# The Deployment of Balanced Scorecard Perspectives and Dimensions to Improve Hospitals' Performance

## Ph.D. Dissertation booklet

Faten Amer

University of Pécs
Faculty of Health Sciences
Doctoral School of Health Sciences
Head of Doctoral School: Prof. Dr. József Bódis
Programme Leader: Prof. Dr. Gábor L. Kovács
Supervisor: Dr. Dóra Endrei
Co-Supervisor: Prof. Dr. Imre Boncz

Pécs, 2023

# Abbreviations:

| Abbreviation | | Abbreviation | |
|---|---|---|---|
| ACC-EXR | Hospital accessibility experience | MTR | Managerial trust |
| AQSS | Availability and quality of supplies and services | NBR | No blame error reporting |
| AVE | Average variance extracted | NGO | Non-governmental organization |
| BSC | Balanced scorecard | PATIENT-ATTs | Patient attitudes |
| BSC-PATIENT | Instrument which engages patients in a comprehensive assessment of BSC perspectives and dimensions | PATIENT-CENT-EXR | Patient-centeredness care experience |
| BSCP-ATTs | Attitudes toward BSC perspectives | PATIENT-ENG | Patient engagement |
| BUILCAP-EXR | Building capacity experience | PATIENT-EXR | Patient experiences |
| BUILENV-EXR | Building environment experience | PI | Perceived image |
| BUIL-EXR | Building experience | PICO | Population, intervention, comparison, and outcome |
| CFA | Confirmatory factor analysis | PMOH | Palestinian Ministry of Health |
| CFI | Comparative fit index | PQ | Perceived quality |
| CITC | Corrected item-total correlation | PR-EXR | Price experience |
| COMP-PI | Perceived image of the complications | PRISMA | Preferred reporting items for systematic reviews and meta-analyses |
| COVID-19 | Coronavirus-19 | QSS | Quality of supplies and services |
| CR | Composite reliability | QUALDEV | Quality and development |
| CVI-UA | Universal agreement among experts for the content validity index | RCTs | Randomized controlled trials |
| CVR | Content validity ratio | REPUT | Community and reputation |
| EFA | Exploratory factor analysis | RESCOMINF | Needs-response, communication, and information provision |
| FINI | Financial incentives | RMSEA | Root mean square error of approximation |
| GFI | Goodness-of-fit index | RoB | The risk of bias |
| HCOs | Healthcare organizations | RoB-2 | Cochrane risk-of-bias tool for randomized trials-version two |
| HCW-CENT | Healthcare workers-centeredness | ROBINS-I | The risk of bias in the non-randomized intervention studies tool |
| HCW-ENG | Healthcare workers' engagement | S-CVI | Scale content validity index |
| HCWs | Healthcare workers | SERV-EXR | Services experience |
| HCW-SCIDEV | Health care workers' scientific development | SQRT | Square root of the AVE |
| I-CVI | Item content validity index | SRMR | Standardized root mean square residual |
| IIC | Interitem correlation | STROBE | Strengthening the reporting of observational studies in epidemiology |
| INFO-EXR | Information experience | TECH | Technology |
| IRB | Institutional review board | TECH-PI | Technology-perceived image |
| ITRODP | Introductory period | TLI | Tucker–Lewis's index |
| KMO | Kaiser–Meyer–Olkin | TQM | Total quality management |
| KPIs | Key performance indicators | WT-EXR | Waiting time experiences |
| LOY-ATTs | Patient satisfaction and loyalty attitudes | WTLB | Workload time- life balance |
| MANAG-PE | Managerial tasks and performance evaluation | | |
| MeSH | Medical subject headings | | |

# Chapter 1: Introduction

## Aim of the dissertation

This dissertation consists of four core sub-studies that aim to engage stakeholders in the strategic performance evaluation of Palestinian hospitals based on the balanced scorecard (BSC) tool. The dissertation core sub-studies have the following detailed main objectives:

1. To perform a systematic review to a) gather all studies that have measured the impact of implementing BSC on healthcare workers' (HCWs') satisfaction, patient satisfaction, and financial performance at healthcare organizations (HCOs), particularly since these three measured outcomes represent the latest affected perspectives in the strategic maps (Kaplan, 2009; Mcdonald, 2012), and b) assess and compare the results among the included studies.

2. To perform a systematic review to a) find and recategorize all the perspectives, dimensions, and key performance indicators (KPIs) that were employed in BSC implementations for unification purposes, b) rank dimensions according to their frequency of use by HCO worldwide, and c) rank dimensions according to their importance from healthcare managers' perspectives.

3. To a) develop an instrument that assesses 1. patient experiences (PATIENT-EXR) in light of BSC perspectives, and 2. patient attitudes (PATIENT-ATTs) in light of BSC perspectives such as perceived image (PI) and attitudes toward BSC perspectives (BSCP-ATTs) including patient satisfaction and loyalty attitudes (LOY-ATTs), and b) customize the developed instrument to Palestinian hospitals, translate it into Arabic, and validate it.

4. To a) develop an instrument that performs a comprehensive assessment of hospitals based on BSC perspectives and dimensions, and b) customize the developed instrument at Palestinian hospitals, translate it into Arabic, and validate it.

Additionally, the dissertation has four sub-studies, which are implementations of the core sub-studies. These implementations have four adjunctive objectives:

1- To implement an assessment for cancer care in Palestine based on BSC perspectives

2- To implement an assessment of BSC perspectives and dimensions in hospitals during the Coronavirus-19 (COVID-19) pandemic.

3- To implement the first developed instrument to perform patient engagement (PATIENT-ENG) in BSC implementation at Palestinian hospitals

4- To implement the second developed instrument to perform HCWs' engagement (HCW-ENG) in BSC implementation at Palestinian hospitals.

## Background

### Performance evaluation (PE) of the healthcare sector

The healthcare sector's PE is quite challenging and complex. Unsatisfactory performance can result from long patient waiting time experiences (WT-EXR), inefficiency, dissatisfactory patients, and HCW burnout (Buathong & Bangchokdee, 2017; Meena & Thakkar, 2014). COVID-19 imposed further burdens on the healthcare system worldwide due to the limited capacity of hospital beds and the increased psychological stress of HCWs during the COVID-19 pandemic (Amer et al., 2021; Lou et al., 2021). There is still a lack of information that would help healthcare managers and policymakers in the era of COVID-19 to improve the delivery of healthcare quality and to learn for the future (Austin & Kachalia, 2020). Higher pandemic burdens, such as HCWs' burnout and stress, will rise when HCOs lack plans and preparedness to strengthen their surge capacity and HCWs' resilience (Afulani et al., 2021; Ali et al., 2020). The World Health Organization initiated the Performance Assessment Tool for Quality Improvement in Hospitals in

2003. It aimed to develop a framework for the assessment of hospital performance. The resulting dimensions from this project were clinical effectiveness, efficiency, HCWs' orientation, responsive governance, safety, and patient-centeredness care experience (PATIENT-CENT-EXR). However, studies have shown that there are still some gaps in this model and issues concerning the dimensions investigated (Carini et al., 2020; Veillard et al., 2005). Additionally, the Organization for Economic Co-operation and Development launched the Healthcare Quality Indicator project in 2006; it aimed to develop KPIs to compare quality in healthcare at the international level and achieve international benchmarking. This project concluded that healthcare must be safe, effective, PATIENT-CENT-EXR, timely, efficient, equitable, acceptable, and hospital accessibility experience (ACC-EXR) (Kelley & Hurst, 2006; OECD, 2019).

The BSC was first suggested by Norton and Kaplan in 1992 (Kaplan & Norton, 1992). The first generation of the BSC, unveiled by Kaplan and Norton in 1992, involved four perspectives: the financial, customer, internal process, and knowledge and growth perspectives, steered by the organizational vision and strategy (Kaplan & Norton, 1992). The customer perspective focused mainly on the patients. However, in some implementations, it also included HCWs or both (Amer et al., 2022a, 2022b). BSC perspective assessment provides managers with a comprehensive PE approach (Amer et al., 2022a). In comparison with other PE tools, most of the available PE models mainly focus on the internal perspective but lack coverage of the other dimensions or perspectives that are also important. BSC was considered different from the other managerial tools for two reasons. First, it offers a holistic approach to PE since it allows managers to highlight both financial and nonfinancial metrics. Second, the BSC is not only a planning or a PE tool. It is also a strategic managerial tool that assigns KPIs compatible with the HCO strategy. However, other managerial tasks and performance evaluation (MANAG-PE) tools, such as total quality management (TQM), lack these comprehensive properties.

## Literature gap

Despite the importance of PATIENT-ENG and HCW-ENG in the literature, recent reviews (Bohm et al., 2021) revealed that there was a lack of engaging stakeholders in BSC implementations. In addition to this literature gap. To our knowledge, no research has utilized the BSC tool to evaluate the performance of Palestinian hospitals to date. Moreover, there is a lack of PATIENT-ENG and HCW-ENG in the PE implementation of Palestinian hospitals in general. Consequently, this dissertation implication will make comprehensive PEs of Palestinian hospitals based on the HCWs' and patients' points of view. This will lead to recommendations for Palestinian hospital managers as well as health policymakers on how to improve the PEs of Palestinian hospitals in their future action plans.

# (Sub-study 1): The Deployment of BSC in HCOs: Is it Beneficial? A Systematic Review[1]

---

[1] This chapter is based on the following paper: **Amer et al.,** 2022. The deployment of balanced scorecard in health care organizations: is it beneficial? A systematic review. BMC Health Services Research, 22(1), 1–14. https://doi: 10.1186/s12913-021- 07452-7. **Impact factor: 2.908**

# Introduction

Until the beginning of 2022, two reviews focused on studying the effect of the BSC on health, one of which analyzed the impact qualitatively (Rabbani et al., 2007), and the other presented a few instances of the positive influence (Mcdonald, 2012). This showed that no complete or rigorous scientific methodology has been reported to evaluate the effect of BSC adoption in HCO. Given the lack of research on this topic, we performed a systematic review in which we assessed the impact of implementing the BSC on three attributes that represent the latest affected perspectives in strategic maps (Kaplan, 2009; Mcdonald, 2012): HCWs' satisfaction, patient satisfaction, and financial performance. Thus, the present systematic review aims to gather all studies that have measured the impact of implementing the BSC on HCWs' satisfaction, patient satisfaction, and financial performance at HCO, particularly since these three attributes represent the latest affected perspectives in strategic maps (Kaplan, 2009; Mcdonald, 2012). Furthermore, this review aims to assess and compare results among the included studies.

# Methods

Our previous systematic review analyzed the dimensions and indicators of the BSC utilized at the PE. of HCO (Amer et al., 2022a). This systematic review was carried out by finding all studies that approached the impact of BSC implementation on HCOs in adherence with the 27-point of the preferred reporting items for systematic reviews and meta-analyses (PRISMA) checklist (Liberati et al., 2009).

The inclusion criteria were as follows: any type of HCO performed a performance assessment of HCOs by implementing the BSC for at least one year and measured the impact on financial indicators or on the patient satisfaction rate or the HCWs' satisfaction rate. The impact should be objective and measured/quantitative. The search strategy was developed by two authors who are experts in healthcare management and BSC and an expert in systematic reviews and meta-analysis. The search strategy was initially developed for the PubMed database based on the population, intervention, comparison, and outcome (PICO) tool (Methley et al., 2014) and depended on using both medical subject headings (MeSH) terms and keywords. Next, the strategy was adapted to the Cochrane Central Register of Controlled Trials, Embase, and Google Scholar databases, as per Cochrane's recommendations (Lefebvre et al., 2021).

The gray literature, pre-prints, and unpublished studies were searched on Google Scholar and Google's search engine websites to reduce publication bias. Furthermore, we attempted to identify other potentially eligible studies or ancillary publications by searching the reference lists of any potentially eligible studies. The databases were searched until October 2020. Afterward, we conducted the search strategies on the electronic databases and removed duplicates using the EndNote X9.2 program. Two authors independently performed the selection of eligible studies. A discussion after each step was made or, if necessary, a third author was consulted for arbitration in case of disagreements. Initially, the titles and abstracts of the studies were examined to eliminate irrelevant studies. In the second step, the full texts of all potentially relevant studies were carefully reviewed to make a final decision based on the criteria mentioned above. Authors of studies with no available full texts or unclear impact duration were contacted to obtain further details and clarification.

Data extraction was performed between June and July 2021 and then compared to discuss differences. The following data were extracted from the eligible studies: 1) author/s, 2) year of publication, 3) country of origin, 4) data collection duration, 5) data collection tool, 6) the number of perspectives, 7) the number of KPIs, 8) availability of weights/importance for perspectives or KPIs, and 9) outcome, which is represented in the KPIs that have been used and their

weights/importance. The frequency of each KPI used at each implementation was plotted in Microsoft Excel, and the sum was calculated. In addition, the weight/importance assigned for each KPI at each implementation was reported on a scale of 100%. In the case of studies that did not give weights/importance explicitly, each KPI weight/importance was calculated by dividing one by the number of KPIs used in that study to assign an equal weight/importance for each KPI. Consequently, we computed an average of the weights/importance assigned for each KPI. Next, we performed regrouping and coding for the KPIs to find the frequency of use and the set weights/importance percentages for each dimension. Then, the resulting major and sub-dimensions were listed and described between August and September 2021.

The risk of bias (RoB) assessment was performed by two authors independently. As per the Cochrane collaboration's guidelines, the Cochrane risk-of-bias tool for randomized trials-version two (RoB-2) was used for the assessment of randomized controlled trials (RCTs) (Higgins et al., 2019). The risk of bias in the non-randomized intervention studies tool (ROBINS-I) was used to assess the observational and quasi-experimental studies (Sterne et al., 2016).

# Results

Initially, the search strategy resulted in a total of 4031 studies. After removing duplicates, a total of 2985 studies remained, which were screened based on their titles and abstracts. Then, irrelevant studies were excluded; thus, 202 studies remained. A careful examination of the full texts of the included studies was made; based on this, only 20 studies were finally included in the current systematic review. The results revealed a positive impact of the BSC on patient satisfaction and financial performance and, to a lesser extent, on HCW satisfaction. The RoB assessment revealed a moderate RoB in many studies.

# Discussion

This systematic review aimed to identify all the studies that measured the impact of BSC implementation on three variables, HCWs' satisfaction, patient satisfaction, and financial performance at HCO, and then proceeded to analyze the effect of these BSC implementations. The analysis of the results reflected a remarkably positive impact of BSC on patient satisfaction in most studies. The same positive impact of BSC implementation holds for financial performance in both currency and percentage indicators. Notably, the authors found that almost all studies showed a positive impact, amounting to several million United States Dollars. However, a few studies have reflected a moderately negative impact on financial performance, which forms three distinct categories.

To our knowledge, this is the first paper that has analyzed all the studies that measured the impact of BSC on patient satisfaction, HCWs' satisfaction, and financial performance in HCO. The results and analysis of this systematic review support the positive impact of applying the BSC in HCO, especially on patient satisfaction and financial performance. Furthermore, a greater emphasis on the role of HCWs is required when implementing the BSC since HCWs' satisfaction showed slightly positive, almost zero, or somewhat negative scores in most studies included. Additionally, the three primary outcome measures concentrated upon in this systematic review are considered the last destination for impact in the strategic maps and the causal effects in most BSC studies. Finally, unlike other BSC reviews (Mcdonald, 2012; Rabbani et al., 2007), which included definitions of biobanks, pharmacies, laboratories, radiology, and medical colleges in HCO, this review limited the definition to primary, secondary, or tertiary HCOs. This strategy leads to the homogeneity of the resulting studies and to more valid comparisons among the results. Nevertheless, this paper has some limitations. First, it focused on the impact of the BSC on the three chosen indicators only, whereas impacts on other types of indicators were not considered for

analysis. Due to the vast variations of indicator types, analysis of these indicators presents a challenge, requiring narrowly specified modes of analysis. Second, no meta-analysis could be applied to this systematic review due to the heterogeneity of studies regarding their data collection tools and the enormous variation in the types of indicators. However, the later variation was clarified in the charts, and the data collection tool was specified for each study. Third, the current review included studies that measured the impact after at least one year of implementation. Fourth, it is essential to mention that impact comparability is roughly more rational for patient satisfaction and HCWs' satisfaction than financial performance. This could be referred to as the comparison ability based on a percentage score of 100 for the satisfaction variables. Additionally, the change in financial performance based on currency could be influenced by other confounding factors, such as the HCO size or the number of health facilities included in the study. Therefore, future studies should consider these confounding factors. Moreover, future studies should reduce the RoB due to the lack of high-quality BSC implementations in the literature. Finally, this review searched for BSC implementation in healthcare databases; consequently, future systematic reviews are recommended to include studies in management and health policy databases.

## Conclusion

In conclusion, this systematic review offers evidence to HCO and policymakers on the benefits of implementing the BSC in HCO. Although the quality assessment revealed that many studies had a high RoB, BSC implementation positively influenced HCO patient satisfaction and financial performance. Based on the findings in the present review, researchers are encouraged to focus on lowering the RoB in BSC implementation in the future. HCO managers are also advised to consider HCWs' satisfaction and HCW-ENG in future BSC implementations. Finally, an additional assessment of the BSC impact on HCO during the COVID-19 pandemic is needed, as we could not find any.

## (Sub-study 2): A systematic review: the dimensions to evaluate healthcare performance and an implication during the pandemic[2]

## Introduction

BSC reviews (Behrouzi et al., 2014; Bohm et al., 2021; Broccardo, 2015; Colbran et al., 2019; T. Gao & Gurd, 2014; Gonzalez-Sanchez et al., 2018; Gurd & Gao, 2007; Mcdonald, 2012; Rabbani et al., 2007; Rouis et al., 2018; Trotta et al., 2013; Zelman et al., 2003) focused only on the general narration of the BSC perspectives and sub-dimensions used. Moreover, none of them summarized the perspectives or dimensions of the BSC based on their importance or frequency of use by healthcare managers. In other words, all the previous systematic reviews lack a systematic methodological categorization of perspectives, dimensions, and KPIs. In correspondence with this research gap, this review aims at a) finding and recategorizing all the perspectives, dimensions, and KPIs that were employed in BSC implementations for unification purposes, b) ranking dimensions according to their frequency of use by HCOs worldwide, and c) ranking dimensions according to their importance from the healthcare managers' perspective.

---

## Methods

The inclusion and exclusion criteria were set based on the PICO tool (Methley et al., 2014) and were the same as those used in the first systematic review. Additionally, all study designs were included. The same search strategy used in the first systematic review (Amer et al., 2022b) was also used for the second systematic review with a different aim.

Data extraction was performed between June and July 2021 and then compared to discuss differences. The following data were extracted from the eligible studies: 1) author/s, 2) year of publication, 3) country of origin, 4) data collection duration, 5) data collection tool, 6) the number of perspectives, 7) the number of KPIs, 8) availability of weights/importance for perspectives or KPIs, and 9) outcome, which is represented in the KPIs that have been used and their weights/importance. The frequency of each KPI used at each implementation was plotted in Microsoft Excel, and the sum was calculated. In addition, the weight/importance assigned for each KPI at each implementation was reported on a scale of 100%. In the case of studies that did not give weights/importance explicitly, each KPI weight/importance was calculated by dividing one by the number of KPIs used in that study to assign an equal weight/importance for each KPI. Consequently, we computed an average of the weights/importance assigned for each KPI. Next, we performed regrouping and coding for the KPIs to find the frequency of use and the set weights/importance percentages for each dimension. Then, the resulting major and sub-dimensions were listed and described between August and September 2021. The same methodology that was utilized in the first systematic review (Amer et al., 2022b) was also used to assess the RoB for the resulting studies in this review.

## Results

A total of 4028 studies resulted from running the search strategy in the four databases. In addition, another three studies were identified through a Google search. Therefore, a total of 4031 studies were included. Duplicates were removed (n=1046) using the EndNote program, and then the remaining articles were screened based on their titles and abstracts (n=2985). Irrelevant papers were excluded (n=2794). Consequently, the remaining 191 studies were examined by reading the full texts. Among these papers, 22 papers were written in non-English languages, including Spanish, German, French, Chinese, and Persian. A full-text translation was performed for each study to decide whether to include or exclude any of them. As a result of reading the full texts, 158 studies were excluded, and only 33 were eligible for this review, in which 36 full implementations of different BSC designs were actually applied.

A total of 797 KPIs were extracted from the resulting implementations. These KPIs were categorized in the studies under 15 perspectives. After regrouping these KPIs into homogenous major dimensions and sub-dimensions, 13 major dimensions resulted, with 45 sub-dimensions (*Figure 1*) in page 22. The selection bias across studies reflected a serious RoB in five studies (Chang et al., 2008; Chu & Wang, 2009; Rowe et al., 2014; Widyasari & Adi, 2019; Yang & Tung, 2006). Therefore, the intervention and the follow-up did not coincide, and a potentially substantial amount of follow-up was missing in their analysis. Studies with a moderate risk of intervention/exposure measurement bias reflected a well-defined intervention status, but some aspects of the assignments of intervention status were determined retrospectively. Furthermore, bias in selecting the reported results was serious in one study that partially reported the results (Widyasari & Adi, 2019). Studies that reported all results but did not have a preregistered protocol or whose outcome measurements were not defined in an initial plan were given a moderate risk.

## Discussion

All the perspectives, dimensions, and KPIs employed in BSC implementations were collected to fulfill the research aims. Categorization and regrouping of the KPIs into major and sub-dimensions were performed. Then, the dimensions were ranked according to their frequency of use and importance. The BSC tool can offer comprehensive planning, monitoring, evaluation, and improvement of HCO's KPIs. Hence, their performance should be improved in the short and long term. Analyzing the results showed that BSC implementations typically utilized four fundamental perspectives: financial, customer, internal, and knowledge and growth. However, the analysis revealed the frequent employment and the importance of other BSC perspectives in BSC implementations. Specifically, the external and managerial perspectives.

We believe that this paper has several strengths. First, this systematic review includes all types of studies with BSC implementations, such as books, theses, conference papers, and letters to the editor. Second, this review contains all implementations despite the country, language, or HCO administrative type, which gives an advantage of generalizing results to HCOs worldwide. Third, unlike other BSC reviews (Mcdonald, 2012; Rabbani et al., 2007), which included definitions of biobanks, pharmacies, laboratories, radiology, and medical colleges in HCOs, this review limited the report to primary, secondary, or tertiary healthcare organizations. However, an initial assessment by top management to evaluate the importance of each dimension and KPI based on the health organizations' strategy could be needed. This strategy leads to the homogeneity of the resulting studies and to more valid comparisons among the results. Fourth, this review calculates the use frequency of perspectives and the weights/importance assigned to them. Fifth, the first review has uniform KPIs in homogenous major dimensions and sub-dimensions despite the categorization differences among implementations, yielding more precise results. The resulting KPIs and dimensions in this review can be generalized or replicable to other HCOs and hospitals. Finally, this study is the first to analyze the implications of BSCs in HCOs during the pandemic based on the literature. This implication provides a guide for future theoretical implications, such as performing systematic reviews for each major dimension during the pandemic. It also provides a guide for practical implications of BSC dimensions to assess HCOs' performance.

However, this systematic review has some limitations. First, unlike previous studies, it excludes some HCOs, such as laboratories, pharmacies, radiology departments, and biobanks, as specified in the inclusion/exclusion criteria. Therefore, our results cannot be generalized to such HCO types. However, we excluded them to arrive at more homogenous KPIs and dimensions that are directly related to HCOs that offer primary, secondary, and tertiary medical services. Second, it includes only the articles that report the complete implementations of BSC while excluding studies that display only the BSC design without reporting the full implementation results. Third, we extracted the KPIs from all resulting implementations despite their RoB. However, we included an ROB assessment for each implementation.

## Conclusion

In conclusion, our review shows that the most frequently used perspectives in BSC papers were internal, financial, patient, learning and growth, HCW, managerial, community, and stakeholder perspectives. The perspectives that had the highest importance were internal, financial, learning and growth, patient, HCW, community, managerial, and stakeholder.

Moreover, this review solves the dilemma of the KPI categorization difference between BSC implementations by dimension unification into 13 major dimensions. The financial, information and innovation, technology (TECH), efficiency and effectiveness, availability and quality of supplies and services (AQSS), error-free and safety, time, healthcare workers-centeredness (HCW-

CENT), PATIENT-ATTs, needs-response, communication, and information provision (RESCOMINF), community and reputation (REPUT), HCO building, and MANAG-PE. The proper utilization of the 13 major dimensions and the 45 sub-dimensions will serve as a planning, monitoring, evaluation, and continuous improvement tool for HCOs, resulting in performance augmentation.

## (Sub-study 3): Assessing patient experience and attitude: BSC-PATIENT development, translation, and psychometric evaluation - a cross-sectional study[3]

## Introduction

BSC implementations utilized different sources to conduct the PE of HCOs (Amer et al., 2022a, 2022b), including hospital records, patient satisfaction questionnaires, patient and HCWs interviews, and observations. Additionally, BSC reviews (Amer et al., 2022a, 2022b) showed that only a few BSC implementations utilized validated scales to evaluate patient satisfaction, such as the Press Ganey questionnaires (Fields & Cohen, 2011; Smith & Kim, 2005). The patient satisfaction perspective is important since patients represent the hospitals' end receivers of healthcare services. However, researchers have pointed to the importance of PATIET-ENG in the process of health policy planning, evaluation, and delivery improvement (Anderson et al., 2021; Gagliardi et al., 2008). Additionally, patient feedback was proven to positively impact performance in HCO (Hammoudeh et al., 2020). Strategies to support PATIENT-ENG include RESCOMINF skills improvement, managing patient conflicts and complaints, maintaining patient confidentiality, patient training, and asking patients to review outputs by assessing their perceptions and experiences (Anderson et al., 2021; Bellows et al., 2015). It is not sufficient to perform the PE of HCO based on manager and hospital records only; a focus on PATIENT-ENG among the selection of the KPIs at HCO was recommended (Gagliardi et al., 2008). However, BSC reviews referred to the lack of patient and family member involvement in the evaluation process of BSC (Amer et al., 2022a, 2022b; Bohm et al., 2021)

The first aim of this research was to develop a comprehensive instrument which engages patients in a comprehensive assessment of BSC perspectives and dimensions (BSC-PATIENT), and can assess the following: 1. PATIENT-EXR in light of BSC perspectives, 2. PATIENT-ATTs include patient PI and BSCP-ATTs, such as patient satisfaction and LOY-ATT. The second aim of this research was to customize the developed instrument at Palestinian hospitals, translate it into Arabic, and validate it.

## Methods

This research is a cross-sectional study reported using strengthening the reporting of observational studies in epidemiology (STROBE) guidelines (Vandenbroucke et al., 2007). The questionnaire was created and validated based on the key authors' Kaplan and Norton theoretical

---

[3] This chapter is based on the following paper: **Amer F et al.**, 2022. Assessing Patient Experience and Attitude: BSC-PATIENT Development, Translation, and Psychometric Evaluation—A Cross-Sectional Study. International Journal of Environmental Research and Public Health. 19(12):7149. https://doi.org/10.3390/ijerph19127149. **Impact factor: 4.614**

framework (Kaplan & Norton, 1992) and the best practices for developing and validating the health and behavioral scales (Boateng et al., 2018).

The items were generated using the four-round Delphi method (Dalkey & Helmer, 1963) and the input of two panels. As a result, 52 items remained. The panelists rated the relevance and importance of each remaining item based on four- and three-point ordinal scales, respectively (Zamanzadeh et al., 2015). Next, we calculated the content validity ratio (CVR), the item content validity index (I-CVI), the scale content validity index (S-CVI), and universal agreement among experts for the content validity index (CVI-UA) to assess the content validity per item and scale (Zamanzadeh et al., 2015). Only the items rated 0.99 or above in CVR were included as per Lawshe guidelines (Lawshe, 1975). However, dimensions that scored 0.80-0.99 indicated the need to be revised. For the CVI, items that scored less than 0.60 were eliminated. Items that scored 0.6-0.79 were revised (Zamanzadeh et al., 2015).

The questionnaire items were initially developed in English. Then, it was translated into Arabic. All translations were prepared as per the translation and validation guidelines (Sousa & Rojjanasrirat, 2011). We performed a final review to produce the final corrected translation. An expert checked the final form in the BSC, and minor modifications were recommended. The first version of the questionnaire was piloted in one non-governmental organization (NGO) hospital in the south of West Bank. For that, 30 patients were asked to answer the first version of the questionnaire. They were asked to write their comments regarding language simplicity. The time needed to complete the questionnaire was also recorded. Items were coded before performing the analysis by IBM SPSS statistics 21 software. Then, Cronbach's alpha was calculated for each perspective to evaluate the internal consistency (Cho & Kim, 2015), and values above 0.6 were considered acceptable. Based on the results, some items were modified or deleted.

Institutional review board (IRB) approval for this research was received on 31 May 2020. All methods described in this study were approved by the Research and Ethics Committee at the Faculty of Medicine and Health Sciences at An Najah National University with the reference code number (Mas, May/20/16). Afterward, requests at 15 hospitals in West Bank and three hospitals in Jerusalem were applied between June and December 2020. The hospitals were selected using a convenience sample. However, the total number of beds per administrative type and governorate was considered when choosing the participants (HCO and patients). Public hospital approval was first applied to the Palestinian Ministry of Health (PMOH). Then, the request was applied to each hospital individually for all hospital types. The final form of the questionnaire was distributed between January and October 2021. The sample size was calculated according to the Steven K. Thompson sample size equation (Thompson, 2012), where n is the sample size, N is the population size, p is the estimated variability in the population (0.5), d is the margin of error (0.05), and the z score is at the 95% confidence interval (1.96). In our study, N was the population volume in the Palestinian territories (PMOH, 2020). Therefore, the needed sample size was found to be $n = 385$ patients. Additionally, studies considered 300 participants as a good sample size to successfully run each exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) or five respondents per parameter (Bentler & Chou, 1987; Comrey & Lee, 2013; Williams et al., 2010). Splitting the sample to perform EFA and CFA is recommended to perform construct validity (Knafl & Grey, 2007). Therefore, a total of 1000 questionnaires were distributed, anticipating a lower response rate during the pandemic.

I collected the data with four medical students at An-Najah University after giving three hours of training on BSC and the data collection steps and ethics to each medical student. Tasks and hospitals were delegated to them according to their living area: eastern Jerusalem and north,

middle, and south of the West Bank. The Gaza Strip was excluded due to the political situation and accessibility obstacles during the study. Moreover, five hospitals were excluded: two military hospitals that were not yet operating, one psychiatric hospital, and two rehabilitation hospitals. We sought variation in our sample regarding hospital size, area, and administrative type. For that, the maximum variation sampling strategy was used. The number of hospitals and the number of beds per administrative type were considered upon recruiting the sample (PMOH, 2020). Patients who existed in the targeted departments at the time of the visit were asked if they were willing to participate in the research.

Printed questionnaires were distributed to respondents instead of sending the questionnaires via email to reduce nonresponse bias (Sedgwick, 2014). Additionally, all participants were asked to agree on participation in a consent form that is coherent with the Declaration of Helsinki ethical principles (World Medical Association Declaration of Helsinki, 2013). Patients were informed that participation was confidential. Additionally, all patients were informed that participation was voluntary, so they could refuse participation in the study or withdraw at any time. To reduce the response bias (Sedgwick, 2014), the "I don't know (neutral)" answer was added as an option since experiences and attitudes can sometimes be uncertain (David Susman, 2021). Second, the data collectors ensured that the number of missing answers was minimized by checking the questionnaires upon retrieval. In case of missing parts, they drew the participant's attention to answer them. When entering data, if any questions were found to be still missing, they were entered as I don't know.

The inclusion and exclusion criteria were set to be a Palestinian patient above 15 years old of any gender. Outpatients should have finished receiving medical care at the assessed hospital or had received medical care at least once previously and returned to the same hospital. Inpatients should have been admitted for at least one day. The following departments were included: emergency room, internal medicine, surgery, gynecology, and pediatrics. In the emergency department, the questionnaires were completed by the patient companions. Additionally, in the pediatric department, the questionnaires were completed by one parent of the child. For the rest, questionnaires were completed by patients themselves; unless they were unable to complete the questionnaire, the questionnaires were read to them by the data collector or a family member and completed according to patient answers. To distinguish, a question was added to ask the respondent if their responses were based on their own, family, or friends' experiences.

Normality was tested using the Shapiro–Wilk test. The frequencies were used to analyze patient socio-demographics and the participating HCO characteristics. Our sample was split based on admission status to assess construct validity using EFA and CFA. EFA was performed for the inpatient sample using principal axis factoring with the Promax rotation method (Henry Mintzberg, 1990) in IBM SPSS statistics 21 software. The Kaiser–Meyer–Olkin (KMO) and Bartlett's sphericity tests were tested to determine the adequacy of the EFA (Kaiser & Rice, 1974). The inclusion or exclusion of a component was determined by an eigenvalue ≥ one (Larsen & Warne, 2010) and the visual assessment of Cattell's scree plot (Cattell, 1966). Item inclusion or exclusion was determined by a factor loading ≥ 0.50 and factor loadings on the assigned factor higher than all cross-loading of other factors (Williams et al., 2010).

CFA was performed for the components that resulted in EFA using the outpatient sample. The maximum likelihood estimation method in the IBM Amos 23 Graphics software (IBM, Wexford, PA, USA) was applied. The goodness of fit for the competing models was evaluated through the most commonly used fit indices. Minimum discrepancies were divided by degrees of freedom less than five and closer to zero, *P* value higher than 0.05, goodness-of-fit index (GFI),

comparative fit index (CFI), Tucker–Lewis's index (TLI), and cutoff values close to 0.95. Additionally, a root mean square error of approximation (RMSEA) < 0.06 and standardized root mean square residual (SRMR) value < 0.08 are needed before we can conclude that there is a relatively good fit between the hypothesized model and the observed data (Hooper et al., 2008; Shi et al., 2019). Item inclusion or exclusion in CFA was determined by a loading ≥ 0.50.

The interitem correlation (IIC) and the corrected item-total correlation (CITC) were calculated (Cohen, 1988). In this study, items with a correlation higher than 0.90 were considered redundant and deleted (Tavakol & Dennick, 2011). A correlation of 0.30 was considered the lower limit. Additionally, the composite reliability (CR) per factor was evaluated after performing CFA. CR is preferred over Cronbach's alpha, specifically in structural equation modeling (Peterson & Kim, 2013). In the current study, a CR ≥ 0.60 was considered sufficient (Ab Hamid et al., 2017; Hair J, Hult GTM, Ringle C, 2014). The Fornell-Lacker criterion was used to evaluate convergent and discriminant/divergent validity (Fornell & Larcker, 1981). The average variance extracted (AVE) was considered adequate for convergent validity if it was higher than 0.50 (Fornell & Larcker, 1981). However, if a value < 0.50 with CR > 0.60, the convergent validity of the factor was still considered adequate (Fornell & Larcker, 1981). To establish discriminant validity, the square root of the AVE (SQRT) should have a greater value than the correlations with other latent factors (Ab Hamid et al., 2017). Additionally, factor uniqueness was evaluated depending on the value of Spearman correlation (r) with other factors at the same scale. Researchers have recommended the separation of dependent and independent variables since the correlation between them can be misleading in assessing discriminant validity (Hair et al., 2015). Therefore, we assessed r for the independent and dependent factors separately. Then, r was described as negligible when r < 0.20, low (r = 0.20-0.49), moderate (r = 0.50-0.69), high (r = 0.70-0.85), or very high (r = 0.86-1.00) (Bookter, 1999; Plichta & Kelvin, 2011). In this study, the absence of high or very high r between the subscale factors indicated discriminant validity (Bookter, 1999).

## Results

The content validity resulted in removing one item and indicated that a revision is needed for eight items. The revised items required either further clarification and rewording or modification for specific participants. For example, the CVR results indicated that financial and price items should not be included for nonprofit hospitals. Additionally, the CVI results showed that particular items were relevant only to inpatients. This step increased the S-CVI, CVI-UA, and CVR from 0.90, 0.63, and 0.95 to 0.95, 0.78, and 0.97, respectively.

The patient socio-demographics and hospital characteristics section included age, gender, scientific degree, working sector, insurance availability, and type. Moreover, the number of visits to the evaluated hospital compares the attitudes of the new and previous customers. The number of earlier visits is considered necessary in the analysis since past customer behavior tends to be a good predictor of future behavior (David Susman, 2021). Moreover, the information source on which the respondent evaluation was built was recorded since perceptions and attitudes may emerge from direct personal experience or from observing other people's experiences, such as family and friends' experiences (American Psychological Association; Edward et al., 2015). The second section of the questionnaire was designed to measure PATIENT-EXR in light of BSC perspectives and their attitudes toward them, including patient perceived quality (PQ), PI, satisfaction, and LOY-ATT.

The pretest was performed at one NGO hospital in the south of the West Bank. Patients found the length of the questionnaire appropriate. Additionally, the layout was well accepted and clear. They gave specific minor comments that were incorporated. These corresponded to the

rewording of a few items. The time for completing the questionnaire was less than 10 minutes. Consequently, few modifications were made after piloting. Cronbach's alpha was calculated per the BSC perspective. All perspectives had a Cronbach's alpha above 0.7 at the pretest, except for the environmental perspective, which was 0.59. Hence, some of its items were moved to other perspectives, and five items were deleted. As a result, 52 and 50 items remained for inpatients and outpatients, respectively. The final English and Arabic questionnaire forms were ready for use. Since the research coincided with the COVID-19 pandemic, hospital approvals took six to nine months until they were received. Only 15 hospitals out of 18 agreed to participate. The data collection was performed between January and October 2021. The data from the pretest at one hospital were excluded. Next, we distributed 1000 questionnaires to the remaining 14 hospitals. As a result, 740 were returned (the response rate was 74%). The variation in the sociodemographic characteristics was a result of the maximum variation technique.

The statistical analysis using the Shapiro–Wilk test showed that the data were not normally distributed, so nonparametric tests were used. Then, construct validation was assessed for the instrument. EFA resulted in 46 items with loadings higher than 0.50 for 16 components. Eigenvalues for all components were higher than one. The KMO was 0.813, reflecting very high sampling adequacy (Ab Hamid et al., 2017; Kaiser & Rice, 1974), and Bartlett's test was also significant. The cumulative variance was 67.414%. The 12 components were BSCP-ATTs, PATIENT-CENT-EXR, services experience (SERV-EXR), price experience (PR-EXR), building experience (BUIL-EXR), ACC-EXR, PI of the complications (COMP-PI), technology perceived image (TECH-PI), information experience (INFO-EXR), HSRP-PI, and WT-EXR. One item loaded on the 12th component. However, this item had a higher loading on the BSCP-ATTs. None of the specific inpatient items had loadings higher than 0.50. Moreover, the scree plot showed the necessity of deleting the last three components.

The resulting nine components in EFA were tested in the Amos program. The model was edited based on the item loadings, model fit indices, and calculations in the convergent, discriminant, CR, IIC, and CITC at the next step until we arrived at the best model. First, adding two items that did not have loadings to the INFO-EXR construct showed good loadings in CFA. The same was true for the BSCP-ATTs and technology-perceived image (TECH-PI) constructs. Second, splitting the BUIL-EXR component into two separate constructs, building environment experience (BUILENV-EXR) and building capacity experience (BUILCAP-EXR), improved the item loadings and the model fit. Third, two items were removed from the PATIENT-CENT-EXR construct because they have loadings lower than 0.50. On the other hand, two items were added to the BSCP-ATTs construct since both had loadings higher than 0.50 and improved the model fit. Moreover, merging the TECH-PI and COMP-PI items at the BSCP-ATTs construct resulted in loadings lower than 0.5 and IIC lower than 0.30. Hence, three separate constructs in the attitude section were decided. Finally, the modification indices in the Amos program were utilized to improve the model. The final model revealed that the CMIN/df, CFI, GFI, TLI, RMSEA, and SRMR indices in CFA were above or close to the cutoff points, reflecting a good fit model. Nevertheless, the $P$ value was <0.001, which can be referred to as its sensitivity to normality.

The CRs for all factors were higher than 0.6 except for the SERV-EXR factor. However, this factor's IIC and CTIC were higher than 0.3. The other factors also had IICs higher than 0.3, and their CITC ranged from 0.328-0.853, reflecting satisfactory IIC and CITC.

The convergent validity was less than 0.5 for BSCP-ATT, BUILENV-EXR, SERV-EXR, and COMP-PI. However, the CR, IIC, and CITC showed satisfactory results (Fornell & Larcker, 1981), except for the SERV-EXR, which had a CR equal to 0.50 but an IIC and CITC higher than

0.3. On the other hand, the square roots of the AVE were higher than the off-diagonal correlations between factors. Additionally, a lower correlation between factors indicates each factor's uniqueness. The correlations between the independent factors were either negligible or low, except between two factors, the PATIENT-CENT-EXR and INFO-EXR, which were moderate. Merging the two constructs lowered the loadings and the model fit indices in CFA.

The same was perceived regarding merging the BUILENV-EXR and BUILCAP-EXR constructs. Consequently, separate factors were determined, as mentioned earlier. Regarding the independent factors, negligible or low correlations existed among them. Neither high nor very high correlations existed between the independent factors. Therefore, this establishes discriminant validity and the uniqueness of the independent factors. The same holds for the dependent factors.

## Discussion

In agreement with this paper's aim, it was possible to build a valid and reliable instrument. BSC-PATIENT is the first validated instrument to engage patients in the evaluation of hospitals by measuring their experiences and attitudes toward the hospital based on the BSC perspectives: the financial, internal, knowledge and growth, customer, and external perspectives. The deployment of this instrument at BSC implementations and PEs, in general, will improve patient satisfaction and allow a better understanding of BSC strategic maps based on patients' experiences and attitudes. In general, this study has several strengths. First, the BSC-PATIENT is the first instrument that engages patients in BSC perspective assessment. Second, this instrument can determine PATIENT-ATTs based on BSC perspectives. Third, to our knowledge, this is the first study to distinguish between PATIENT-EXR and PATIENT-ATTs, which will allow us to examine the relationship between PATIENT-EXR and PATIENT-ATTs in future studies. Fourth, this instrument was customized to be used for all insurance, leadership, and admission statuses. Fifth, this instrument was designed based on KPIs extracted from BSC implementations in primary, secondary, and tertiary healthcare settings in low-, middle-, and high-income countries worldwide. Hence, the implementation of BSC-PATIENT can be generalized to different healthcare settings and countries. However, the instrument may need some customization based on the healthcare setting strategy and the country's properties. For example, we customized the BSC-PATIENT from the environmental perspective based on Palestinian culture, the financial perspective based on administrative type, the knowledge and growth perspective based on the health information system in Palestine, and a few items specific for inpatients based on admission status. Finally, no study has assessed Palestinian hospital performance during this era, so using BSC-PATIENT will offer a comprehensive hospital assessment from patient perspectives during COVID-19. However, this instrument has some limitations. Despite this instrument assessing items such as patient education on IC, it lacks COVID-19-specific items, as this instrument was designed before the COVID pandemic, so COVID-19-related items can be considered in future versions of the BSC-PATIENT instrument. Second, patient literacy was not assessed. However, academic qualifications were evaluated at the demographic level to be considered in the analysis. Third, measuring PATIENT-EXR in the past may involve a bias of recall. Additionally, participant bias may have occurred since the sample was convenient and the included hospitals agreed on participation. However, the high percentage of the included hospitals (30%) from the total number of hospitals at West Bank and including all administrative type types from all regions may have reduced the selection bias. Another limitation is that we could not validate this instrument in English due to our inaccessibility to English-speaking patients. Future research needs to consider testing the psychometric properties of the BSC-PATIENT in an English-speaking country.

## Conclusion

The BSC-PATIENT instrument was developed to enhance PATIENT-ENG in the PE of hospitals. This instrument was validated in Arabic and customized for Palestinian hospitals. This is the first instrument to engage patients in evaluating their experiences and attitudes toward the BSC perspective. It consists of 38 items: 21 items assessing patient experience observations and 15 items assessing PATIENT-ATTs. Both experiences and attitudes were designed based on BSC perspectives. The findings of this research showed adequacy in the psychometric properties of this instrument and suggest some recommendations for future research. First, we tested the psychometric properties of the BSC-PATIENT in English and other languages in different countries. Second, we consider it vital to engage other stakeholders, such as doctors, nurses, and managers, in the PE of hospitals' BSC perspectives. Third, this instrument can be used to assess the impact of PATIENT-EXR on PATIENT-ATTS toward the hospital, specifically the PI, PQ, satisfaction, and LOY-ATTs. Fourth, managers must consider using a comprehensive approach for the PE of hospitals instead of limiting it to financial or internal indicators. Fifth, BSC-PATIENT will allow comparing the differences in patient experience and attitudes based on patient and hospital characteristics. Finally, enhancing PATIENT-ENG in the evaluation process instead of focusing on satisfaction alone must be considered in future BSC and PE implementations. Involving stakeholders in the evaluation of the BSC will lead to a better and deeper understanding of hospital PE.

## (Sub-study 4): How to engage HCWs in the evaluation of hospitals: development and validation of BSC-HCW1 - a cross-sectional study[4]
## Introduction

The analysis of the BSC implementations in our two previous systematic reviews (Amer et al., 2022a, 2022b) revealed that most of the implementations focused on measuring HCWs' satisfaction as a sole indicator. Second, although strategic maps were utilized based on hospital record data in BSC implementations, there has been a lack of analysis of the factors that impact or predict HCWs' satisfaction based on HCWs' opinions and observations (Amer et al., 2022b). Third, although the researchers have pointed to the importance of patient and HCW-ENG in the process of PE and delivery improvement (Anderson et al., 2021; Gagliardi et al., 2008; Korlén et al., 2018), the reviews (Amer et al., 2022a; Bohm et al., 2021) revealed that there had been a lack of engaging stakeholders in BSC implementations, such as engaging patients and HCWs. Based on the review (Amer et al., 2022b), we recommend that HCW-ENG in BSC implementations might provide a solution to the issue of stagnant levels of satisfaction among HCWs in BSC implementations. In addition, the participation of HCWs will aid HCO managers and researchers in their efforts to obtain a better grasp of the BSC strategic maps as well as the causal relationships between KPIs based on the perspectives of HCWs. Moreover, we think that the participation of HCWs in BSC implementations will result in an even greater improvement in both the financial

---

[4] This chapter is based on the following paper: **Amer et al**., 2022. How to Engage Health Care Workers in the Evaluation of Hospitals: Development and Validation of BSC-HCW1—A Cross-Sectional Study. International Journal of Environmental Research and Public Health. 19, 9096. https://doi.org/10.3390/ijerph19159096. **Impact factor: 4.614**

performance and the level of satisfaction perceived by patients. We also found that in a manner similar to the inadequate emphasis placed on HCWs' satisfaction during BSC implementations in HCOs, the notion LOY-ATT of HCWs was rarely taken into consideration. (Amer et al., 2022a, 2022b). A review in the business field (Curtis et al., 2011) found a strong positive relationship between satisfaction and LOY-ATT. However, these relationships were found to be moderated by different factors, such as demographics and setting type. We think that understanding HCWs' LOY-ATTs may assist hospital managers in expecting HCWs' future behavior. This will provide insight to managers when evaluating their hospitals' performance, building their plans, and allocating their resources.

In this research, the first aim is to develop an instrument that performs HCW-ENG in a comprehensive assessment of BSC perspectives and dimensions (BSC-HCW1). The second aim of this research is to customize the developed instrument at Palestinian hospitals, translate it into Arabic, and validate it.

## Methods

BSC-HCW1 was developed using the previously reported methodology for BSC-PATIENT development (Amer et al., 2022c) with HCW adaptation. As a consequence, 58 items remained. We calculated the I-CVI, the S-CVI, and the CVI-UA (Zamanzadeh et al., 2015) to examine the content validity per item and scale. The same methodology used for the linguistic validation of BSC-PATIENT was also used for BSC-HCW1. Internal consistencies of the instrument's perspectives in the initial edition of the questionnaire were evaluated. The first version of the questionnaire was pretested on 30 HCWs in one NGO hospital in the south of the West Bank. As a consequence, few elements were changed or removed. The same sampling procedure and HCO sample used to produce BSC-PATIENT (Amer et al., 2022c) was also used to develop BSC-HCW1. Therefore, a total of 800 questionnaires were distributed. The same ethical considerations, IRB of BSC-PATIENT (Amer et al., 2022c), and hospital approvals were also considered and used for BSC-HCW1. The same data collection process used in BSC-PATIENT (Amer et al., 2022c) was also used. The inclusion and exclusion criteria were established as a Palestinian doctor or nurse of either gender who had worked at the examined hospital for at least three months. The included departments were emergency, internal medicine, surgery, gynecology, and pediatrics. The HCWs were conveniently selected in this study based on their presence at the departments during the data collection. The same statistical analysis of BSC-PATIENT (Amer et al., 2022c) was used.

## Results

The section on HCWs' sociodemographics included age, gender, profession type, working department, years of experience, and total monthly income. Moreover, the questionnaires were coded based on the hospital name, administrative type, location, and JCI accreditation. The second section of the questionnaire was designed to evaluate HCWs' satisfaction predictors based on BSC perspectives and to directly measure their LOY-ATTs. The pretest revealed acceptable Cronbach's alpha for the instrument, which was 0.94, as well as for its construct. The final questionnaire forms in English and Arabic were completed and ready to be used. We delivered 800 questionnaires at 14 hospitals, out of which 454 valid questionnaires were retrieved (the response rate was 57%). The characteristics and socio-demographics of the respondents reflected the maximum variation technique.

The data were not normally distributed. Therefore, nonparametric tests, specifically Spearman correlations, were chosen in the following steps. EFA for the 51 items resulted in 35 item loadings higher than 0.50 for 15 components. All the components had eigenvalues greater

than one. The KMO was 0.832 with a significant Bartlett's test, indicating a high level of sample adequacy (Ab Hamid et al., 2017; Kaiser & Rice, 1974). The total variation was 66.72%. The 15 components were TECH, health care workers' scientific development (HCW-SCIDEV), MANAG-PE, workload time- life balance (WTLB), LOY-ATTs, quality of supplies and services (QSS), financial incentives (FINI), HCW-ENG, REPUT, MANAG-COMM, ACC-EXR, introductory period (ITRODP), safety, and no blame error reporting (NBR). However, no item had a loading higher than 0.5 on the 15th component. The scree plot results confirmed only 10 components out of 15, so these 10 were tested in the next step.

CFA was performed for the resulting ten components in EFA. The CMIN/DF was 1.966. However, the other model fit indices were CFI= 0.885, GFI= 0.841, TLI= 0.860, RMSEA= 0.064, and SRMR= 0.0692, with a significant $P$ value. Hence, in the next phase, the model was tweaked based on the item loadings, model fit indices, and computations in the convergent, discriminant, CR, IIC, and CITC until the optimal model was reached. For example, the ESC4 item was removed from the MANAG-COMM and was covered with a single-item construct measuring managerial trust (MTR). Additionally, the REPUT component was converted to the PTR construct. Two items were moved to the LOY-ATTs construct. Moreover, items with loadings less than 0.5 were also removed or relocated to other constructs on which they had better loadings. Moreover, two items were added to the MANAG-PE construct. Two constructs, QSS and HCW-SCIDEV, were merged into one construct: quality and development (QUALDEV). This was due to the very high correlation between them. This merging also increased the fitness of the model. Finally, eight modification indices were utilized to improve the fit of the model. As a result, the optimal model consisted of nine constructs. The CMIN/DF was 1.334. Additionally, the other model fit indices were CFI= 0.958, GFI= 0.875, TLI= 0.948, RMSEA= 0.041, and SRMR= 0.0557. However, the $P$ value was significant.

The CRs for all factors were higher than 0.6. Additionally, all factors' IIC and CTIC were higher than 0.3. The IIC ranged from 0.334-0.703, and the CITC ranged from 0.466-0.729, reflecting satisfactory internal consistency. For the five factors MANAG-PE, HCW-ENG, QUALDEV, WTLB, and LOY-ATTs, the convergent validity was between 0.4 and 0.5. However, the CRs for all were greater than 0.6, indicating acceptable convergent validity (Fornell & Larcker, 1981). Correlations between the independent factors were insignificant or low in this context, except for the moderate association between the MANAG-PE factor and HCW-ENG. No high or very high correlations were found between factors. On the other hand, the square roots of the AVE were higher than the off-diagonal correlations between factors. In other words, convergent and discriminant validity were fulfilled for all factors.

## Discussion

In line with this paper's aim, we developed, translated, and validated the BSC-HCW1 instrument to perform successful HCW-ENG in the evaluation process of BSC perspectives: the financial, internal, knowledge and growth, customer, external, and managerial perspectives. Our findings showed that the final model of BSC-HCW1 resulted in nine factors.

The BSC-HCW1 has several strengths. First, it is the first validated instrument designed to apply HCW-ENG in a comprehensive assessment of the following BSC perspectives: financial, customer, internal process, knowledge and growth, external, and managerial. Second, this is the first validated instrument to conduct PE for Palestinian hospitals based on HCWs' opinions and observations. BSC-HCW1 will help PMOH and health policymakers improve the performance of the health sector and overcome many challenges. For example, there is a lack of existing data measuring such KPIs in the records of many Palestinian hospitals. Additionally, there was a lack

of transparency and the unwillingness of many hospitals to share the data extracted from their hospital records externally. The success in using the BSC-HCW1 in the Palestinian healthcare context, which is characterized by fragility and fragmentation both geographically and administratively, may indicate that this instrument can be utilized successfully in other hospitals in low- to middle-income countries or countries that reside under complex situations. Finally, the BSC-HCW1 will solve the heterogenicity in KPIs that were perceived in the previous BSC implementations and will offer a uniform assessment. This will facilitate PE comparisons among hospitals based on area and administrative style. It will also enhance data sharing among hospitals and recommendations among researchers, which will lead to improving hospital performance and a better understanding of HCWs' LOY-ATTs predictors worldwide.

On the other hand, this instrument has some limitations. First, the external perspective dimensions were ultimately excluded during the validation process. A refinement of these perspective items may be included in future versions of the BSC-HCW1. Second, this instrument is solely intended for use by two specific categories of HCWs: physicians and nurses. Both categories are important, as they spend the majority of their time with patients and are ultimately in charge of providing care. However, other categories of HCWs who work in hospitals, such as technicians, pharmacists, and nonclinical HCWs, were not included in this study. Therefore, future versions to include these categories can be beneficial. Third, despite the validation of this instrument during the pandemic, it was developed before it, so it lacks essential items. For example, the assessment of personal protective equipment availability at hospitals during the pandemic. It also lacks an assessment of customer-related variables in this era, such as HCWs' stress and fear and items related to the development and knowledge pertaining to COVID-19 updates. Therefore, it is recommended to consider adding such items to future versions. Moreover, it is recommended to include items that measure types of burnout other than emotional exhaustion from the customer perspective. Additionally, it is advised to include family-related factors and marital status in the instrument since they may work as modifiers for HCWs' LOY-ATTs. Moreover, we recommend adding items that assess motivation, work control, work stability, access to resources, and prior achievements since they may be predictors of HCWs' satisfaction. Furthermore, some HCWs noted that they were hesitant to provide negative feedback regarding their managers' performance, which may have biased the responses. However, all respondents were informed of the consent form's anonymity and privacy to lower this bias. Additionally, this was explained to them verbally by the data collectors. Additionally, participant bias may have occurred since the sample was convenient and the included hospitals agreed to participate in the research. Nevertheless, the high percentage of the included hospitals (30%) from the total number of hospitals at West Bank and including all administrative style types from all regions may have reduced the selection bias. Another limitation is that due to our inability to access English-speaking patients, we could not verify this instrument in English. Future studies should include the psychometric properties of the BSC-HCW1 in an English-speaking country. Last, because of the vast number of KPIs, the developers of this instrument have decided to only include those dimensions that are directly relevant to the demands of HCWs from each BSC perspective. The development of the second version of BSC-HCW1 that adds the unrelated dimensions to HCWs' demands at each BSC perspective has the potential to significantly improve the level of -ENG in the PE of their hospitals and BSC implementations.

# Conclusion

Researchers and hospital administrators who want to adopt the BSC in hospitals may benefit from utilizing the BSC-HCW1. This instrument might help understand the performance of

the perspectives and dimensions of the BSC based on the opinions and observations of HCWs. Most BSC implementation studies did not include HCWs at all or included them simply to gauge their level of satisfaction. Additionally, HCWs' LOY-ATTs were rarely taken into account. None of the BSC implementations were able to get the HCWs to participate in the process of evaluating the perspectives and dimensions of the BSC. The BSC-HCW1 is the first instrument that has been designed specifically to include HCWs in the process of conducting PE using BSC perspectives and dimensions. BSC-HCW1 might let hospital managers look at BSC strategic maps based on what HCWs have observed and what they think. Therefore, it is strongly recommended that researchers make use of BSC-HCW1 in any future BSC implementations. Another study is needed to produce a second version of this instrument that utilizes HCW-ENG in evaluating the BSC dimensions that are not directly relevant to their needs but are nonetheless related to the PE of HCOs. In addition to HCWs, other stakeholders, such as patients and hospital administrators, must be included in the implementation of BSCs. Palestinian health policymakers and hospital management will be able to assess their strengths and shortcomings based on the observations and views of their HCWs using this instrument. It is possible to make use of this validated instrument in its Arabic form in other Arab nations. However, validation in more languages is still required for this instrument.

# A summary of novel findings

## - The first sub-study
The results showed that HCWs' satisfaction and the RoB need to be further improved in future BSC implementations.

## - The second sub-study
1. This review solved the dilemma of the KPI categorization difference in BSC implementations, yielding more precise results. The resulting BSC perspectives were financial, customer, internal process, external, knowledge and growth, and managerial, under which 13 major dimensions and 45 subdimensions were defined.
2. This review calculated the use frequency of perspectives and the weights/importance assigned to them. The most frequently used perspectives in BSC papers were internal, financial, patient, learning and growth, HCW, managerial, community, and stakeholder perspectives. The perspectives that had the highest importance were internal, financial, learning and growth, patient, HCW, community, managerial, and stakeholder.
3. This review found a lack of PATIENT-ENG and HCW-ENG in BSC implementation. Additionally, LOY-ATTs of patients and HCWs were rarely taken into account in BSC implementations.

## - The third sub-study
1. BSC-PATIENT is the first validated instrument designed to engage patients in BSC perspectives' PE.
2. This instrument was validated in Arabic and customized for Palestinian hospitals with adequate psychometric properties.

## - The fourth sub-study
1. The BSC-HCW1 is the first validated instrument designed to engage HCWs in a comprehensive assessment of BSC perspectives: financial, customer, internal process, knowledge and growth, external, and managerial perspectives based on the opinions and observations of HCWs.

2. The BSC-HCW1 is the first validated instrument to conduct PE for Palestinian hospitals based on HCWs' opinions and observations.

# Theoretical and practical implications

The results of the core studies in this dissertation have implications that resulted in drawing recommendations for HCO managers and policymakers.

- **The first implementation (theoretical):**[5]
A study in which we proposed theoretical implementation solutions to cancer care crises in Palestine based on BSC perspectives.

- **The second Implementation (theoretical):**[6]
*Figure 1* is a theoretical implication to assess the 13 major dimensions and 45 subdimensions that resulted in chapter 4 during the COVID-19 era. The evaluation was performed based on a rapid analysis by searching for independent studies in Google Scholar and the Google search engine during the COVID-19 pandemic until June 2021 (Amer et al., 2022a).



**Figure 1. An implication of BSC dimensions. An assessment of the resulting 13 major dimensions and 45 sub-dimensions in the COVID-19 era.**

---

[5] This analysis was published in: **Amer F.**, 2022. Al-Nawati tragedy: a 16-year-old patient with leukaemia and no access to cancer care. Lancet Oncology; 23(4):447–9. https://doi:10.1016/S1470-2045(22)00091-2. **Impact factor: 54.433**

[6] This analysis was published as short report preprint https://doi.org/10.21203/rs.3.rs-1970297/v1

- **The third implementation (practical):**[7]

An implication using the BSC-PATIENT instrument (Amer et al., 2022c). The variance analysis based on admission status revealed that PATIENT-EXR and PATIENT-ATTs for outpatients need improvement. The multiple regression and path analysis provided strong evidence for the INFO-EXR impact on patients' attitudes. Palestinian health policymakers must prioritize the design and delivery of patient education programs into their action plans and encourage two-way information communication with patients. Strong evidence for the roles of PATIENT-CENT-EXR, SERV EXR, and BUIL-EXR in improving BSCP-ATT was found. Recommendations for Palestinian health policymakers based on this implementation include developing a formal training plan for healthcare workers to improve the information provided to patients upon their admission and discharge, including oral and written information. Second, healthcare workers should invest in formal training to improve patients' education, such as education on infection control measures. Third, improving the receipt of information and feedback from patients through the distribution of surveys. Additionally, serious consideration to solve patients' complaints is encouraged. Fourth, increasing the variety of specialties and departments available at Palestinian hospitals, as well as the availability of medical services at night, on vacations, and weekends. Additionally, ensuring the availability of female doctors and nurses in all departments is a demand that can be referred to in Palestinian culture. Fifth, building dimensions, including the environment, such as the cleanliness, infrastructure, and capacity of departments, should be improved. Many patients reported that the number of chairs in the waiting area had to be increased in Palestinian hospitals. Sixth, improving Palestinian outpatients' experiences related to patient care, services, and accessibility, as well as outpatient attitudes toward balanced scorecard perspectives and dimensions. Seventh, engaging patients in hospital performance evaluations by utilizing the developed instrument. This should be carried out routinely to monitor the change and improvement in the quality of health services from patients' observations.

- **The fourth implementation (practical):**[8]

A practical implication of the BSC-HCW1 instrument (Amer et al., 2022d). The variance analysis revealed no difference between doctors' and nurses' evaluations. The multiple regression and path analysis provided evidence of the importance of improving HCWs' WTLB, QUALDEV, and MANAG-PE in improving the LOY-ATTs of HCWs. All factors revealed PE gaps. Recommendations for Palestinian health policymakers based on this implementation include reviewing the system of the financial incentive and linking it with healthcare workers' appraisals and achievements. Second, healthcare workers should be trained and counseled on how to improve their time management and workload time-life balance. Third, they should invest in action plans

---

[7] This study is published as: **Amer et al**., 2022. Engaging patients in balanced scorecard evaluation - An implication at Palestinian hospitals and recommendations for policy makers. Front. Public Health. 10:1045512. doi://10.3389/fpubh.2022.1045512
**Impact factor: 6.461**

[8] This study is published as preprint: **Amer F.**, 2022. Engaging physicians and nurses in balanced scorecard evaluation - An implication at Palestinian hospitals and recommendations for policy makers. PREPRINT (Version 1) available at Research Square:
https://doi.org/10.21203/rs.3.rs-2235199/v1

on how to increase the time that healthcare workers spend with their patients. Fourth, continuous educational programs should be planned and executed to update healthcare workers with information regarding diseases and medication related to their fields. Future utilization of mobile health for such purposes is recommended. Fifth, perform a periodic evaluation of available equipment that requires maintenance or replacement. Additionally, investments in electronic decision support systems can improve the quality and development factor. Sixth, monitoring the performance of healthcare workers and designing an appraisal system that explains their strengths and weaknesses. Moreover, to discuss with them how to utilize their strengths and what actions or development programs are needed to improve their weaknesses. Seventh, managers must strengthen healthcare workers' engagement in planning and decision processes. Eighth, the managerial early awareness of the high-risk groups who intend to leave their jobs and invest in improving their experiences encourages their loyalty attitudes, such as the improvement of workload time-life balance, quality and development initiatives, managerial performance, healthcare workers engagement, and financial incentives. Ninth, focusing on improving the factors that affect the respect of healthcare workers' direct managers, particularly their engagement, managerial performance, and loyalty attitude. Finally, the factors that affect perceived patient respect should be improved, particularly quality and development initiatives, healthcare workers' workload time-life balance, loyalty attitudes, financial incentives, and managerial performance evaluations.

# List of publications during Ph.D.

## Published full-text articles related to the dissertation

1. **Amer F,** Hammoud S, Khatatbeh H, Lohner S, Boncz I, Endrei D (2022). The deployment of balanced scorecard in health care organizations: is it beneficial? A systematic review. BMC Health Services Research. 22(1), 1–14. https://doi: 10.1186/s12913-021- 07452-7. **Impact factor: 2.908**

2. **Amer F**, Hammoud S, Khatatbeh H, Lohner S, Boncz I, Endrei D (2022). A systematic review: the dimensions to evaluate health care performance and an implication during the pandemic. BMC Health Services Research. 22(1), 621. https://doi.org/10.1186/s12913-022-07863-0. **Impact factor: 2.908**

3. **Amer F**, Hammoud S, Onchonga D, Alkaiyat A, Nour A, Endrei D, Boncz I (2022). Assessing Patient Experience and Attitude: BSC-PATIENT Development, Translation, and Psychometric Evaluation—A Cross-Sectional Study. International Journal of Environmental Research and Public Health. 19(12):7149. https://doi.org/10.3390/ijerph19127149. **Impact factor: 4.614**

4. **Amer F**, Hammoud S, Khatatbeh H, Alfatafta H, Alkaiyat A, Nour A. I, Endrei D, Boncz I (2022). How to Engage Health Care Workers in the Evaluation of Hospitals: Development and Validation of BSC-HCW1—A Cross-Sectional Study. International Journal of Environmental Research and Public Health. 19(15): 9096. https://doi.org/10.3390/ijerph19159096. **Impact factor: 4.614**

5. **Amer F**, Neiroukh H, Abuzahra SE, AlHabil Y, Afifi M, Shellah D, Boncz I, Endrei D (2022). Engaging patients in balanced scorecard evaluation - An implication at Palestinian hospitals and recommendations for policy makers. Front. Public Health. 10:1045512. doi://10.3389/fpubh.2022.1045512 **Impact factor: 6.461**

6. **Amer F** (2022). Al-Nawati tragedy: a 16-year-old patient with leukaemia and no access to cancer care. Lancet Oncology. 23(4):447–9. https://doi:10.1016/S1470-2045(22)00091-2. **Impact factor: 54.433**

## Additional published full-text articles "health sciences"

1. **Amer F**, Hammoud S, Farran B, Boncz I, Endrei D (2021). Assessment of Countries' Preparedness and Lockdown Effectiveness in Fighting COVID-19. Disaster Medicine and Public Health Preparedness. 15(2): e15-e22. https://doi:10.1017/dmp.2020.217. **Impact factor: 5.556**

2.  (This article was cited by the World Health Organization-Western Pacific and the European Commission in their guidelines to indicate the role of lockdown in mitigating the pandemic consequences).

3.  Hammoud S, Onchonga D, **Amer F**, Kocsis B (2021). The Burden of Communicable Diseases in Lebanon: Trends in the Past Decade. Disaster Medicine and Public Health Preparedness. 1–3. https://doi.org/10.1017/dmp.2021.200. **Impact factor: 5.556**

4.  Hammoud S, **Amer F**, Lohner S, Kocsis B. (2020). Patient education on infection control: A systematic review. American Journal of Infection Control. 48(12): 1506–1515. https://doi.org/10.1016/j.ajic.2020.05.039. **Impact factor: 4.303**

5.  Biancovilli P, Makszin L, **Amer F**, Csongor A (2022). Celebrities and Breast Cancer: A Multidimensional Quali-Quantitative Analysis of News Stories Shared on Social Media. International Journal of Environmental Research and Public Health. 19(15):9676. https://doi.org/10.3390/ijerph19159676. **Impact factor: 4.614**

6.  Hammoud S, **Amer F**, Khatatbeh H, Alfatafta H, Zrínyi M, Kocsis B (2022). Translation and validation of the Hungarian version of the Infection Control Standardized Questionnaire: A Cross-sectional study. BMC NURSING. 21 (244):1-12. https://doi.org/10.1186/s12912-022-01024-8. **Impact factor: 3.198**

7.  Hammoud S, **Amer F**, Kocsis B (2021). Examining the Effect of Infection Prevention and Control Awareness among Nurses on Patient and Family Education: A Cross-sectional Study. Nursing & Health Sciences. 24(1), 140-151. https://doi.org/10.1111/nhs.12905. **Impact factor: 2.214**

8.  Onchonga D, Várnagy Á, **Amer F**, Viktoria P, Wainaina P (2021). Translation and validation of the Swahili version of the Wijma Delivery Expectancy/Experience Questionnaire version A (W-DEQ-A). Sexual & Reproductive Healthcare. 29, 100626. https://doi.org/10.1016/j.srhc.2021. **Impact factor: 2.194**

9.  Khatatbeh H, Al-Dwaikat T, Oláh A, Onchonga D, Hammoud S, **Amer F**, Prémusz V, Pakai A. (2021). The relationships between pediatric nurses' social support, job satisfaction and patient adverse events. Nursing Open. 8(6): 3575-3582. https://doi.org/10.1002/nop2.907. **Impact factor: 1.942**

10. Khatatbeh H, Pakai A, Al-Dwaikat T, Onchonga D, **Amer F**, Prémusz V, Oláh A (2021). Nurses' burnout and quality of life: A systematic review and critical analysis of measures used. Nursing Open. 9(3): 1564-1574. https://doi.org/10.1002/nop2.936. **Impact factor: 1.942**

11. Alfatafta H, Alfatafta M, **Amer F**, Hammoud S, Zhang L, Molics B, Boncz I (2022). Quality of life of patients with severe knee Osteoarthritis in Hungary. STUDIA UNIVERSITATIS BABES-BOLYAI EDUCATIO ARTIS GYMNASTICAE. 67(2): pp. 5-15 https://doi.org/10.24193/subbeag.67(2).09.

## H-index and citation metrics

**MTMT:** total citations= 95, independent citation= 55
**SCOPUS:** citations= 66, H-index= 5
**Web Of Science:** citations= 57, h-index= 5, peer-reviews= 15
**Google Scholar:** citations= 122, H-index= 6, i10-index= 5
**Research Gate:** citations= 127, H-index= 7