ISSN 2029-9338 (ONLINE)
ISSN 1822-6515 (CD-ROM)

MEASUREMENT ISSUES IN HEALTH CARE SERVICE QUALITY

Zaneta Piligrimienė

Kaunas University of Technology, Lithuania, zaneta.piligrimiene@ktu.lt

crossref http://dx.doi.org/10.5755/j01.em.17.1.2280

Abstract

No system can be designed and operated effectively if the quality of the product or service is not understood or correctly measured. Therefore, it seems that discussions about service quality, its dimensionality, structure and measurement will never end. Logically, without a valid measure, it would be difficult to establish and implement appropriate tactics or strategies for service quality management. Therefore, for the success of health care organizations, as well as for other service industries, accurate measurement of service quality is of ultimate importance.

This study to some extent reiterates the researches done by Lee, Delene, Bunda and Kim (2000) and Walbridge & Delene (1993) and can be treated as efforts to examine the effect of alternative measurement methods on the results of health care service quality as perceived by physicians, patients and administrators.

Keywords: Health care, service quality, quality measurement, measurement scales.

JEL Classification: C18, I11, I19, M10, M31.

Introduction

Health care service quality measurement issues have been discussed in numerous studies (Choi et al., 2005; Kang, 2006; Dagger et al., 2007; Qin, 2009; Bowers, Kiefe, 2002; Piligrimienė, Bučiūnienė, 2011; De Man et al., 2002; Eiriz, Figueiredo, 2005; Gupta, 2008; Chang et al., 2006; etc.). However, only few studies (Walbridge & Delene, 1993; Lee, Delene, Bunda & Kim, 2000), involved systematic examination of the psychometric properties of measurement scales, using more than one measurement method for the same concept. Usually studies on service quality deal more with pragmatic and managerial issues for health care services, or focus on checking dimensionality of service quality, employing one measurement method, and are inadequate for testing the construct validity of the scale. Construct validity is defined as the degree of correspondence between constructs and their measures (Babbie, 2004). The two main sources of variance in measures of a construct are the construct being measured and the measurement error. Single measures do not allow us to make an assessment of measurement error. With a single method, we cannot separate trait variance from unwanted method variance. Thus, for construct validation we need to use multitrait - multimethod data, which provides us with the correlation matrix for two or more traits where each trait is measured by two or more methods. Moreover, it is well known that measurement method can affect the nature of a respondent's evaluation. The possibility to assess the differences in respondent's understanding about the object being researched is only available when we perform measurement with more than one method. Thus, **the problem of this article** can be formulated as the following question: *How different measures affect the results of perception of health care service quality?*

The objective of this paper – to explore the effect of different measurement methods on the results of health care service quality perception as assessed by physicians, patients and managers of health care organizations.

Research methods: systematic and comparative analysis of scientific literature; empirical quantitative survey research employing self-administered questionnaires.

Theoretical background

Multi-method research. Multi-method research usually means the application of two or more sources of data or research methods to the investigation of a research question. Such research is also frequently referred to as mixed methodology (Diamantopoulos et al., 2012). The rationale for mixed-method research is that findings in social sciences usually are deriving from a single research method and are vulnerable to incorrect inferences and conclusions if measurement error is affecting those findings. The mixed-method research is based on the principle of triangulation. Triangulation is a powerful technique that facilitates validation of data through cross verification from more than two sources. In particular, it refers to the application and combination of several research methodologies in the study of the same phenomenon (Bogdan & Biklen, 2006). So, triangulation is often used to indicate that more than two methods are used in a study with a view to double (or triple) checking results. It has to ensure that researchers are not too reliant

on a single research method when investigating one or another research problem. Denzin (1970) identified 4 basic types of triangulation [4]:

- Data triangulation: involves time, space and persons
- Investigator triangulation: involves multiple researchers in an investigation
- Theory triangulation: involves using more than one theoretical scheme in the interpretation of the phenomenon
- Methodological triangulation: involves using more than one method to gather data.

The latter one can be further divided into *within-method* and *between-method* triangulation. Between-method triangulation involves contrasting research methods, for example, survey and observation. Within-method triangulation involves the use of varieties of the same method to investigate a research issue; for example, two contrasting scales to measure service quality in the same self-administered questionnaire.

Single-item versus Multi-item scale. In the context of measurement considerations, the use of several contrasting scales for measurement of the same construct raises new questions, in particular about advantages and disadvantages of Single-item and Multi-item measures. Many researchers agree that when used in social science, multi-item measures can be superior to a single, straightforward question. Most of them identify two major reasons for that. First, the reliability of a multi-item measure is higher than of a single-item measure. With a single question, respondents are less likely to give consistent answers over time. A multi-item measure has several questions targeting the same social issue, and the final composite score is based on all questions, so the composite score is more consistent over time. Thus, the multi-item measure is more reliable than a single question. Second, the validity of a multi-item measure can be higher than a single question. Many measured social characteristics, especially such constructs as service quality, are broad in scope and simply cannot be assessed with a single question. A multi-item measure will allow subjects to describe their feelings about different aspects of the subject under study. This will greatly improve the precision and validity of the measure. Multi-item measures will be necessary to cover more content of the measured characteristic and to fully and completely reflect the construct domain.

On the other hand, some researchers argue that single-item measures may elicit responses that are as reliable as the multi-item measures when knowledgeable respondents are involved (Lee et al., 2000). The single-item method has the potential advantage of parsimony for the respondent. Therefore, when there is no or little difference between the explanatory power of single and multi-item methods, the single-item method may be preferable.

Measures in health care quality. Health care service quality (HCSQ) definition and understanding, as well as its measurement, has always been problematic. HCSQ is a multidimensional complex concept and its perception and evaluation depends on the perspective of different stakeholders who make an evaluation. Health care providers (health care professionals, managers of HQ organizations) and patients have different perception of what is quality in health care. Enumeration of different definitions of health care service quality shows that in looking for the way to define and measure HCSQ it may be more advisable to study quality through dimensions that constitute it. According to the leading author in health care literature, Donabedian (2003), health care quality can be characterized by seven attributes that include efficacy, effectiveness, efficiency, optimality, acceptability, legitimacy and equity. These quality dimensions, taken singly or in a variety of combinations, constitute a definition of health care quality. Two other leading concepts / instruments for service quality measurement that have an extensive use in various fields, are those of Grö nroos (1984) model of „technical-functional“ quality and Parasuraman, Zeithaml and Berry (1988) SERVQUAL model, which distinguishes 5 service quality dimensions: reliability, assurance, tangibles, responsiveness and empathy. There is a large number of various modifications and replications of use of these three leading concepts in health care quality research field. Naturally, multi-item scales were used for measurement of such a multi-dimensional construct as health care quality, independently from the type of respondent being questioned.

Within this research we try to explore if there some differences in understanding what constitutes health care service quality when using more than one method for construct measurement. That would allow making some inferences about the construct validity. At the same time, we hope to reveal the appropriateness of multi-item and single-item measures in the research of health care service quality context. Moreover, we try to find out differences in perception among three distinct groups of quality evaluators: patients, health care professionals and administrators/ managers of health care organizations.

ISSN 2029-9338 (ONLINE)
ISSN 1822-6515 (CD-ROM)

Research Design

Questionnaire Development. The survey method, employing the self-administered questionnaire was chosen as the most appropriate data collection method for this particular research. As the base of multi-method research, two forms of triangulation were used: data triangulation (the data was collected from three different samples of respondents) and methodological within-method triangulation (two different scales for measuring health care service quality was used in the same questionnaire). This decision was made in order to reduce the measurement error that might be produced with only one measurement method. The second rationale for this was the intent to evaluate the degree to which different respondents understands the same concept as expressed in different ways, and to compare that degree of similarity across the samples of health care professionals, patients and managers of health care organizations.

The research instrument was developed using two methods of measurement in it: single-item global rating method and multi-item rating method. The single-item global rating scale provided the respondent with 13 "global" dimensions of health care service quality. Respondents were asked to indicate the relative importance of 13 global service quality dimensions for their perception of health care service quality, based on 5-point scale, where 1 – unimportant and 5 – of critical importance, respectively. Respondents also were asked to indicate the relative importance of 72 different health care service attributes that were designed to operationally define the 13 global dimensions of service quality, based on the same 5-point scale. The 13 dimensions, their origins, definitions and number of items in questionnaire, corresponding each of the dimension, can be found in *appendix 1*.

Sampling. 5 hospitals in Lithuania were selected for data collection, after the probability cluster sample procedures and all of the possible respondents in these 5 hospitals (patients, health care professionals and administrators/managers) were intended to be questioned. Overall 2150 questionnaires have been distributed, and 618 properly answered questionnaires were collected with the response rate of 29 per cent. The number of collected questionnaires according to the respondent group was the following: 225 questionnaires from patients sample (response rate – 23 per cent), 318 questionnaires from health care professionals sample (response rate – 32 per cent) and 75 from administrators (response rate – 50 per cent).

Data was analysed using descriptive statistics, Factor analysis, Spearman's correlation analysis tests. The reliability of the scales was statistically verified using Cronbach's alpha coefficient, with the indexes higher than $\alpha = 0.7$.

Research Findings

Table 1 presents a summary profile of three samples. Sample of patients was based on 4 variables: gender, age, education and hospital department (in which they were treated during the survey). Sample of health care professionals was based on 5 variables: gender, age, position, work experience and department/ speciality. Sample of administrators was based on 4 variables: gender, age, position and administrative work experience. Gender distribution in all three samples shows the dominance of female respondents, with a little bit larger distribution in patients' sample.

Table 1. Sample profile

		Patients, %	Health care professionals, %	Administrators, %
Gender:	Male	35,0	7,3	14,7
	Female	65,0	92,7	85,3
Age:	to 45 years	65,6	55,3	41,3
	over 45 years	34,4	44,7	58,7
Education	Without high education	75,2		
	High education	24,8		
Position:	Physician		16,4	
	Nurse		77,8	
	other		5,8	
Position:	Head of department			46,3
	Senior nurse			46,3
	Manager of hospital			7,4
Work experience:	to 10 years		6,4	
	11-20 years		36,4	
	over 20 years		57,2	

ISSN 2029-9338 (ONLINE)
ISSN 1822-6515 (CD-ROM)

		Patients, %	Health care professionals, %	Administrators, %
Administrative work experience:	to 5 years			27,1
	5-10 years			22,9
	over 10 years			50,0
Specialty/ department:	Surgery	17,5	25,5	
	Obstetrics-gynaecology	20,9	8,9	
	Internal medicine	48,3	22,7	
	Admission – diagnostics		13,5	
	Trauma – rehabilitation	6,6	5,0	
	Paediatrics		10,6	
	Neurology	6,6	9,2	
	Sustaining nursing		4,6	
		N-225	N - 318	N - 75

The subjects in patient's sample ranged in age from 16 to 82 years, with a mean age of 43 years (SD=8.49). Almost 25 per cent of respondents in this sample had a high (university) education; more than a half (54 per cent) of them had higher that secondary formal education. During the survey we reached respondents from 18 different hospital departments, and the majority were taking cure from diseases that could be attributed to internal medicine.

The subjects in professional's sample ranged in age from 24 to 71 years, with a mean age of 45.5 years (SD=7.90). The work experience ranged from 1 to 44 years in health care setting, with a mean of 22.86 (SD = 8.25), indicating a sample of very experienced medical personnel. Distribution of respondents according to work department (speciality) in the hospital encompasses 8 departments.

The subjects in manager's sample ranged in age from 34 to 66 years, with a mean age of 48.12 years (SD = 7.87). Distribution according to administrative work experience ranged from 1 to 30 years, with a mean of 12.5 (SD = 7.88).

The distribution of health care professionals and managers' samples on the indicated characteristics generally reflect those of health care providers' population in Lithuanian hospitals with some bias in gender distribution. Thus although the findings might be not generalizable nationally across all kind of health care services (e.g., primary health care centres) or internationally, they still can bring the valuable and representative information on the question under investigation.

Research results

Reliability of the multi-item scale. At first, in order to make a meaningful comparison of the results, it is necessary to derive a composite score for each of the 13 HCSQ dimensions measured by the multi-item scale. For this purpose, the confirmatory factor analysis for each dimension in multi-item scale was performed, that allowed to confirm that all 13 dimensions maintained their initial structure. The level of internal consistency was checked as a way of assessing the homogeneity of items comprising each dimension. The Cronbach's alpha coefficients for the 11 dimensions (2 dimensions were operationalized with only one item) ranged from 0.74 to 0.86, with a mean of 0.83 (see *Table 2*). This high degree of internal consistency showed that multiply variables were good descriptors of a particular quality dimension and allowed us to sum the ratings to get composite scores for each of the 13 dimensions. Each composite score indicated a measure of each HCSQ dimension obtained by the multi-item scale. These composite scores were used for a comparative analysis along with the scores assessed by the single-item global scale.

Table 2. Results of confirmatory factor analysis on multi-item scale

Dimension	KMO	Factor loadings % of Variance	Cronbach α
Tangibles	0,873	41,5	0,854
Accessibility	0,843	35,1	0,813
Competence	0,596	40,8	0,745
Responsiveness	0,822	57,4	0,813
Communication	0,828	61,3	0,870
Reliability	0,760	52,0	0,814
Respect & Caring	0,838	49,0	0,850

ISSN 2029-9338 (ONLINE)
ISSN 1822-6515 (CD-ROM)

Dimension	KMO	Factor loadings % of Variance	Cronbach α
Appropriateness	0,768	70,8	0,861
Continuity	0,687	65,7	0,739
Safety	0,790	59,8	0,832
Patients outcome	0,697	71,5	0,800
Effectiveness*	-	-	-
Efficiency*	-	-	-

* „Effectiveness" and „Efficiency" were operationalized with one item and particular indicators were not calculated

Results of correlation analysis. After getting a proof about possibility to compare the results obtained by two different methods, the correlation analysis between the results of different measurement methods was performed in a total sample and in each of the three samples separately (see *Table 3*). Spearman's rho correlation coefficient was calculated, because the Kolmogorov-Smirnov test showed that data were not distributed normally.

Table 3. Correlation between the results of different measures*

Quality dimension	Correlation coef. Total sample	Correlation coefficients		
		Patient sample	Professional sample	Administrator sample
Tangibles	0,484	0,563	0,431	0,444
Accessibility	0,450	0,427	0,457	0,525
Competence	0,318	0,374	0,276	0,296
Responsiveness	0,404	0,416	0,406	0,327
Communication	0,499	0,508	0,518	0,394
Reliability	0,459	0,459	0,473	0,332
Respect & Caring	0,511	0,543	0,473	0,597
Appropriateness	0,399	0,438	0,352	0,439
Continuity	0,543	0,581	0,535	0,454
Safety	0,587	0,581	0,566	0,690
Patient outcome	0,461	0,482	0,416	0,517
Effectiveness	0,472	0,456	0,461	0,575
Efficiency	0,597	0,648	0,539	0,706
Total	*0,804*	*0,827*	*0,777*	*0,839*

* - p< 0,01

Correlation analysis showed the statistically significant, but only moderately strong relations between the results of measurement on each dimension (ρ_{max}= 0,597 in total sample). Significant correlations between each 13 dimensions suggest that both scales – multi-item as well as single-item – measure the same construct (health care service quality), but the low correlation coefficients indicate the difference of the meaning of each dimension for respondents when expressed in different ways.

However, the high total correlation coefficients allow suggesting that both measures provide a view about health care service quality, and a total score for HCSQ would be similar if measured either with multi-item scale or single-item scale.

Analysis of correlation in different samples showed the highest correlation coefficients in administrator's sample. This sample produced the highest correlation on 7 dimensions in comparison to other two samples. Relatively highest correlation coefficients in a sample suggest that administrators perceive HCSQ similarly (in the way the researcher wants them to understand) despite of the measurement method used. It confirms the presumption that the more knowledgeable the respondents are about the subject of investigation, the more consistent results they would provide, independently from measurement method.

Sample of patients showed the highest correlation coefficients on 5 dimensions when comparing to other samples. Such dimensions as „tangibles", „competence", „responsiveness", „appropriateness" and „continuity" had relatively similar meanings for patients no matter expressed in one definition or by several different aspects.

The sample of health care professionals showed the least similar evaluation scores as performed with different methods. In comparison with other two groups, only two dimensions, namely, „communication"

ISSN 2029-9338 (ONLINE)
ISSN 1822-6515 (CD-ROM)

and „reliability“ had correlation coefficients higher than in other samples. These are a little bit unexpected results, taking into account that this group should also be knowledgeable about HCSQ.

Comparison of means. The moderate relations between the results of two measurements enabled to make an analysis of mean ranks on each SQ dimension as received from two methods (see *Table 4*).

Table 4. Comparison of the results of different measures

Quality dimension	Multi-item scale		Quality dimension	Single-item global scale	
	Mean*	SD		Mean	SD
Appropriateness	4,44	0,56	Competence	4,45	0,64
Effectiveness	4,26	0,69	Responsiveness	4,35	0,67
Safety	4,09	0,64	Appropriateness	4,35	0,67
Continuity	4,04	0,64	Effectiveness	4,35	0,62
Reliability	3,98	0,62	Patient outcome	4,3	0,63
Patient outcome	3,92	0,64	Reliability	4,25	0,70
Responsiveness	3,81	0,62	Respect & Caring	4,06	0,75
Communication	3,81	0,65	Safety	4,06	0,82
Efficiency	3,79	0,87	Communication	3,92	0,78
Accessibility	3,76	0,50	Continuity	3,92	0,75
Competence	3,66	0,51	Efficiency	3,8	0,82
Respect & Caring	3,66	0,59	Accessibility	3,49	0,83
Tangibles	3,65	0,56	Tangibles	3,33	0,76

* - means represent the average sum of importance scores assigned to each item group, representing separate SQ dimension.

As we can see, the positions of SQ dimensions ranks are not the same when looking at their means from different measurements. Only one dimension „tangibles” is on the same, the last, position on both scales. The lowest rating on „tangibles” with both methods indicates that intangible elements dominate over tangibles when talking about what is important in health care service quality. That finding corresponds with the similar findings of Walbridge & Delene (1993).

Relative positions of such SQ dimensions as „appropriateness“, „effectiveness“, „reliability“ and „patient outcome“ differ only slightly and keeps positions in the top six. It should be noted that „appropriateness“ dimension was ranked as the most important with multi-item scale (mean 4.44) and as in the second-third place of importance (mean 4.35) with single-item scale. It allows suggesting that „appropriateness" is perceived by respondents as very important part of health care service quality.

Dimension „responsiveness” was rated as second of importance (mean 4.35) when measuring it on the single-item scale, but on the multi-item scale it was rated only as seven in a row of importance (3,81). Taking into account that „responsiveness“ was operationalized with 5 items on multi-item scale, we may conclude that the way the service quality dimensions are described (by one definition or by multiply characteristics) leads to different understanding, and, consequently, to different ratings.

Ratings of importance on „communication“, „accessibility” and „efficiency” differ only slightly and these dimensions are at the end of the importance row independently from the measurement method. And on the contrary, very obvious difference can be noticed between importance ratings of such dimensions as „safety”, „continuity” and „respect & caring”. „Safety” and „continuity” have bigger importance ratings on multi-item scale, meanwhile „respect & caring” – on single-item scale.

Still, the biggest gap was found between the ratings on quality dimension „competence“. This dimension on global item scale was rated as the most important (mean 4.45), meanwhile on multi-item scale it was scored as the third from the end of the row, i.e., on the 11th position (mean 3.66). Such results might be explained with the use of such „indirect“ quality measures on multi-item scale, describing „competence“ as „physicians that performs scientific research“ or „physicians that publish the results of their research in scientific journals“, that might artificially lowered the importance score of this dimension. That raises a question should the indirect measures be used when defining health care service quality at all. It‘s worth noting that very similar results were found in other empirical studies (Walbridge & Delene, 1993; Gupta, 2008), showing that scientific performance of medical personnel for most of respondents is not associated with health care service quality. The mean score for „competence” without two mentioned items would be

4,19 (SN – 0,53), showing the third rank of importance (on multi-item scale), which confirms the soundness of the above reasoning.

Conclusions

The findings allow concluding that both terminology, used for describing a study construct, and the measurement method affect the measurement results. Health care service quality, operationalized with single-item „global" measures and operationalized with multi-item measures (describing specific characteristics of each quality dimension) can have and do have a different meaning to respondents.

However, the results show that the use of single-item scale allows getting as reliable results as the use of multi-item scale, especially if the respondents are well acquainted with the subject under investigation. According to the results of our empirical research, managers (administrative employees) of health care organizations (hospitals) were the most knowledgeable about health care service quality. The correlation coefficients in this sample were the highest on 7 quality dimensions comparing to other samples (patients and health care professionals). Comparison among three samples showed that patients provided a little bit more consistent results on both scales than health care professionals did. This may be a result of numerous sub-attributes associated with separate dimensions, which may or may not have been evident to health care professionals in their ratings of „global" dimensions. As such, the multi-item scale, representing different quality characteristics, is probably more meaningful to health care professionals.

Even so, the use of single-item scale might be very useful if the purpose of the research is to get the general understanding of what constitutes health care service quality. High total correlation between both measurements suggests that they both measure the same construct. In the case of development of general understanding about health care service quality, single-item scale has a big advantage over multi-item scale, because it has a shorter length, less monotonous and time consuming for respondents, and may lead to greater survey effectiveness, especially in clinical populations. Still, the assessment of reliability level for single-item measures is problematic. When the research is diagnostic in nature, when the focus is on particular service characteristics, trying to identify the areas for improvement, the multi-item scale is more efficient. It can generate the detailed information on specific aspects of service quality that can be then used for actions in the quality improvement arena. So, the researcher should always weight the goals of the research and the level of needed depth of information before he/she selects the measurement method.

References

1. Babbie, E. (2004). Practice of Social Research. 10th Edition, Wadsworth, Cengage Learning.
2. Bogdan, R. C. & Biklen, S. K. (2006). Qualitative research in education: An introduction to theory and methods. Allyn & Bacon. ISBN 978-0205512256.
3. Bowers, M.R. & Kiefe, C.I. (2002). Measuring Health Care Quality: Comparing and Contrasting the Medical and the Marketing Approaches. American Journal of Medical Quality, vol.17, no.4, 136-144.
4. Bryman, A. Triangulation. Available: http://www.referenceworld.com/sage/socialscience/triangulation.pdf
5. Chang, J.T., Hays, R.D., Shekelle, P.G., et al. (2006). Patients' global ratings of their health care are not associated with the technical quality of their care. Annals of Internal Medicine, vol.144, 665-672.
6. Choi, K., Lee, H., Kim, C., Lee, S. (2005). The service quality dimensions and patient satisfaction relationships in South Korea: Comparisons across gender, age, and types of service. Journal of Services Marketing, vol.19, no.3, 140-149.
7. Dagger, T.S., Sweeney, J.C., Johnson, L.W. (2007). A Hierarchical Model of Health Service Quality: Scale Development and Investigation of an Integrated Model. Journal of Service Research, vol.10; no.2, 123-142.
8. De Man, S., Gemmel, P., Vlerick, P., Van Rijk, P., Dierckx, R. (2002). Patients' and personnel's perceptions of service quality and patient satisfaction in nuclear medicine. European Journal of Nuclear Medicine, vol. 29, no. 9, 1109-1117.
9. Diamantopoulos, A., Sarstedt, M., Fuchs, Ch., Wilczynski, P., & Kaiser, S. (2012). Guidelines for Choosing between Multi-item and Single-item Scales for Construct Measurement: a Predictive Validity Perspective. Journal of the Academy Marketing Science. DOI 10.1007/s11747-011-0300-3, Available at: http://www.springerlink.com/content/1742m74342561752/
10. Donabedian, A. (2003). An Introduction to Quality Assurance in Health Care. Oxford: Oxford University Press.
11. Eiriz, V. & Figueiredo, J.A. (2005). Quality evaluation in health care services based on customer-provider relationships. International Journal of Health Care Quality Assurance, vol.18, no.6, 404-412.
12. Grö nroos, Ch. (1984). A service quality model and its marketing implications. European Journal of Marketing, 1984, vol.18, 36-44.

13. Gupta, H.D. (2008). Identifying Health Care Quality Constituents: Service Providers' Perspective. Journal of Management Research, vol.8, no.1, 18-28.
14. Kang, G.D. (2006). The hierarchical structure of service quality: integration of technical and functional quality. Managing Service Quality, vol.16, no.1, 37-50.
15. Lee, H., Delene, M.L., Bunda, M.A., Kim, Ch. (2000). Methods of measuring Health-Care service quality. Journal of Business Research, vol.48, 233-246.
16. Parasuraman, A., Berry, L.L., Zeithaml, V.A. (1988). SERVQUAL: a multi-item scale for measuring consumer perceptions of the service quality. Journal of Retailing, vol.64, no.1, 12-40.
17. Piligrimienė, Ž. & Bučiūnienė, I. (2011). Exploring managerial and professional view to health care service quality. Economics and Management, no.16, 1304-1315. Available at: http://internet.ktu.lt/lt/mokslas/zurnalai/ekovad/16/1822-6515-2011-1304.pdf
18. Qin, H. (2009). Links among perceived service quality, patient satisfaction and behavioural intentions in the urgent care industry: empirical evidence from college students. (Doctoral dissertation). University of North Texas. Available at: http://digital.library.unt.edu/permalink/meta-dc-11014
19. Walbridge, S.W. & Delene, L.M. (1993). Measuring Physician Attitudes of Service Quality. Journal of Health Care Marketing, winter, 6-15.

Appendix 1. Health care quality dimensions, their description and number of items in questionnaire

Quality dimension	Description	Number of items	Authors
Tangibles	Physical evidence of service: facilities, number and appearance of personnel, tools or equipment used to provide service.	11	Parasuraman *et al.*, 1988; Haywood-Farmer & Stuart, 1988; Brown & Swartz, 1989; Walbridge & Delene, 1993; Lee *et al.*, 2000; Lim, Tang & Jackson, 1999; etc.
Accessibility	Ease of contact, waiting time, convenient location, etc. Ease with which health care services are reached. Access can be physical, financial and psychological.	11	Parasuraman *et al.*, 1985; Koch, 1991; Ware *et al.*, 1983; Maxwell, 1984, 1992; Bowers *et al.*, 1994; Mittal & Baldasare, 1996; Jun *et al.*, 1998; Rees, 1998; O'Brien, 1991; JCAHO, 1997.
Responsiveness	The degree to which patient is brought to the centre of health care service providing. Willingness of service providers to provide prompt service. It also involves the timeliness of service.	5	WHO, 2000; PZB, 1985, 1988.
Communication	Keeping patients informed in language they can understand, listening to them, education, etc.	6	PZB, 1985; Arnetz & Arnetz, 1996; Raper, 1996.
Reliability	The degree to which a promised service is performed dependably and accurately.	6	PZB, 1985; 1988;
Respect and Caring	Respect for patient values, preferences and needs. The degree to which the patient is involved in the decision-making process and to which services are provided with care and respect for his values and expectations.	8	Rees, 1998; JCAHO, 1997.
Competence	The degree to which health care personnel have the training and abilities to assess, treat and communicate with their clients.	7	OECD, 2006; Walbridge & Delene, 1993; Lee *et al.*, 2000; Haywood-Farmer & Stuart, 1988.
Effectiveness	The extent to which attainable improvements in health are, in fact, attained.	1	Donabedian, 1980; 2003; Arah *et al.*, 2003; WHO, 2000; Juran & Godfrey, 2000.
Appropriateness	The degree to which provided healthcare is relevant to the clinical needs, given the current best evidence.	4	OECD, 2006.
Continuity	The extent to which healthcare for specified users over time is coordinated across providers and institutions.	3	OECD, 2006; JCAHO, 1997; Jun *et al.*, 1998.
Safety	Minimization of various kind of risk for patients, personnel and environment.	5	PZB, 1985; WHO, 2006, etc.
Patient outcomes	Patients' opinion about their health outcomes.	4	Bowers *et al.*, 1994; Jun *et al.*, 1998; Mittal & Baldasare, 1996;
Efficiency	Optimal use of available resources to yield maximum benefits or results.	1	JCAHO, 1997; Donabedian, 2003.