradley RV, Byrd TA, Pridmore JL, *et al*. An empirical examination of antecedents and consequences of IT-governance in US hospitals. *J Inf Technol*. 2012 Jun; 27(2): 156-77. <http://dx.doi.org/10.1057/jit.2012.3>
- [48] Horváth P. *Controlling*. 12th ed. München: Verlag Franz Vahlen; 2012: 832p.
- [49] Komonen K. A cost model of industrial maintenance for profitability analysis and benchmarking. *Int J Production Economics*. 2002; 79(1): 15-31. [http://dx.doi.org/10.1016/S0925-5273\(00\)00187-0](http://dx.doi.org/10.1016/S0925-5273(00)00187-0)
- [50] Guimaraes T, Langley K. Developing Innovation Benchmarks: An Empirical Study. *Benchmarking for Quality Management & Technology*. 1994; 1(3): 3-20. <http://dx.doi.org/10.1108/14635779410073283>
- [51] Adler-Milstein J, DesRoches CM, Furukawa MF, *et al*. More than half of US hospitals have at least a basic EHR, but stage 2 criteria remain challenging for most. *Health Aff*. 2014 Sep; 33(9): 1664-71. PMID:25104826.
<http://dx.doi.org/10.1377/hlthaff.2014.0453>
- [52] Diana ML, Harle CA, Huerta TR, *et al*. Hospital characteristics associated with achievement of meaningful use. *J Healthc Manag*. 2014 Jul/Aug; 59(4): 272-86. PMID:25154125.
- [53] Kruse CS, Bolton K, Freriks G. The Effect of Patient Portals on Quality Outcomes and Its Implications to Meaningful Use: A Systematic Review. *J Med Internet Res*. 2015 Feb; 17(2): e44. PMID:25669240. <http://dx.doi.org/10.2196/jmir.3171>