Structural Modeling of Health Services' Quality Level as a Determinant of User Satisfaction of the Tertiary Health Care

Aldin Brajić[1] | *International University of Travnik, Travnik, Bosnia and Herzegovina*
Saliha Brajić[2] | *NGO EconVita, Zenica, Bosnia and Herzegovina*

Received 14.5.2024 (revision 6.9.2024), **Accepted** (reviewed) 24.10.2024, **Published** 14.3.2025

Abstract

Tertiary level of health care should provide highly specialized health services, that include the most complex methods and procedures of diagnosis, treatment and rehabilitation. The aim of the research is to examine the impact of the quality of health services on user satisfaction of medical services at the tertiary level of health care. For the purposes of this research, clinical centers of the tertiary health care level operating in the territory of Bosnia and Herzegovina were selected, which also represents the basic set of research. The survey was conducted on a sample of 1 022 users of health services provided by clinical centers in the territory of Bosnia and Herzegovina, where the cities represented the strata in the research: Sarajevo, Banja Luka, Tuzla, Mostar and Foča. The results indicate a strong influence of independent constructs on dependent constructs, that is, the quality of health services has a strong influence on the level of user satisfaction with (non)medical services.

Keywords	***DOI***	***JEL code***
Quality of health services, user satisfaction, tertiary level of health care, structural equation modeling	*https://doi.org/10.54694/stat.2024.26*	*C38, C83, I11, M31*

INTRODUCTION

Service quality and customer satisfaction are often used as indicators of competitiveness. However, their mutual relationship is relatively unclear. Namely, in some studies these two concepts were used as synonyms (Zeithaml, Berry and Parasuraman, 1993), while in other studies a distinction was made

[1] International University of Travnik, Aleja konzula, Meljanac bb, 72270, Travnik, Bosnia and Herzegovina. E-mail: brajicaldin@gmail.com, phone: (+387)66991393. ORCID: <https://orcid.org/0000-0002-3838-7373>.
[2] NGO EconVita, Zenica, Bosnia and Herzegovina. E-mail: saliha.cabro@yahoo.com.

between these concepts (Cronin and Taylor, 1992). Although today the dominant concepts of satisfaction (disconfirmation paradigm) and service quality (gap models) start from two different paradigms, both use expectations and perceptions as key determinants in their explanation, which is the reason why two comparative groups of satisfaction researchers have developed in the literature users and service quality researchers.

Regardless of the obvious overlaps of the previously mentioned concepts, some important differences were also established between them. Very often, the subject of controversy is the question of what precedes what, user satisfaction with service quality or quality precedes satisfaction? According to a growing number of authors, these two concepts are largely related. Of course, they differ from each other in the duration of the experience with the service, in the levels of expectations, the degree of affectivity and depending on stability of duration of the relationship between the user and service provider (Snoj, Savić and Rajtmajer, 1999). According to the integral approach, advocated by Klaus (Snoj, 1995), service quality could be understood as the value of the service for the user. It is thought to become more congruent with longer-term attitudes (Stafford, Stafford and Wells, 1998). On the other hand, satisfaction is considered more of an emotional reaction to the experience with a product or service, quite similar to the individual emotional state of mood deterioration that forms the basis of the level of regret. Cronin and Taylor (Oliver, 1993) conducted an empirical test of the reciprocity of satisfaction and quality in several service industries. Their research showed the fact that service quality can be seen as one of the determinants of user satisfaction, and satisfaction itself affects the user's future purchasing decisions. Oliver points out in his studies that satisfaction is the result of the user's overall reaction, and that it can have a potential impact on future perceptions of quality (Oliver, 1993). On the other hand, some authors in their research point out that satisfaction precedes the perceived quality of services (Bolton and Drew, 1991). However, regardless of the divided attitudes in the earlier period, today the prevailing opinion is that quality precedes satisfaction (Mikulić, 2007). Accordingly, this paper has the task of examining whether the quality of health services has an impact on the achieved level of satisfaction of users of tertiary level health care, both medical and non-medical services.

1 LITERATURE REVIEW

Among the first research efforts on the quality of services at the tertiary level of health care, the research of Alaloola and Albedaiwi (2008) who surveyed a total of 1983 users on the case study of the King Abdulaziz Medical Center in Riyadh stands out. A significant degree of satisfaction was expressed regarding the comfort of the hospital rooms (88.5%), the temperature of the rooms (78.1%), the emergency call of the staff (87.9%), the cleanliness of the rooms (79.6%) and respect from the staff (87.4%). On the other hand, dissatisfaction was expressed regarding the clarification of procedures (57.2%) and presentation by the doctor (59.1%).

A comparative study of satisfaction with the quality of health services at the tertiary level of health care by 383 users and 162 nurses, using the SERVQUAL survey questionnaire, was carried out by Nashrath, Akkadechanunt and Chontawan (2011) as a case study of a clinical hospital in the Maldives. The uniqueness of this research is reflected both in the consistency of the application of the quality dimensions according to SERVQUAL, and in the examination of the gap analysis between the two categories. The results of the research are also extremely interesting, since they show a certain overlap in the perceptions of users and nurses regarding the poorly rated quality dimensions, especially regarding the "identification" dimension. The authors believe that feedback from nurses can be an excellent input for creating a quality management system.

Lee and Kim (2017) attempted a comparative analysis of health service quality measurements in a clinical hospital in Seoul, on two sample groups: current users and their family members, and former users and the general public. It is a study on two different samples of different sizes – the first sample

of 365 and the second sample of 232 participants. The survey is based on the HEALTHQUAL questionnaire with dimensions: empathy, safety, tangibility, efficiency and improvement of health services. Testing was performed through t-test, ANOVA and confirmatory factor analysis (SEM). The results of the research showed that there is no statistically significant difference between the presented samples, and it is recommended to act simultaneously on all five dimensions of the quality of health services.

In their research on the evaluation of the quality of health services at the tertiary level of health care in India, Natarajappa and a group of authors (2020) distinguished 13 dimensions of the quality of health services, namely: 1) reception, 2) social responsibility, 3) staff behavior, 4) service quality and service availability, 5) confidence, 6) continuity, 7) communication, 8) environment, 9) treatment costs, 10) customer loyalty, 11) hospital discharge, 12) medical services, and 13) overall services. The presented dimensions are mainly derived by the authors from the SERVQUAL questionnaire, creating their own model of quality management at the tertiary level of health care. On a sample of 30 users (patients), the authors come to the conclusion that the highest level of correlation of the dimension "user loyalty" was achieved with the dimensions "overall service" and "discharge from the hospital", although a positive correlation is also noticeable in other dimensions.

In a sample of 410 tertiary level health care users in Bangladesh, Dilshad et al. (2020) confirms the existence of a growing concern on the part of users regarding the level of quality of health services of public clinical centers. The overall level of satisfaction with all services, technical equipment and interpersonal relations was at an extremely low level.

We can see a somewhat different approach to examining the satisfaction of users of health services in a study of a tertiary clinical hospital in Karachi (Pakistan). The researchers, on their own created a questionnaire, examined the satisfaction of a total of 173 users, excluding users of maternity, psychiatry and chemotherapy departments. What is extremely interesting in this research is the fact that satisfaction was measured by the frequency of problems that occurred in certain departments. The less often the problems happened, the more satisfied the users were and vice versa (Imam et al. 2007).

Garg et al. (2014) conducted a user satisfaction survey of a tertiary specialist hospital in India. The research was carried out by surveying users who were hospitalized during the two months of monitoring. For the survey questionnaire, they used the previously created Canadian questionnaire for measuring satisfaction with medical services (The Northwest Territories Hospital satisfaction questionnaire). More than 88% of users rated the services as excellent and good. The areas in which dissatisfaction was noted were the cleanliness, especially in the toilets, and the quality of the food served to the users. Also, the research results point to the need to develop the soft-skills of the medical staff, in order to better understand the users.

In South Korea, a more extensive study was conducted in 29 outpatient clinics of the University Clinical Center in Seoul with a sample of 1 194 users. The authors made a significant effort in creating measurement scales based on the theoretical framework of Donabedian and the National User Satisfaction Index questionnaire. The research makes a clear distinction between health service quality and satisfaction. The shortcoming of the research is that the quality of the health service was developed in 5 constructs (doctor's examination, services of nurses/technicians, technical services, amenities and physical surroundings), while satisfaction was measured exclusively by one question: "Are you satisfied with the overall health service of this institution (Ham, et al. 2015)?"

Kulkarni (2018) developed his own scale for measuring the satisfaction of users of the tertiary level of health care, inspired by the HCAHPS survey questionnaire. The scale was tested on a sample of 100 randomly selected users of a tertiary clinical hospital in Maharashtra (India). The results of the descriptive analysis showed that the users were satisfied with the availability of services, the professionalism

of care, the waiting time for an examination, the behavior of consultants, medical and other staff. Overall satisfaction level of 73% was excellent or good and 22% average. Dissatisfaction was expressed mostly with regard to toilets and drinking water.

Another study on the satisfaction of users of tertiary care in Pakistan was carried out by Maroof et al. (2019). Like their predecessors, this team created its own satisfaction measurement scale. The questionnaire consisted of 38 questions, and the statements used to measure satisfaction were formulated with a negative sign. Using a Likert scale from 1 – completely disagree, to 5 – completely agree. A total of 110 users gave their ratings. The results showed that there is significant user dissatisfaction and that their needs are not being met at an adequate level.

A very interesting survey was conducted on a sample of as many as 136 hospitals of the tertiary level of health care in China, which showed that the users were mostly satisfied with the level of services provided. Through statistical analysis, 12 variables were singled out: 1) Medical skill of the doctor, 2) Inquiry about the medical history and current situation of the user, 3) Convenience of using the elevator in the hospital, 4) Feeling of respect from the medical staff, 5) Timely instructions from the staff, 6) Explanation of treatment and medication, 7) Waiting time before consultation, 8) Waiting time for medical examination, 9) Privacy protection, 10) Waiting time for bill payment, 11) Bathroom cleanliness, and 12) Drinking water supply in waiting areas. The identified variables were then allocated into four categories: 1) Waiting time, 2) Service and treatment, 3) Costs and 4) Environment. The research showed that Chinese users were most satisfied with the category "Service and treatment", and somewhat less with "Waiting time", "Costs" and "Environment", for which a greater degree of investment by the management of clinical centers is recommended (Hu, 2019.). By the way, this is one of the rare ones among hospital researches with the author's own measurement scales.

By using Gaps Model of Service Quality and the SERVQUAL instrument Ozretić et al. (2020) performed an analysis of deviations in the perception and expectations of users of university clinics regarding the quality of health services. Although the data was collected for each of the 18 departments, it was established that there was a significant deviation of the variables at the level of the entire clinical center. The largest gaps were indentified "responsiveness" and "tangibility".

Ojeniweh et al. (2021) investigate the satisfaction of 141 users of tertiary level health care in Nigeria. The highest degree of satisfaction was achieved with communication, technical equipment and interpersonal aspects of the health service, contrary to the previously mentioned research. On the other hand, dissatisfaction was expressed in terms of costs and waiting time for service provision.

2 RESEARCH METHODOLOGY

Within the elaboration of certain theoretical and methodological origins of the observed problem, and certain applied considerations, the work used: hypothetical-deductive method, method of induction and deduction, method of analysis and synthesis and statistical methods (descriptive and inferential statistics) with a systematic approach to research.

2.1 Sample and data collection

Data collection was carried out through primary research among University-Clinical Centers in Bosnia and Herzegovina. In the process of empirical research, a survey questionnaire was used as a method for data collection. Content validity is ensured by the use of validated measuring instruments, and by consulting a group of experts when formulating, translating and adjusting the measurement scales. The questionnaire used to measure the various variables in the predicted models consists of a series of questions. The questions were chosen based on a systematic review of the literature, the subject of which was the quality of health services and user satisfaction.

The survey questionnaire also contains a group of questions related to the sociodemographic characteristics of the respondents. The questions were mostly taken over and adapted from earlier validated empirical research.

Content and nomological validity was conducted during and after operationalization of measurement models. All indicators are thoroughly checked for wording, specificity and sentence length to ensure relevance to the context in which the research is conducted. A panel of experts from the academic community in Bosnia and Herzegovina checked the content validity and relevance of the measuring instrument. They received the questionnaires via e-mail, and were able to fill out the questionnaire, as well as send written comments to the indicators of the measured constructs. Minor changes, i.e. rewording, were made to several indicators based on comments received from panel experts.

The invitation to participate in the research was distributed by registered mail to the addresses of all clinical centers, as well as by e-mail. The survey was conducted using a combination of field research and online via Google Forms. Field data collection was carried out by expert and trained persons who had the necessary information in case of ambiguities of respondents. The survey questionnaires that were submitted online contained additional information for each question, and contact information was provided in case of additional questions and ambiguities from respondents. Also, in order to avoid missing data in Google survey questionnaires, the questions were arranged as mandatory, that is, they could not be skipped. Data collection was carried out in the period from July 2021 to February 2022. For the purposes of the research, clinical centers of the tertiary level of health care operating in the territory of Bosnia and Herzegovina were selected, which also represents the basic set of research. Data collection was carried out on the basis of a stratified sample, since it belongs to the category of random samples and allows to evaluate the degree of reliability of drawing conclusions about the investigated parameters. The survey was conducted on a sample of 1 022 users of health services of clinical centers in Bosnia and Herzegovina with an extremely high return rate (70%), with the cities representing the strata in the survey: Sarajevo, Banja Luka, Tuzla, Mostar and Foča. In Bosnia and Herzegovina there are the following tertiary health care institutions: University Clinical Center Sarajevo; University Clinical Center of Republika Srpska Banja Luka; University Clinical Center Tuzla; University Clinical Hospital Mostar and University Hospital Foča. Individual filling in of questionnaires took 7 minutes. All statements of the constructs were measured with a Likert scale ranging from 1 – "completely disagree" to 5 – "completely agree". The second group of questions related to the sociodemographic characteristics of the respondents.

2.2 Structural model and hypotheses

The basic structural model consists of the independent variable "quality of health services" and the dependent variable "user satisfaction", which are connected to the following hypotheses:

H1: There is a statistically significant influence of the quality of health services on the level of user satisfaction with medical services in Bosnia and Herzegovina.

H1a: Tangibility has a statistically significant effect on the level of user satisfaction with medical services.

H1b: Reliability has a statistically significant effect on the level of user satisfaction with medical services.

H1c: Response has a statistically significant effect on the level of user satisfaction with medical services.

H1d: Safety has a statistically significant effect on the level of user satisfaction with medical services.

H1e: Empathy has a statistically significant effect on the level of user satisfaction with medical services.

H2: There is a statistically significant influence of the quality of health services on the level of user satisfaction with non-medical services in Bosnia and Herzegovina.

H2a: Tangibility has a statistically significant effect on the level of user satisfaction with non-medical services.

H2b: Reliability has a statistically significant effect on the level of user satisfaction with non-medical services.
H2c: Response has a statistically significant effect on the level of user satisfaction with non-medical services.
H2d: Safety has a statistically significant effect on the level of user satisfaction with non-medical services.
H2e: Empathy has a statistically significant effect on the level of user satisfaction with non-medical services.

Quality of health services (QHS) is a second-order construct consisting of a total of 23 statements, and it consists of five first-order constructs, of which five statements measure the construct tangibility, six statements measure the construct reliability, three statements measure the construct response, four statements measure the construct safety and five statements measure the construct empathy.

Table 1 Instruments for measuring the quality of health services

Dim.	Subdimension	Code	Indicators (assertions)
QHS	Tangibility	Tangibility_1	The health institution has modern equipment.
		Tangibility_2	The exterior and interior of the healthcare facility is visually acceptable.
		Tangibility_3	The employees of the health care facility look neat.
		Tangibility_4	The accessories and devices of the healthcare facility are clean.
		Tangibility_5	The health institution has equipment and facilities in accordance with the services it provides.
	Reliability	Reliability_1	In the health facility, examinations, treatments and treatment services are quick and precise.
		Reliability_2	User review schedule is on time.
		Reliability_3	The service procedure is performed correctly on the first attempt.
		Reliability_4	Ease of contacting hospital staff.
		Reliability_5	The health institution insists on providing a health service without errors.
		Reliability_6	The employees of the health care facility have the knowledge to respond to the user's inquiry.
	Response	Response_1	Employees of the health care facility warn when the user needs help.
		Response_2	User complaints are resolved successfully and promptly.
		Response_3	Employees of the health care facility provide clear and understandable information.
	Safety	Safety _1	Sufficient attention is paid to the user.
		Safety_2	Employees of the health facility are available when needed by the user.
		Safety_3	The employees of the healthcare facility are capable of analyzing the user's illness.
		Safety_4	The medical staff accurately and precisely treats the user's ailments.
	Empathy	Empathy_1	Employees show extreme patience in dealing with users.
		Empathy_2	The employees are friendly and hospitable.
		Empathy_3	Users can easily file complaints.
		Empathy_4	Moral support is provided to users.
		Empathy_5	Services are provided to all users regardless of social status.

Source: Own construction

User satisfaction (US) is a second-order construct consisting of a total of 30 statements, and it consists of two first-order constructs: user satisfaction with medical services, which is measured by 15 statements, and user satisfaction with non-medical services, which is measured by 15 statements.

Table 2 Instruments for measuring user satisfaction

Dim.	Subdimension	Code	Indicators (assertions)
US	User satisfaction with medical services	USMS1	I am satisfied with the reception upon arrival at the health facility.
		USMS2	I am satisfied with the presentation of doctors and nurses/technicians in the health institution.
		USMS3	I am satisfied with the clarity of the information provided about upcoming procedures and interventions by doctors and nurses/technicians.
		USMS4	I am satisfied with the length of the conversations that the doctors and nurses/technicians spent with me.
		USMS5	I am satisfied with the professional approach of the doctors and nurses/technicians in the health facility.
		USMS6	I am satisfied with the doctor's ability to diagnose the health problem.
		USMS7	I am satisfied with the expediency (quickly provided services and no waiting).
		USMS8	I am satisfied with the explanation for the delay in the ordered examination.
		USMS9	I am satisfied with the success of the treatment.
		USMS10	Satisfied with the treatment process.
		USMS11	The doctors and nurses/techs did everything possible to ease my pain.
		USMS12	Before I receive the medicine, the doctors and nurses/technicians explain what it is for and the possible side effects.
		USMS13	The doctors and nurses/technicians have provided written information about the symptoms I have or recommendations that I must follow after I leave the healthcare facility.
		USMS14	After leaving the health care facility, I understand my health condition and the procedures I am responsible for implementing for the benefit of my health.
		USMS15	Every experience I had with a healthcare facility has met my expectations in terms of medical services.
	User satisfaction with non-medical services	USNMS1	I am satisfied with the resolution of the complaint.
		USNMS2	I am satisfied with the hospital environment.
		USNMS3	I am satisfied with the accommodation services.
		USNMS4	I am satisfied with the food services.
		USNMS5	I am satisfied with the prices of health services.
		USNMS6	I am satisfied with the low level of corruption in the institution where I stayed.
		USNMS7	I am satisfied with the clarity of information provided about upcoming procedures by non-medical staff.
		USNMS8	I am satisfied with the application of information technologies (e-cards, e-orders, e-prescriptions, etc.) in the institution where I stayed.
		USNMS9	I am familiar with the rights arising from compulsory health insurance, and they refer to the right to use health care and the right to certain financial benefits and assistance.
		USNMS10	I am familiar with the rights arising from extended health insurance.
		USNMS11	I am familiar with the rights arising from voluntary health insurance for myself and my family.
		USNMS12	I am satisfied with the method of financing health institutions and health services.
		USNMS13	I am satisfied with the payment of medicines, surcharges, co-payments, additional payments, etc.
		USNMS14	I am satisfied with the friendliness of the staff.
		USNMS15	Every experience had with a healthcare facility has met my expectations, from the aspect of non-medical services.

Source: Own construction

For structural equation modeling analysis we used IBM SPSS AMOS 21.0. The steps for structural equation modeling (SEM) analysis are: (1) Descriptive statistical analysis, (2) Data testing and verification, (3) Assessment of model fit, (4) Model reliability testing, (5) Model validity testing, and (6) Assessment of structural relationships / hypothesis testing. All the mentioned methodological steps of testing are carried out on the sample itself, since the shape of the structural model that is ultimately tested depends on the character of the sample. Therefore, each individual step will be explained in chapter 3 Results.

Finally, the greatest contribution of this research is the simultaneous examination of user satisfaction with health services for all five functioning health institutions of the tertiary level of health care: University Clinical Center Sarajevo, University Clinical Center of Republika Srpska Banja Luka, University Clinical Center Tuzla, University Clinical Hospital Mostar and University Hospital Foča. In this way, a higher level of research objectivity was ensured and the making of generalizing conclusions at the level of Bosnia and Herzegovina, making this research endeavor unique in these areas.

3 RESULTS

3.1 Descriptive statistical analysis

Based on the descriptive analysis, we concluded that among the respondents, women dominated, which make up 58.6% of the sample, while men make up the remaining 41.4%. Looking at the age structure, the largest number of respondents are between the ages of 21 and 29, 253 of them, or 24.8%, followed by the age group between 30 and 39, relatively 23.5%, that is, 240 in absolute terms. Furthermore, we have 191 respondents between the ages of 40 and 49, i.e. 18.7%, while the least number of respondents are over 60, 97 of them, i.e. 9.5%. The majority of respondents have completed secondary vocational education (45%) and belong to the category of workers (37.2%).

The arithmetic mean of all indicators for the independent construct *"quality of healthcare services"* ranges from 2.80 to 3.78 (on a scale of 1 to 5, 1 – completely disagree, 5 – completely agree). The mode and median are 3, and the standard deviations range between 1.138 and 1.398. The arithmetic mean of all indicators of the dependent construct *"user satisfaction"* ranges from 2.64 to 3.51. The mode and median are 3, and the standard deviations range between 1.209 and 1.407.

3.2 Data testing and verification

Multivariate analysis was performed using the *Mahalanobis Distance test*, the calculation of which enables the identification of outliers through an approximate test of statistical significance. After converting the *Mahalanobis Distance test* to its probabilities, we eliminated all observations that are less than or equal to 0.001 ($p \leq 0.001$). The total number of outliers among users of health services was 136, leaving 886 respondents in the sample.

Based on the results of the *Kolmogorov-Smirnov* and *Shapiro-Wilk tests,* $p < 0.001$ we concluded that the null hypothesis is rejected, that is, that the assumption of normal distribution is not satisfied. However, consulting the literature resulted in the conclusion that data deviations from the assumption of normality do not represent a problem, if the analysis uses the Maximum Likelihood (ML) estimation method. Namely, Nwabueze et al. (2009) confirmed in their study that the ML method is robust in five different distributions. A similar conclusion was reached by Fuller and Hemmerle (1966), confirming the robustness of the ML technique in six distributions, including the one in which the assumption of normality of the data was violated. In order to test the data for the presence of homoscedasticity, the Breusch-Pagan test was applied, where the null hypothesis assumes the presence of homoscedasticity. Based on the obtained results, we concluded that the statistical significance of the χ^2 test is < 0.001. Therefore, the results indicate a violation of the assumption of homoscedasticity.

The analysis of the Scatter plot concluded that the data are linearly distributed, thereby fulfilling the assumption of data linearity. Potential multicollinearity was examined using the tolerance index

(TOL) and the variance inflation factor (VIF). For all variables, TOL values are greater than 0.1, and all VIF values are less than 10. Previously, it led to the conclusion that no variable causes the problem of multicollinearity in the research.

3.3 Assessment of model fit

In order to examine the suitability of the model, a confirmatory factor analysis of the measured models of the quality of health services (QHS) and user satisfaction (US) was carried out. However, the obtained results did not show satisfactory suitability of the model. Accordingly, the model was respecified. After analyzing the modification index, it was determined that some statements have too high a correlation, that is, they measure almost the same concept. By reviewing the standardized factor loadings, it was observed that all the values of the manifest variables are above the recommended value of 0.5, and therefore we kept all the variables of the observed models. Based on the above, the CFA analysis was repeated, and based on the obtained results, we concluded that the models achieved a good level of suitability. Namely, the GOF indicators are above/below the recommended limit values.

Table 3 Evaluation of the suitability of the measurement model (GOF), the quality of health services (QHS) and the satisfaction of users of health services (US)

Measures	Threshold value	QHS construct	US construct
P-value	> 0.05	0.001	0.001
Minimum Discrepancy Function by degrees of freedom divided (CMIN/df)	< 5	4.522	4.815
Root mean square error of approximation (RMSEA)	< 0.08	0.063	0.066
Standardized root mean squared residual (SRMR)	< 0.09	0.0248	0.0345
Comparative Fit Index (CFI)	> 0.90	0.967	0.953
Normed Fit Index (NFI)	> 0.90	0.957	0.942
Relative Fit Index (RFI)	> 0.90	0.95	0.934
Incremental Fit Index (IFI)	> 0.90	0.967	0.953
Tucker Lewis – non-normed fit index (TLI – NNFI)	> 0.90	0.961	0.947
Goodness of Fit Index (GFI)	> 0.90	0.902	0.869
Adjusted Goodness of Fit Index (AGFI)	> 0.80	0.874	0.843
Parsimonious Normed Fit Index (PNFI)	> 0.50	0.814	0.838
Parsimonious Comparative Fit Index (PCFI)	> 0.50	0.821	0.848

Source: Own construction

Based on the data from the Table 3, we can conclude that the absolute, parsimonious and incremental indicators are acceptable and therefore suitable for the QHS construct. Chi square coefficient (CMIN) is significant and is 972 274 with 215 degrees of freedom (df), while CMIN/df = 4 522. Considering the sample of 886 observations and 23 manifest variables, it is expected that the chi square test will be significant. Absolute suitability indicators include the value of the square root of the standard error of assessment (RMSEA), which should be less than 0.08, which is 0.063, as well as the standardized root

mean square (SRMR), whose value should be less than 0.09. and is 0.0248, which additionally confirms the suitability of the measurement model. Parsimony indicators such as AGFI = 0.874 and PNFI = 0.814 are above the recommended values and also indicate a good fit of the model. The values of the incremental indicators – the standardized fit index NFI = 0.957 and the comparative fit index CFI = 0.967 are above the limit of 0.9, and indicate a satisfactory fit of the model.

We can also conclude that absolute, parsimonious and incremental indicators are acceptable and therefore suitable for the US construct. Chi square coefficient (CMIN) is significant and is 1 863 353 with 387 degrees of freedom (df), while CMIN/df = 4 815. Considering the sample of 886 observations and 30 manifest variables, it is expected that the chi square test will be significant. Absolute suitability indicators include the value of the square root of the standard error of assessment (RMSEA), which should be less than 0.08, which is 0.066, as well as the standardized root mean square (SRMR), whose value should be less than 0.09. and is 0.0345, which additionally confirms the suitability of the measurement model. Parsimony indicators such as AGFI = 0.843 and PNFI = 0.838 are above the recommended values and also indicate a good fit of the model. The values of the incremental indicators – the standardized fit index NFI = 0.942 and the comparative fit index CFI = 0.953 are above the limit of 0.9 and indicate a satisfactory fit of the model.

While evaluating the suitability of the model, it is always necessary to display the chi-square value with the associated p-value. However, the sensitivity of the chi-square test to the sample size is high, given that its value increases along with the sample size. Higher chi-square values are associated with lower p-values, which indicate a statistically significant result, that is, poor model fit. This may encourage the use of smaller samples, which may be counterproductive, as they may only ostensibly mask poorer fit indicators and provide less precise parameter estimates. Also, when the variables do not satisfy multivariate normality and/or when the sample is small, then the test size is unlikely to follow a chi-square distribution. With such models, depending on the method and degree of deviation from normality, wrong conclusions are more common, so the chi-square can be overestimated and thus show the model worse than it is in reality, or else the chi-square can be underestimated and show the model better than it is. In addition, the chi-square increases with the increase in the number of manifest (observed) variables (indicators), i.e. it also depends on the complexity of the model. Therefore, more complex models with more parameters will show a smaller chi-square value, due to the reduction in the number of degrees of freedom.

Therefore, the acceptable value of this indicator can be the result of adding free parameters to the model, and not exclusively the result of a correctly specified model. Due to the limitations of the chi-square test with its associated p-value, it is recommended to use additional indicators of model suitability during the analysis, instead of relying on only one indicator. It is useful to combine the so-called goodness-of-fit and badness-of-fit measures, where higher values in the former indicate better suitability, and worse in the latter. According to Sharif and Sharif (2018) for the model to be eligible it is necessary to show:

- the results of the chi-square test, even if the p-value is significant;
- at least three incremental suitability indicators, the values of which are greater than 0.90 (CFI, NFI, TLI, etc.);
- RMSEA although the values are not good, and the acceptable value of RMSEA is < 0.08;
- SRMR whose value should be less than 0.08.

3.4 Model reliability testing

To test the reliability of measuring models, the quality of health services and the satisfaction of users of health services, the following were used: Cronbach alpha coefficient, Standardized Cronbach alpha coefficient, compozite reliability (CR) and maximal reliability [MaxR(H)]. Based on the obtained results, we established that all indicators in the models consistently represent the corresponding latent construct or factor, that is, that the measurement models are reliable.

Table 4 CFA results for the measurement model of health service quality

Dim.	Subdimension	Code	Factor loads	Cronbach alpha	Standardized Cronbach alpha	CR	MaxR(H)
QHS	Tangibility	Tangibility_1	0.791	0.911	0.912	0.901	0.907
		Tangibility_2	0.776				
		Tangibility_3	0.761				
		Tangibility_4	0.803				
		Tangibility_5	0.879				
	Reliability	Reliability_1	0.829	0.940	0.940	0.936	0.938
		Reliability_2	0.815				
		Reliability_3	0.858				
		Reliability_4	0.834				
		Reliability_5	0.886				
		Reliability_6	0.836				
	Response	Response_1	0.864	0.900	0.900	0.901	0.901
		Response_2	0.854				
		Response_3	0.882				
	Safety	Safety _1	0.898	0.934	0.934	0.930	0.933
		Safety_2	0.901				
		Safety_3	0.839				
		Safety_4	0.867				
	Empathy	Empathy_1	0.897	0.939	0.939	0.940	0.945
		Empathy_2	0.911				
		Empathy_3	0.815				
		Empathy_4	0.896				
		Empathy_5	0.829				

Source: Own construction

Cronbach's alpha coefficient was used to show the reliability of the "quality of health services (QHS)" scale. Cronbach's alpha coefficient for the first-order latent variables tangibility, reliability, response, safety and empathy ranges from 0.900 to 0.940, which shows excellent reliability and internal consistency for this sample. In addition, the value of the "Standardized Cronbach coefficient alpha", which also ranges from 0.900 to 0.940, further confirms the reliability of the scale. Therefore, it can be concluded that all indicators in the model consistently represent the appropriate latent construct or factor, i.e. that the measurement model "quality of healthcare services" is reliable.

Based on the data presented in Table 5, we can draw the same conclusions regarding the reliability of the US model. Since Cronbach's alpha coefficient and Standardized Cronbach coefficient alpha for User satisfaction with medical services are 0.976 and for User satisfaction with non-medical services are 0.964, it can be concluded that all indicators in the model consistently represent the appropriate latent construct or factor, i.e. that the measurement model "user satisfaction" is reliable.

Table 5 CFA results for the measurement model of user satisfaction

Dim.	Subdimension	Code	Factor loads	Cronbach alpha	Standardized Cronbach alpha	CR	MaxR(H)
US	User satisfaction with medical services	USMS1	0.835	0.976	0.976	0.976	0.977
		USMS2	0.863				
		USMS3	0.879				
		USMS4	0.864				
		USMS5	0.900				
		USMS6	0.870				
		USMS7	0.847				
		USMS8	0.799				
		USMS9	0.865				
		USMS10	0.889				
		USMS11	0.878				
		USMS12	0.809				
		USMS13	0.796				
		USMS14	0.846				
		USMS15	0.886				
	User satisfaction with non-medical services	USNMS1	0.866	0.964	0.964	0.962	0.967
		USNMS2	0.846				
		USNMS3	0.829				
		USNMS4	0.808				
		USNMS5	0.798				
		USNMS6	0.819				
		USNMS7	0.874				
		USNMS8	0.780				
		USNMS9	0.684				
		USNMS10	0.672				
		USNMS11	0.658				
		USNMS12	0.752				
		USNMS13	0.729				
		USNMS14	0.846				
		USNMS15	0.881				

Source: Own construction

3.5 Model validity testing

To determine the convergent validity of the factors in the models, indicators of standardized factor loading and average variance extracted (AVE) were calculated, and finally we compared the values of CR and AVE. Based on Tables 4 and 5, we can state that all factor loadings are statistically significant, at a significance level of 1%. The standardized factor loadings of all indicators are greater than 0.5, which indicates the fact that they well reflect the latent variable they measure. In Table 6 we compared the values of CR and AVE.

Table 6 Results of convergent validity of measurement models of quality of health services and user satisfaction

Dimension	Subdimension	CR	AVE	CR > AVE
QHS	Tangibility	0.901	0.645	Fulfilled
	Reliability	0.936	0.711	Fulfilled
	Response	0.901	0.751	Fulfilled
	Safety	0.930	0.768	Fulfilled
	Empathy	0.940	0.758	Fulfilled
US	User satisfaction with medical services	0.976	0.732	Fulfilled
	User satisfaction with non-medical services	0.962	0.628	Fulfilled

Source: Own construction

Based on the data from the Table 6, we can see that all AVE values are above the recommended threshold value of 0.5 (from 0.628 to 0.768), which indicates that each construct in the model explains at least 50% of the variance in its indicators. The highest explanation of the variance is among the indicators of the latent variable of security (0.768, i.e. 76.8%). AVE is a more conservative assessment of the validity of the measurement model, and it is considered that based on the CR indicator, a conclusion can also be made about the convergent validity of the model. After all, this indicator also indicates, as we previously stated, the reliability of the construct, where the lower limit of good reliability is the value 0.7, which is satisfied in this model (from 0.901 to 0.976). Since all values of CR > AVE, it can be concluded that the measurement model meets the conditions of convergent validity. According to all analyzed indicators of validity and reliability, it was concluded that the selected indicators explain well the factor (latent variable) they represent. Discriminant validity was tested using Fornell-Larcker criteria, HTMT and HTMT2.

Table 7 Results of discriminant validity of measurement models of quality of health services and satisfaction of users of health services using Fornell-Larcker criteria

Fornell-Larcker

Latent variables	Tangibility	Reliability	Response	Security	Empathy
Tangibility	0.803				
Reliability	0.883***	0.843			
Response	0.852***	0.983***	0.867		
Security	0.848***	0.971***	1.000***	0.877	
Empathy	0.819***	0.911***	0.960***	0.956***	0.871

Fornell-Larcker

Latent variables	User satisfaction with medical services	User satisfaction with non-medical services
User satisfaction with medical services	0.856	
User satisfaction with non-medical services	0.890***	0.793

Source: Own construction

Table 7 shows the values of the square root of the AVE index, while below are the correlations between the constructs, on the basis of which discriminant validity testing can be performed. Testing is done by comparing the square root of the AVE index value with the correlation of the given construct with all other constructs. Based on the presented results, we can conclude that the discriminant validity of the measuring model of the QHS and US is impaired, given that many values on the diagonal are smaller than the correlation coefficients in the relevant rows and columns.

As the discriminant validity according to the Fornell-Larcker criterion was violated, it was retested using the HTMT and HTMT2. Namely, in the simulation study conducted by Henseler, Ringle and Sarstedt (2015) it was shown that the traditional criterion for determining discriminant validity was not effective in finding the problem of discriminant validity when it really exists. The Fornell-Larcker criterion successfully identified 139 problems in only 15% of cases. For this reason, a new measure of discriminant validity was proposed, called the ***heterotrait-monotrait ratio*** (HTMT). After calculating the discriminant validity using HTMT and HTMT2, we can say that it is still impaired, because the ratio of individual constructs hovered around the value of 1 (from 0.80 to 0.99). Of all the latent constructs in the QHS, Response achieved the highest values (0.97 and 0.99). In this case, the literature suggests that the construct should be removed and its indicators should also be removed or joined to other constructs, and that the testing should be repeated. Respecting the content of the indicator, we attached Response_1 and Response_3 to the latent construct Safety, and we attached Response_2 to the construct Empathy.

Table 8 Results of discriminant validity of measurement models of quality of health services and satisfaction of users of health services using HTMT and HTMT2

HTMT				
Latent variables	**Tangibility**	**Reliability**	**Security**	**Empathy**
Tangibility				
Reliability	0.84			
Security	0.84	0.92		
Empathy	0.80	0.90	0.91	
HTMT2				
Latent variables	**Tangibility**	**Reliability**	**Security**	**Empathy**
Tangibility				
Reliability	0.84			
Security	0.84	0.92		
Empathy	0.80	0.90	0.91	
HTMT				
Latent variables	**User satisfaction with medical services**		**User satisfaction with non-medical services**	
User satisfaction with medical services				
User satisfaction with non-medical services	0,87			
HTMT 2				
Latent variables	**User satisfaction with medical services**		**User satisfaction with non-medical services**	
User satisfaction with medical services				
User satisfaction with non-medical services	0,87			

Source: Own construction

The next step was to find the manifest variables of one latent construct that are highly correlated with the manifest variables of other latent constructs. Based on a detailed analysis, the existence of a high correlation (correlations above 0.7) with the manifest variables of other latent constructs was identified for the following manifest variables: Safety_2, Empathy_2, Security_1, Reliability_5, Response_2 which was added to latent construct Empathy and Response_3 which was added to latent construct Safety, and therefore they were eliminated. Within each latent construct, the condition of having at least three manifest variables is met.

After the transformation, we repeated the testing, and the final results are shown in the Table 8. The results showed that values of HTMT and HTMT2 are around the threshold value of 0.90, indicating the absence of discriminant validity problems in the models, with the QHS measurement model retaining four latent constructs of the first order: tangibility, reliability, security and empathy. Namely, smaller values of HTMT and HTMT2 show that the correlations between indicators measuring different constructs are smaller compared to the correlations between indicators measuring the same construct. Precisely from this comes the conclusion that the discriminant validity in the models has been confirmed, that is, it has been established that the constructs are mutually different and that their associated indicators measure them well. In the case of US, the aforementioned problem did not manifest itself, and therefore no transformation was performed.

3.6 Assessment of structural relationships/hypothesis testing

Based on the conducted analyses, we concluded that all measurement models met the assumed criteria of suitability, reliability and validity, and as such were the subject of analysis and hypothesis testing. The basic work model is a recursive structural model, in which all paths go from predictor to dependent variables, which means that no two-way relationships are defined. The basic structural model now consists of four exogenous (independent) latent constructs: tangibility, reliability, security and empathy (the result we obtained based on discriminant validity). Furthermore, the structural model is made up of two endogenous (dependent) latent constructs: the level of user satisfaction with medical services and the level of user satisfaction with non-medical services. The effects of exogenous on endogenous latent constructs are defined by hypotheses. Looking at the basic structural model of the subject research, it can be concluded that the model consisted of 2 latent constructs of the second order with 6 latent constructs of the first order and a total of 47 manifest variables. The total number of parameters is 1 128 (based on the formula: $p \cdot (p + 1)/2$), the number of parameters to be calculated is 131, while the number of degrees of freedom is 997 (1 128 – 131). Assessment of the suitability of the basic model was carried out using the GOF indicator. GOF indicators are above/below the recommended limit values.

Based on the tabular presentation, we can conclude that the absolute, parsimonious and incremental indicators are acceptable and thus we have confirmed the suitability of the basic structural model. Chi square coefficient (CMIN), as an absolute indicator, is significant and amounts to 3 860 119 with 997 degrees of freedom (df), while CMIN/df = 3 872. Considering the sample of 886 observations and 47 manifest variables, it is expected that the chi square test will be significant. The values of RMSEA, which should be less than 0.08 and equal to 0.057, and SRMR, whose value should be less than 0.09 and equal to 0.0324, are classified as absolute indicators of suitability, which additionally confirms the suitability of the measurement model. Parsimony indicators such as AGFI = 0.797, PNFI = 0.851 and PCFI = 0.868 are above the recommended values and also indicate a good fit of the model. The values of incremental indicators – NFI = 0.923, CFI = 0.941, TLI – NNFI = 0.936, RFI = 0.916 and IFI = 0.941 are above the limit of 0.9, and indicate a satisfactory fit of the model. Based on the presented results of the suitability of the model, it is possible to conclude that the proposed model of the influence of the quality of health services on the level of user satisfaction with (non)medical services meets the methodological requirements.

Table 9 GOF fit index of the structural model

Measures	Threshold value	Structural model
P-value	> 0.05	0.001
Minimum Discrepancy function by degrees of freedom divided (CMIN/df)	< 5	3.872
Root mean square error of approximation (RMSEA)	< 0.08	0.057
Standardized root mean squared residual (SRMR)	< 0.09	0.0324
Comparative Fit Index (CFI)	> 0.90	0.941
Normed Fit Index (NFI)	> 0.90	0.923
Relative Fit Index (RFI)	> 0.90	0.916
Incremental Fit Index (IFI)	> 0.90	0.941
Tucker Lewis – Non-Normed Fit Index (TLI – NNFI)	> 0.90	0.936
Goodness of Fit Index (GFI)	> 0.90	0.82
Adjusted Goodness of Fit Index (AGFI)	> 0.80	0.797
Parsimonious Normed Fit Index (PNFI)	> 0.50	0.851
Parsimonious Comparative Fit Index (PCFI)	> 0.50	0.868

Source: Own construction

Given that the overall suitability of the model is acceptable, it is possible to approach the analysis of the structural part of the model, with the aim of assessing whether the proposed theoretical relations, that is, hypotheses are supported in the specific research context of Bosnia and Herzegovina. In the analysis, it is necessary to check the signs of the parameters, and the statistical significance of the parameters measured by the t-value. The results of the path analysis, for the H1 and H2 hypothesis, are presented in the following Table 10.

Table 10 Presentation of the results of testing the basic structural model

Hypotheses	Non-standard. evaluation parameters	Standard. evaluation parameters	t	P	R^2	Cohen's *f*
H1a: tangibility → USMS	0.199	0.184	3.730	0.001	0.843	2.32
H1b: reliability → USMS	–0.034	–0.030	–0.226	0.821		
H1d: safety → USMS	0.619	0.597	3.743	0.001		
H1e: empathy → USMS	0.189	0.197	2.711	0.007		
H2a: tangibility → USNMS	0.030	0.028	0.443	0.658	0.886	2.78
H2b: reliability → USNMS	0.892	0.795	4.098	0.001		
H2d: safety → USNMS	–0.951	–0.921	–3.651	0.001		
H2e: empathy → USNMS	0.457	0.477	5.111	0.001		

Source: Own construction

When it comes to the results of testing the H1 and H2 hypotheses, it can be confirmed that six of the eight sub-hypotheses are supported by this study. The results show that tangibility ($t = 3.730$, $p < 0.01$), security ($t = 3.743$, $p < 0.01$) and empathy ($t = 2.711$, $p < 0.01$) have a statistically significant positive impact on the level of customer satisfaction medical services, while reliability does not have a statistically significant positive influence on the level of user satisfaction with medical services ($t = – 0.226$, $p > 0.1$). According to the obtained results, it is obvious that the users of the tertiary level of health care do not sufficiently perceive the provided reliability as a dimension of the quality of the health service, however, although such user perception was identified, it does not affect the overall level of satisfaction with medical services. Therefore, other isolated dimensions of the quality of health services at the tertiary level of health care have a much greater impact on user satisfaction. Of course, this should not mean for management that reliability improvement activities should not be affirmed.

The results further indicate that reliability ($t = 4.098$, $p < 0.01$) and empathy ($t = 5.111$, $p < 0.01$) have a statistically significant positive influence, while security ($t = – 3.651$, $p < 0.01$) has statistically significant, but negative impact on the level of user satisfaction with non-medical services. The results of the latent construct tangibility at the tertiary level of health care show that it has no statistically significant effect on the level of satisfaction with non-medical services ($t = 0.443$, $p > 0.1$). The reason for this lies in the fact that tangibility is directly related to the implementation of quality medical services, and is very little represented, that is, noticeable in the case of non-medical services. In the case of the latent construct safety, a statistically significant negative impact on the level of satisfaction with non-medical services is noticeable, which certainly has its logical basis. Namely, in order to provide a higher level of security such as better quality medicines, especially in the case of more complex and demanding hospital conditions, the users have to pay a higher price and undergo more complex internal procedures in clinical centers, which directly contributes to a decrease in their satisfaction and vice versa.

The table also shows indicators of the coefficient of determination (R^2) and effect size and for the calculation of which Cohen's f indicator was used. As a "rough rule" it can be said that R^2 values of 0.25–0.50 indicate weak, 0.50–0.75 moderate, and above 0.75 high explanatory power of the model (Hair et al., 2019). However, for social sciences, lower limit values of the coefficient of determination (0.02, 0.13 and 0.26) were proposed. Therefore, when interpreting R^2, one should always take into account the context of the research, i.e. the scientific discipline in which the model is observed (Sarstedt, Ringle and Hair, 2017). Looking at the explanatory power of the estimated basic structural model in Table 10, it can primarily be determined that the coefficients of determination (R^2) are high. For both endogenous constructs, the indicators are higher from the stated limit, which confirms the high explanatory power of the estimated model. The determination coefficients show that a high proportion of the variance in the endogenous constructs (84.3% and 88.6%) is explained by the combination of the influence of exogenous latent variables on the endogenous latent variable bias towards complex models, an *adjusted coefficient of determination* R^2_{adj} can be calculated. Adjusted coefficients of determination (R^2_{adj}) of the estimated basic structural model, which can be used to compare different models, have similar values as the coefficient of determination (84.2% i 88.5%). Additionally, the change in R^2 value when a particular exogenous construct is omitted from the model can be used to estimate the magnitude of its effect on the endogenous construct. This measure is called the effect size (f), and it assesses how strongly an exogenous construct contributes to explaining a certain endogenous construct in terms of R^2 (Avkiran, 2018), which is calculated according to the following formula (Cohen, 1988):

$$\textbf{Cohen's } \boldsymbol{f} = \sqrt{\frac{\mathbf{R}^2}{\left(1-\mathbf{R}^2\right)}}\,. \tag{1}$$

Guidelines for interpreting effect sizes are as follows: 0.01 – weak effect, 0.20 – moderate effect, and 0.40 strong effect. Based on the tabular presentation, we can state a strong influence of independent constructs on the dependent construct, that is, the quality of health services has a strong influence on the level of user satisfaction with medical services (f = 2.32) and on the level of user satisfaction with non-medical services (f = 2.78). Based on the basic structural model and empirical research, the following table presents the hypotheses.

Table 11 Conclusions of hypothesis testing – basic structural model

Hypotheses	Content of the hypothesis	Conclusion of the analysis
H1a	Tangibility has a statistically significant effect on the level of user satisfaction with medical services.	**Confirmed**
H1b	Reliability has a statistically significant effect on the level of user satisfaction with medical services.	**Not confirmed**
H1d	Safety has a statistically significant effect on the level of user satisfaction with medical services.	**Confirmed**
H1e	Empathy has a statistically significant effect on the level of user satisfaction with medical services.	**Confirmed**
H1	***There is a statistically significant influence of the quality of health services on the level of user satisfaction with medical services in Bosnia and Herzegovina.***	**Confirmed**
H2a	Tangibility has a statistically significant effect on the level of user satisfaction with non-medical services.	**Not confirmed**
H2b	Reliability has a statistically significant effect on the level of user satisfaction with non-medical services.	**Confirmed**
H2d	Safety has a statistically significant effect on the level of user satisfaction with non-medical services.	**Confirmed**
H2e	Empathy has a statistically significant effect on the level of user satisfaction with non-medical services.	**Confirmed**
H2	***There is a statistically significant influence of the quality of health services on the level of user satisfaction with non-medical services in Bosnia and Herzegovina.***	**Confirmed**

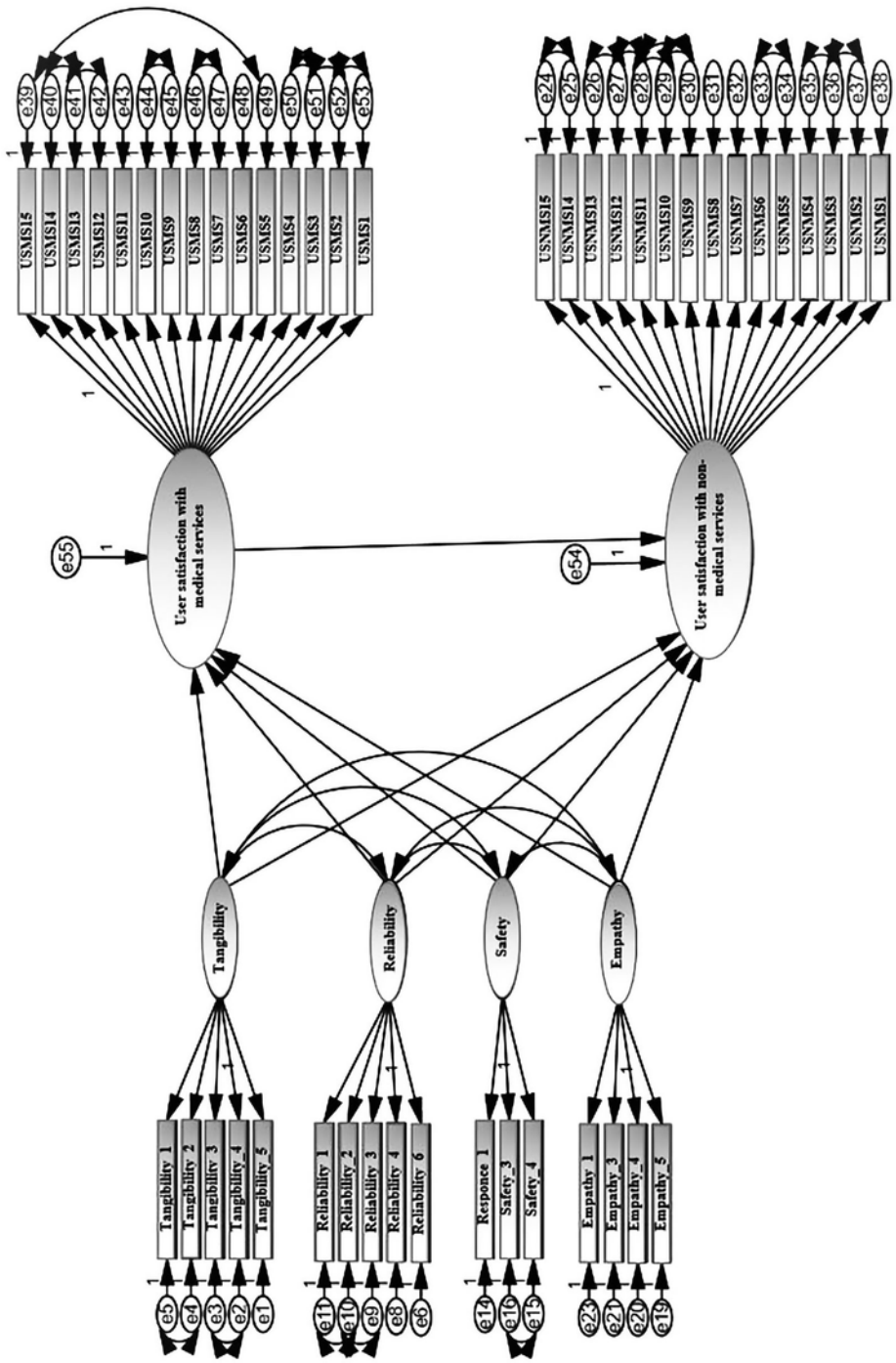
Source: Own construction

Considering the complex connections that can be analyzed by SEM, it is common to visually present the so-called model path diagram. The symbolism in the diagram is as follows (Ho et al., 2012):

- Manifest variables are shown as squares or rectangles;
- Latent variables are shown as circles or ellipses;
- The assumed influence of one variable on another (direct effect) is shown by a straight arrow with one tip;
- Covariances (in the non-standardized solution) and correlations (in the standardized solution) between independent (exogenous) variables are shown by a curved line with arrows.

In addition to the above, error components of manifest and latent variables are also displayed. SEM model parameters include: direct influences on endogenous variables (either from exogenous or other endogenous variables), factor loadings that connect indicators with the corresponding latent variable, and variances and covariances of exogenous variables (Kline, 2011). The following Figure shows the structural model obtained after testing the reliability, validity and hypotheses.

Figure 1 Conclusions of hypothesis testing – basic structural model



Source: Own construction

4 DISCUSSION

A direct comparison of the obtained results with previous research is not possible due to the use of modified measuring scales, as due to the fact that very little research is based on comparative SEM analysis.. However, we can indirectly make a comparison with research based on SERVQUAL or HEALTHQUAL analyses of tertiary levels of health care. The confirmation of the hypotheses about the influence of the quality dimension on user satisfaction (medical and non-medical services) is primarily in accordance with previous research on this topic (Nashrath et al., 2011; Lee and Kim, 2017; Natarajappa et al., 2020). However, there are certain deviations. Thus, Nashrath et al. (2011) point out that Reliability is the greatest source of user satisfaction of tertiary level health care, while the results of the presented research indicate that Reliability is not the source of user satisfaction with medical services, but it is in the case of non-medical services. Also, in their research, Response achieved the lowest values, and as we saw in the results of our research, Response did not correspond to the construct, and was excluded from the hypothesis testing process.

Although it is based on the application of the HEALTHQUAL questionnaire, the research from Lee and Kim (2017) can be partially compared with this one since there are common quality dimensions such as Empathy, Safety and Tangibles. Using confirmatory factor analysis, these authors proved that it was necessary to act simultaneously on all dimensions of the quality of the tertiary level of health care in order to achieve user satisfaction, while in the research in Bosnia and Herzegovina it was established that Tangibility has no effect on satisfaction with non-medical services and Reliability on medical services.

In their work, Natarajappa et al., (2020) focused more on the impact of medical and non-medical dimensions of quality on user loyalty. Although our research does not deal directly with the loyalty of users of the tertiary level of health care, taking into account the theoretical assumption that satisfaction is a prerequisite for loyalty, we can state that the conclusions of these two studies coincide.

While interpreting the results, it is necessary to take into account certain limitations, as well as recommendations for future research arising from the aforementioned limitations. The first limitation is reflected in the limitations of literature bases, which can lead to the exclusion of certain publications that would have a significant impact on the formation of a theoretical concept and its empirical testing. Another limitation relates to the time frame for conducting the empirical research. Considering the spatial limitation of research at the level of Bosnia and Herzegovina, it should certainly be aimed at confirming the created structural model and research at the level of other countries. Such a comparative analysis would contribute to greater objectivity and originality of the presented research. Also, it is recommended to carry out identical research at other levels of health care in order to examine the (non)consistency of differences in the importance of individual dimensions of the quality of health services to user satisfaction. Finally, it is suggested to check the measuring scales on the same or a similar sample in order to test the suitability, reliability, validity and objectivity of future research on the impact of the quality of health services on the satisfaction of users.

CONCLUSION

The focus of business in the twenty-first century is to achieve quality in order to satisfy the increasing demands of users. Quality becomes a basic assumption for the survival and development of any organization. Quality management in healthcare is gaining more and more importance, although it has not been fully explored. Health care institutions, especially public ones, by their very nature are not profit-oriented. However, the essence of their existence is the provision of the highest quality health services to the population and the well-being of society as a whole. In order to achieve the primary goal of their existence, it is necessary to introduce systems, internal processes, organization of people and other resources in a way that will contribute to user satisfaction and the fulfillment of their mission. Achieving

the satisfaction of users of health services is not an easy task, considering that it is a question of the results of treatment and the safety of the health of each individual user.

The problem of user satisfaction of health services becomes more complex with the increasing level of health care provided to the user. Therefore, the main motive was aimed at providing assistance to the management of health institutions that operate at the tertiary level of health care in identifying key factors, that is, quality dimensions that contribute the most to the satisfaction of their users. Since satisfaction is a broad and multidimensional concept, a distinction was made between satisfaction with medical and non-medical services. In this way, the direct effects of certain dimensions of quality on the main (medical) services and on secondary (non-medical) services are shown, which together form the overall perception of clinical centers, that is, tertiary hospitals. In the previous publication, the authors primarily focus on the medical aspects of the service, ignoring the fact that admission, discharge and other non-medical aspects have an impact on the overall level of satisfaction of users of health services. This problem was overcome through scientific work and an original and unique construct of the quality of health services and user satisfaction at the tertiary level of health care was created by applying SEM analysis.

The construct confirmed the persistence of the cause-and-effect relationship between the quality of health services and the level of satisfaction with both medical and non-medical services. The conclusion of the construct is the persistence of the inextricable link between the mentioned variables, and that management must simultaneously look at all aspects of the services provided in tertiary level health care institutions with additional engagement in improving the quality dimensions that contribute most to the satisfaction of health care users.

Considering that the statistically significant influence of the quality of health services on the satisfaction of users of health services has been proven, the inevitable conclusion is the necessity of improving working conditions as an initial assumption of user satisfaction. The improvement of working environment conditions can be realized through a greater degree of integration of employees in business decision-making, valorization of their opinions, and enabling open and closer interaction between employees, as well as between management and employees. Employees of healthcare institutions are an inexhaustible source of information and recommendations for improving the quality of services and the satisfaction of users of healthcare services, and are an excellent indicator of successful quality management in healthcare.

References

ALALOOLA, N. A., ALBEDAIWI, W. A. (2008). Patient satisfaction in a Riyadh Tertiary Care Centre. *International Journal of Health Care Quality Assurance*, 21(7): 630–637.

AVKIRAN, N. K. (2018). Rise of the partial least squares structural equation modeling: an application in banking. In: AVKIRAN, N. K., RINGLE, C. M. *Partial least squares structural equation modeling-Recent Advances in Banking and Finance,* Cham: Springer, 1–29.

BOLTON, R. N., DREW, J. H. (1991). A Longitudinal analysis of the impact of service changes on customer attitudes. *Journal of Marketing,* 55: 365–384.

COHEN, J. (1988). *Statistical power analysis for the behavioral sciences*. 2nd Ed. New Jersey: Lawrence Erlbaum.

CRONIN, J. J., TAYLOR, S. A. (1992). Measuring service quality: a reexamination and extension. *Journal of Marketing*, 6: 55–68.

DILSHAD, S., AKHTAR, A., HUDA, S. S., SAMAD, N. (2020). Assessment of Healthcare Service Quality: Tertiary Care Hospitals of Dhaka City. *Global Issues and Innovative Solutions in Healthcare, Culture, and the Environmen*t, 271–291.

FULLER, E. L., HEMMERLE, W. J. (1966). Robustness of the maximum-likelihood estimation procedure in factor analysis. *Psychometrika*, 31(2): 255–266.

GARG, N., GUPTA, S. K., MAHESH, R. (2014). Patient Satisfaction Survey at a Tertiary Care Speciality Hospital. *Int. J. Res. Foundation Hosp. Healthc Adm*, 2(2): 79–83.

HAIR, J. F., RISHER, J. J., SARSTEDT, M., RINGLE, C. M. (2019). When to use and how to report the results of PLS-SEM [online]. *European Business Review*, 31(1): 2–24, <https://doi.org/10.1108/EBR-11-2018-0203>.

HAM, H. S., PECK, E. H., MOON, H. S., YEOM, H. A. (2015). Predictors of Patient Satisfaction withTertiary Hospitals in Korea [online]. *Hindawi Publishing Corporation Nursing Research and Practice*, 1–6. <https://doi.org/10.1155/2015/749754>.

HO, M.-H. R., STARK, S., CHERNYSHENKO, O. (2012). *Graphical representation of structural equation models using path diagrams.* Psychology Faculty Publications.

HU, L., DING, H., HU, G., WANG, Z., LIU, S., LIU, Y. (2019). How perceived quality of care affects outpatient satisfaction in China: a cross-sectional study of 136 tertiary hospitals. *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*, 56: 1–8.

IMAM, S. Z., SYED, K. S., ALI, S. A., ALI, S. U., FATIMA, K., GILL, M., HASSAN, M. O., HASHMI, S. H., SIDDIQI, M.T., KHAN, H. M., JAMEEL, O. F. (2007). Patients' satisfaction and opinions of their experiences during admission in a tertiary care hospital in Pakistan – a cross sectional study [online]. *BMC Health Services Research*, Vol. 7. <https://doi.org/10.1186/1472-6963-7-161>.

HENSELER, J., RINGLE, C., SARSTEDT, M. (2015). A New Criterion for Assessing Discriminant Validity in Variance-based Structural Equation Modeling [online]. *Journal of the Academy of Marketing Science*, 43: 115–135. <https://doi.org/10.1007/s11747-014-0403-8>.

KLINE, R.B. (2011). *Principles and Practice of Structural Equation Modeling*. 3rd Ed. New York: The Guilford Press.

KULKARNI, S. K., (2018). A study of patient satisfaction level in Out Patient Department (OPD) in a Tertiary Care Hospital in Maharashtra. *Journal of Dental and Medical Sciences,* 17(3).

LEE, D., KIM, K. K. (2017). Assessing healthcare service quality: a comparative study of patient treatment types. *Int. J. Qual. Innov.,* 3(1): 1–15.

MAROOF, B., TASNEEM, S. S., JABEEN, R. J., TASKEEN, R. (2019). The Assessment of Patient's Satisfaction and Opinions of their Experience during Admission in the Tertiary Care Hospital [online]. *International Journal of Academic Research Business and Social Sciences,* 9(5): 555–577. <http://dx.doi.org/10.46886/IJARP/v9-i5/5905>.

MIKULIĆ, J. (2007). Quality Management of Airportservices – an Analysis of the Multifactorstructure of Customer Satisfaction. *Market*, Vol. XIX(1), Zagreb: Faculty of Economics.

NASHRATH, M., AKKADECHANUNT, T., CHONTAWAN, R. (2011). Perceived nursing service quality in a tertiary care hospital Maldives [online]. *Nursing & Health Sciences*, 13(4): 495–501. <https://doi.org/10.1111/j.1442-2018.2011.00648.x>.

NATARAJAPPA, SH., RAVI, M. D., CHANDRASHEKAR, B. R., NAGESH, P. (2020). Evaluation of Service Quality in a Tertiary Care Hospital Using a Novel Patient Satisfaction Survey Tool in Mysore City, India – a Pilot Study. *Indian Journal of Public Health Research & Development,* 11(7): 858–865.

NWABUEZE, C. J., ONYEAGU, S. I., IKPEGBU, O. (2009). Robustness of the maximum likelihood estimation procedure in factor analysis. *African Journal of Mathematics and Computer Science Research,* 2(5): 81–87.

OJENIWEH, O., POPOOLA, O., ADESINA, A. D., MARIERE, U., ROTIFA, S. (2021). Patient satisfaction with physical therapy care, in a tertiary healthcare facility: an evaluation of quality of physical therapy services provided in Federal Medical Centre [online]. *Yenagoa, Bayelsa State, Yen Med*, 3(2): 116–123. <https://doi.org/10.13140/RG.2.2.31742.10567>.

OLIVER, R. L. (1993). A conceptual model of service quality and service satisfaction: compatible goals, different concepts. *Advances in Service Marketing and Management*, 2: 65–85.

OZRETIĆ, D. Đ., ŠKARE, V., ČERFALVI, V., BENCEKOVIĆ, Ž., KOMARAC, T. (2020*).* Assessment of the Quality of Public Hospital Healthcare Services by using SERVQUAL [online]. *Acta Clin. Croatia*, 59(2): 285–293. <https://doi.org/10.20471/acc.2020.59.02.12>.

SARSTEDT, M., RINGLE, C. M., HAIR, J. F. (2017). *Partial least squares structural equation modeling*. In: HOMBURG, C., KLARMANN, M.,VOMBERG, A. E. *Handbook of market research*, Cham: Springer.

SHARIF, P., SHARIF, N. (2018). *Structural Equation Modeling with AMOS.* Teheran: Artin Teb.

SNOJ, B. (1995). The profiles of importance of service quality components in health spas [online]. *Der Markt,* 34(2): 95–104. <https://doi.org/10.1007/BF03032046>.

SNOJ, B., SAVIĆ, V., RAJTMAJER, B. (1999). Measuring the perceived quality of ISS Servisystem from Maribor. *Akademija MM*, 5: 33–42.

STAFFORD, M., STAFFORD, T. F., WELLS, B. P. (1998). Determinants of customer satisfaction in the auto casualty claims process. *The Journal of Services Marketing*, 12(6).

ZEITHAML, V. A., BERRY, L. L., PARASURAMAN, A. (1993). The nature and determinants of customer expectations of service [online]. *Journal of the Academy of Marketing Science*, 21(1): 1–12. <http://dx.doi.org/10.1177/0092070393211001>.