The deployment of balanced scorecard in health care organizations: is it beneficial?  
A systematic review

Faten Amer1,2*, Sahar Hammoud1, Haitham Khatatbeh1, Szimonetta Lohner3, Imre Boncz2 and Dóra Endrei2

Abstract
Background: Balanced Scorecard (BSC) has been implemented for three decades to evaluate and improve the performance of organizations. To the best of the researchers’ knowledge, no previous systematic review has performed a comprehensive and rigorous methodological approach to figure out the impact of BSC implementation in Health Care Organizations (HCO).

Aims: The current work was intended to assess the impact of implementing the BSC on Health Care Workers’ (HCW) satisfaction, patient satisfaction, and financial performance.

Methods: The authors prepared the present systematic review according to PRISMA guidelines. Further, the authors customized the search strategy for PubMed, Embase, Cochrane, Google Scholar databases, and Google’s search engine. The obtained studies were screened to isolate those measuring scores related to HCW satisfaction, patient satisfaction, and financial performance. The Risk of Bias (RoB) in the non-Randomized Intervention Studies (ROBINS-I) tool was used to assess the quality of observational and quasi-experimental studies. On the other hand, for the Randomized Controlled Trials (RCTs), the Cochrane (RoB 2) tool was used.

Results: Out of 4031 studies, the researchers included 20 studies that measured the impact of BSC on one or more of the three entities (HCW satisfaction, patient satisfaction, and financial performance). Throughout these 20 studies, it was found that 17 studies measured the impact of the BSC on patient satisfaction, seven studies measured the impact on HCW satisfaction, and 12 studies measured the impact on financial performance.

Conclusion: This systematic review provides managers and policymakers with evidence to support utilizing BSC in the health care sector. BSC implementation demonstrated positive outcomes for patient satisfaction and the financial performance of HCOs. However, only a mild impact was demonstrated for effects related to HCW satisfaction. However, it is worth noting that many of the studies reflected a high RoB, which may have affected the impacts on the three primary outcomes measured. As such, this systematic review reflects the necessity for further focus on this area in the future. Moreover, future research is encouraged to measure the real and current impact of implementing BSC in HCO during the pandemic since we did not find any.

Keywords: Performance, Evaluation, Assessment, Health, Hospital, Impact, Effect

*Correspondence: amer.faten@etk.pte.hu; faten.pharmix@gmail.com
1 Institute for Health Insurance, Faculty of Health Sciences, University of Pécs, Pécs, Hungary
Full list of author information is available at the end of the article

Introduction
Since its development by Norton and Kaplan in 1992 [1], the Balanced Scorecard (BSC) has been utilized by many health care managers for the performance evaluation (P.E.) of Health Care Organizations (HCO) worldwide.
Moreover, BSC can also be used as a strategic managerial tool by linking it to the organization's strategy [2].

The first generation of the BSC was used to evaluate four organizational perspectives: the financial perspective, the customers' perspective, the internal processes' perspective, and finally, the learning and growth perspective, all of which were steered by the organizational vision and strategy [1]. See Fig. 1. In this regard, it should be noted that a recent review re-categorized the BSC four perspectives for HCO into further 45 sub-dimensions or categories [3]. In the second generation of BSC, strategic maps were added to describe the cause-effect relationships between strategic objectives of each perspective [4]. In the third generation of BSC, destination statements, measures, and action plans were added to achieve the intended targets [5]. It is worth mentioning that Duke Children's Hospital in the United States of America (USA) was the first HCO to implement BSC in 2000. As a result, the hospital was able to convert 11 million dollars of losses into four million dollars of profits [6]. See Fig. 2, which shows Duke University's health system strategic map [7]. The BSC strategic maps show that the process flow of the cause-and-effect relationships ends with the customer and financial perspectives [4, 8].

More recently, the pandemic of Coronavirus Disease 2019 (COVID-19) imposed financial burdens on many countries and health care systems worldwide. In addition, the pandemic increased the psychological stress of patients and HCW [3, 9, 10]. Moreover, implementing tasks with particular standards and guidelines was vital in tackling the spread of COVID-19 [11]. However, conflicting managerial decisions and the lack of standardization capability were factors that brought dissatisfaction about HCO [12]. Consequently, it is deemed essential to evaluate previous BSC implementation's effectiveness in HCO and determine if actual benefits remain to justify their continued use. It is worth noting that numerous systematic reviews regarding the impact of BSC implementation in non-health care-related fields, such as architecture [13], management, marketing, and accounting, already exist [14]. However, BSC reviews in health care only described the application of BSC or its perspectives [3, 15]. They lacked evaluation of the effectiveness of using BSC in HCO through systematically reviewed and evaluated literature. In this regard, only two reviews discussed the impact of BSC; one discussed the impact qualitatively [16], while the other mentioned a few examples of the positive impact [8]. To reiterate, this indicates that, until now, no comprehensive or rigorous methodological approach to assess the impact of BSC implementation in HCO has been recorded. Based on this gap in the literature, it was deemed essential to determine whether the previous BSC implementations at HCO were beneficial.

Thus, the present systematic review aims to gather all studies which have measured the impact of implementing BSC on HCW satisfaction, patient satisfaction, and

![Fig. 1 Balanced Scorecard Perspectives [1].](image)
financial performance at HCO; particularly, since these three attributes represent the latest affected perspectives in the strategic maps [4, 8]. Further, this review aims to assess and compare results among the included studies.

**Materials and methods**

Our previous systematic review analyzed the dimensions and indicators of BSC utilized at the P.E. of HCO [3]. This systematic review was carried out by finding all studies that approached BSC implementation's impact in HCO in adherence with the 27-point of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist [17]. See (S1 Appendix).

**Eligibility criteria**

The inclusion and exclusion criteria were set as shown in Table 1 below.

**Data sources, search strategy, and study selection**

In the present systematic review, the search strategy was developed by the first, second, and fourth authors; because the first two authors are experts in health care management and BSC, and the fourth author is an expert in systematic reviews and meta-analysis. The search strategy was initially developed for the PubMed database based on the PICO (Population, Intervention, Comparison, and Outcome) tool [18], and depending on using both MeSH (Medical Subject Headings) terms and keywords. See (Table 1). Next, the strategy was adapted to Cochrane Central Register of Controlled Trials (CENTRAL), Embase, and Google Scholar databases, as per Cochrane's recommendations [19]. The strategies developed for these databases can be found in (S2 Appendix).

The grey literature, pre-prints, and unpublished studies were searched on Google Scholar and Google's search engine websites to reduce publication bias. Furthermore, the authors 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, the first author conducted the search strategies on the electronic databases and removed the duplicates using the EndNote X9.2 program.

The first and second authors independently performed the selection of eligible studies. A discussion after each step was made or, if necessary, the 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 and analysis**

The following types of data were extracted from each of the final eligible studies: 1) author/s, year of publication, 2) country, 3) type of study, 4) duration of data
The research design of eligible studies was extracted directly from them. In case the research design was not explicitly mentioned, it was determined based on the role of the investigator of a given study. Specifically, if the BSC exposures were naturally determined and the investigator had no part, the study was considered observational. On the other hand, when the investigator assigned the BSC intervention, the study was deemed to be experimental.

The impact of BSC in eligible studies was explicitly mentioned or determined by calculating the difference between before and after implementation values. After that, the unification of the units was performed. Next, charts plotting for each outcome were performed by the first author and then reviewed by the first and second authors separately. If the impact measurement unit was not reported in the eligible studies, the authors of these studies were contacted. Lastly, all differences were compared, discussed, and judged by the two authors in the final step.

Quality assessment
The Risk of Bias (RoB) assessment was performed by the first and second authors independently between March and June 2021 to assess the quality of the included studies. As per the Cochrane collaboration's guidelines, the Cochrane (RoB 2) tool was used for the assessment of randomized controlled trials (RCTs) [20]. The Risk of Bias in non-Randomized Intervention Studies (ROBINS-I) tool was used to assess the observational and quasi-experimental studies [21]. As per the Cochrane Handbook, authors should avoid summarizing the overall RoB [22, 23]. Therefore, the RoB was analyzed at the study level and across studies.

In (RoB 2) tool, five types of bias were assessed: bias arising from the randomization processes, bias due to deviations from intended interventions, bias due to missing outcome data, bias in the measurement of outcomes, and bias in the selection of the reported results.

On the other hand, in the ROBINS-I tool, seven types of bias were assessed: bias due to confounding, bias in the selection of participants in a study, bias in measurement/classification of interventions/ exposures, bias due to deviations from intended interventions/ exposures, bias due to missing data, bias in the measurement of the outcomes, bias in the selection of the reported results.

While using the RoB 2 tool, each type of bias was assessed as low, high, or unclear. While using the ROBINS-I tool, each type of bias was evaluated into five...
categories: low, moderate, serious, critical, or no information. Afterward, the assessment results of the two reviewers were compared. Where there was disagreement, the fifth and sixth authors were consulted. Figures for RoB were prepared using the ROBVIS (Risk Of Bias VISualization) tool [24]. Lastly, it was recommended not to advocate quality appraisal as a criterion for inclusion in reviews [25]. Therefore, the authors decided to include all studies in this systematic review regardless of their quality assessment.

**Results**

**Study selection**

Initially, the search strategy resulted in a total of 4031 studies. After removing the 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 included studies’ full texts was made; based on this, only 20 studies were finally included in the current systematic review. Details of the study selection process are shown in the PRISMA flow-chart (Fig. 3).

**Study characteristics**

The main characteristics of the included studies are shown in Table 2.

**Location/country**

Regarding the implementation location, nine studies were implemented in North America, two in Europe, one in Africa, seven in Asia, and one did not specify the location. It should be noted that 14 studies were performed in high-income countries, two in upper-middle-income countries, one in a lower-middle country, and only two in low-income countries.
| Author/s, year of publication | Country | Design of the study | Duration of data collection | Setting | No. of health facilities | No. of participants | Data collection tool/ data sources |
|-------------------------------|---------|---------------------|----------------------------|---------|--------------------------|-------------------|----------------------------------|
| Harber, 1998 [26]            | Canada  | Experimental uncontrolled pretest-posttest | NR                          | Peel Memorial Hospital (Hospital in general + Laboratory) | 1                  | NR                  | NR                              |
| Meliones, 2000 [5]           | The USA- North Carolina | Experimental uncontrolled Interrupted time series | 1996–2000                  | Duke Children Hospital | 1                  | NR                  | NR                              |
| Pink et al., 2001 [27]       | Canada  | Observational prospective longitudinal | 1997–1998                   | Markham Stouffville hospital | 1                  | NR                  | Surveys + hospitals’ reports     |
| Gumbus et al., 2003 [28]     | USA     | Case Study/ observational retrospective longitudinal | 1999–2001                  | Bridge port hospital | 1                  | NR                  | The patient satisfaction measurement system |
| Smith & Kim, 2005 [29]       | USA     | Observational prospective and retrospective longitudinal | 2001–2004                  | Summa’s Food & Nutrition Service Department at Summa Health System (STH & ACH hospitals) | 2                  | NR                  | Press Ganey’s standard inpatient survey + audit checklists |
| Devitt et al., 2005 [30]     | Canada  | Observational prospective and retrospective longitudinal | 2004–2005                  | Toronto East General Hospital | 1                  | NR                  | Data extraction from hospital records |
| Yang & Tung, 2006 [31]       | Taiwan  | Retrospective longitudinal/ observational | 2000–2002                   | General hospitals & their supervisor agency | 21                 | NR                  | Secondary data from the department of health + primary data structured questionnaire measuring hospital’s organizational learning and growth perspective |
| Lorden et al., 2008 [32]     | NR      | Multimethod quantitative and qualitative case study/ experimental uncontrolled Interrupted time series | Jan, 1998– June, 2004       | Community hospital | 1                  | 300 Inpatient/quarter, 700 outpatient/quarter, 227 employees (1st survey), 191 employees (2nd survey) | Employee satisfaction survey + patient satisfaction survey (via email) |
| Josey & Kim, 2008 [33]       | USA-Ohio| Observational retrospective longitudinal | 2006                       | Barbent Citizens Hospital (BCH) | 1                  | NR                  | Patient satisfaction survey      |
| Chang et al., 2008 [34]      | Taiwan  | Observational retrospective longitudinal | 2001–2005                   | Mackay Memorial Hospital | 1                  | NR                  | NR                              |
| Hansen et al., 2008 [24]     | Afghanistan | Observational retrospective longitudinal | July to October of (2004/2005/2006) | Health facilities | >600               | 1700 HCWs, 5800 patients-provider interaction | NR |
| Chu et al., 2009 [35]        | Taiwan  | Case study/ experimental uncontrolled Interrupted time series | 2004–2006                   | The nursing department at a public teaching hospital in Taiwan | 1                  | 13 reference nurses’ group | Financial data from hospital + questionnaires to executives (the weights of indicators) |
Table 2 (continued)

| Author/s, year of publication | Country | Design of the study | Duration of data collection | Setting | No. of health facilities | No. of participants | Data collection tool/data sources |
|------------------------------|---------|---------------------|-----------------------------|---------|-------------------------|---------------------|----------------------------------|
| Edward et al., 2011 [36]     | Afghanistan | Observational retrospective longitudinal | 2004–2008 | Health facilities in Afghanistan | 700 | 1500 HCWs, 5000 patients | National health services performance Assessment + interviews with patients and HCWs |
| Fields & Cohen, 2011 [37]    | USA     | Experimental uncontrolled Interrupted time series | 2009–2010 | Oregon Health and science university family medicine (Clinics) | 1 | NR | Press Ganey survey for patient satisfaction + medical records. |
| Koumpouras, 2013 [38]        | Greece  | Case study/experimental uncontrolled Interrupted time series | 18 months but not specified when exactly | General Panarcadian Hospital of Tripolis | 1 | NR | Questionnaires and interviews |
| Smith et al., 2014 [39]      | Canada/Alberta | Experimental uncontrolled pretest-posttest | 2010–2011 (12-month trial), March 31, 2013 (results) | Hospitals in Alberta including hip and knee surgeries | 12 | NR | NR |
| Abdullah et al., 2014 [40]   | Indonesia | Cross-sectional/observational prospective and retrospective longitudinal | April–December, 2013 | Cibto Mangunkusumo Hospital-Digestive endoscopy center | 1 | 76 patients | Endoscopy reports + interviews based on structured questionnaires |
| Mutale et al., 2014 [25,41]  | Zambia  | Cluster randomized intervention/experimental RCT | 2011–2013 | Health facilities in Zambia | 12 | 96 HCWs, 429 patient interviews, 410 patient observations | A survey in facilities + interviews with HCWs and patients + patient observation + survey with households |
| Catuogno et al., 2017 [42]   | Italy   | Case study/Experimental uncontrolled pretest-posttest | 2007–2008, & 2014–2015 | Hematology department at a Research hospital in Italy | 1 | 14 | Stakeholder satisfaction; questionnaires + care processes + hospital discharge report + charity report + research process + departmental report + economic and financial hospital database + departmental report + charity report |
| Widyasari & Adi, 2019 [43]   | Indonesia | Descriptive Quantitative longitudinal/observational prospective and retrospective | During the year 2018 | Bali Mandara Hospital of Bali (Governmental hospital) | 1 | 30 | Participant observation + structured interviews + semi-structured interviews + documentation |

Note: NR Not Reported, HCWs Health Care Workers. Italics are designs based on our classification but not reported.
Setting
Out of the 20 selected studies, 16 were performed in hospitals or hospital departments, and four in health care facilities or clinics. See Table 2.

Language
Even though no limitation was imposed on language, all of the selected 20 studies measuring the impact of BSC implementation were written in English.

Study designs
Out of the 20 selected studies, only three studies reported their study designs explicitly. However, our classification showed that 11 studies were observational since the investigators were not involved in implementing BSC; instead, these investigators only observed the results of already implemented BSCs at HCO. On the other hand, the remaining nine studies were experimental. One out of the nine was RCT, while the other eight were quasi-experimental studies, which included three pretest-posttest components and five Interrupted Time Series. See Table 2. Notably, only three studies [32, 36, 37] randomly selected HCO, participants, or both.

Variations regarding the data collection instruments
Variances among the data collection instruments used in the 20 studies are shown in Table 2. Notably, the employed instruments were validated only in six studies [27, 28, 32, 35, 39, 42]. Additionally, only five studies [27, 28, 39, 42, 43] assessed the instruments’ feasibility. The pre-testing of the instruments was carried out only in three studies [27, 32, 35]. In addition, only five studies [27, 31, 38, 42, 43] assigned weights for the indicators or assessed their importance before implementation. Further, only one study [29] evaluated the indicators depending on more than one source for the same variable.

BSC’s impact type
The included studies assessed different outcomes for implementing BSC. Out of the final 20 eligible studies, 17 studies [6, 26–29, 31–38, 42–44] measured the impact of BSC on patient satisfaction, 7 measured HCW satisfaction [27–30, 36, 37, 45], and 12 studies measured financial performance [6, 27–30, 33, 34, 40, 41, 43–45]. However, the measured variables varied among studies, even in terms of the same dependent variable (Figs. 4, 5, 6, 7). For example, BSC’s impact on patient satisfaction varied from overall satisfaction to the satisfaction of specific categories, such as adults, children, inpatients, outpatients, patients in the emergency room, patients in rehabilitation. In addition, the measured variables varied based on the service type, such as satisfaction of home care services and departmental services.

Regarding HCW satisfaction, the name assigned to the targeted population varied from staff, employees to HCW. Further, the HCW satisfaction type varied, for instance, from HCW satisfaction towards their job to HCW satisfaction towards their superiors. However, the financial variable had the greatest variation among all three primary outcomes measured. Specifically, it was found that there exists a reduction in costs, expenditures, HCW budget, expenses, catering expenses, expenses/net revenues, bad debt expenses per net revenue, and supply per net revenue. On the other hand, an increase in revenues types included; returns, profits, aggregate surplus, funds, the value of drug-related groups, and return on assets.

Moreover, the unit used for financial impact assessment differed among studies. For example, all studies used currencies for assessment, where these currencies also varied between studies, except for a few studies [27, 33, 34] which used a percentage method. As an attempt to reduce bias, all currencies were converted to United
to standardize and make the comparison across studies more consistent regarding the financial outcomes in the systematic review. Further, the authors of one study [28] were contacted for clarification since they did not report the currency. As a result, Figs. 6 and 7 were designed as seen below; one for the impact in currencies and the other for the impact in percentages.

Most studies used a percentage score to measure the impact on patient and HCW satisfaction, except for three studies [28, 30, 45], which performed the measurement based on a four or five point-Likert scale. However, to make the comparison consistent, all Likert scales were converted to percentages (scores out of 100%). It should be noted that only two of the 20 studies [35, 45] discussed the statistical significance of the results. Hence, the magnitude of change (in percentage) was taken into consideration in our analysis.

**Time of measuring the outcomes**

To make an objective comparison and avoid falling into bias, we reported the time between implementing BSC and assessing its impact in the included studies. Further analysis demonstrated that the intervals between the initiation of BSC implementations and the measurements of impacts varied among studies. In particular, one study [29] reported the results based on 18 months of implementation, two studies [38, 44] took 2 years, one study [45] took 3 years, two studies [6, 35] took 4 years, and one study [28] took 7 years of implementation. The remaining 13 studies reported results based on 1 year of implementation. Due to the previously mentioned variations in measured variables, duration of implementing BSC, and differences in data collection instruments or data sources, refer to Table 2, the authors decided that conducting a meta-analysis would not lead to meaningful results. Instead, a comparison of the impact was performed using the bar charts. See Figs. 4, 5, 6, 7.

**The impact of BSC implementation**

The outcomes of BSC implementation in each of the 20 included studies are shown above in Figs. 4, 5, 6, 7, which illustrate the impact on patient satisfaction, HCW satisfaction, and financial performance (in percentage and currency). The exact impact type and duration were...
Fig. 5 HCW Satisfaction Impact. Increase or decrease in HCW satisfaction rate after BSC implementation (%)

Fig. 6 Financial Impact (%). Increase or decrease in financial performance after BSC implementation (%)
specified for each measurement, and the results are discussed further in the discussion section below.

Since we did not limit our search strategy to the three BSC impacts defined in our aim, eight studies of different impact types resulted in the abstract screening process. Examples of the other BSC impact types include; influencing the market share, the number of new patients, the number of consultations and visits, community satisfaction, percentage of residents receiving outreach activities, number of sampled children, availability of essential drugs, and decreasing the average length of stay [46, 47]. However, after reading the full texts, the authors decided to exclude these studies as they lacked sufficient relevance to our research aim. See (S3 Appendix).

Quality assessment
As illustrated in S4 Appendix, each study was evaluated in terms of RoB. For that purpose, the RoB 2 tool was employed to assess the sole RCT study [32], in which the assessment was deemed fair, except for the performance bias. On the other hand, the RoB in the quasi-experimental and observational studies was measured using the ROBINS-I tool; and it was found that there was no information about analysis methods of confounders’ adjustments except in four studies [33, 36, 39, 42]. The confounding agents were apparent in three studies [27, 35, 37]. However, the three studies failed to adjust for the confounders, which may have affected the precision of the measurement.

Furthermore, the selection bias across studies reflected serious RoB in five studies [27, 30, 31, 33, 40]. A possible reason the intervention and the follow-up did not coincide together and a potentially substantial amount of follow-up time was missing in the analyses. The moderate risk of bias showed that the intervention status was well defined, but some aspects of the assignments of intervention status were determined retrospectively. Further, outcome measurements bias was raised either due to the non-blinding of intervention among assessors [37] or because the outcome measure was subjective and likely to be influenced by other factors [29, 38]. See (S4 Appendix).

Discussion
Discussion of the main results
This systematic review aimed to identify all the studies which measured the impact of BSC implementation on three variables: HCW 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 dollars. However, a few studies reflected a moderately negative impact on financial performance, which form three distinct categories. The first category includes
study [29], which explained the occurrence of unintended events that may have negatively affected financial performance. The second category comprises studies [30, 40] that revealed a highly positive impact on financial performance in previous or subsequent years, which may reflect a sloth in following up. The third category includes studies [28, 39] that showed a positive impact on financial performance on one or more of the other impact types. On the other hand, the analysis of BSC impact for HCW satisfaction revealed a less remarkably positive impact. See (Figs. 4, 5, 6, 7).

**Agreements and disagreements with other studies or reviews**

The findings obtained from the present systematic review are in line with a systematic review [14] that reviewed BSC’s benefits in business, management, and accounting fields. Furthermore, the present study is the first to summarize all BSC implementations and their impacts on the health care sector based on quantitative comparisons. Moreover, the current study was compared with other reviews in the health care sector. For instance, a review [15] carried out a mere description regarding the application of BSC. In contrast, a review [3] only summarized the perspectives and dimensions utilized. Lastly, a review [16] only mentioned examples about BSC impact.

One probable explanation for the mild impact on HCW satisfaction can be referred to the lack of managerial engagement with the non-managerial HCW upon BSC implementation, the lack of understanding by HCW about the advantages of BSC implementation, or the fear of potential responsibility and accountability placed upon HCW due to BSC implementation. As a result, HCW may have declined to implement BSC, contributing to a lower satisfaction score. In conclusion, future researchers should consider increasing employee participation in BSC implementations.

For instance, in a study [29], the employees did not have incentives or motives to participate in BSC since they were permanent employees. Further, the study showed that HCW above 40 years old negatively influenced creativity and productivity upon BSC implementation. Other researchers in [48] also referred to this challenge, who noted that major deficiencies arose from qualified personnel and HCW aging. However, those researchers have also suggested that the high-ranking qualifications of HCW, driving learning and a growth perspective, will eventually generate motivation for new HCW to resolve this issue. Other proposed ideas to solve this problem were creating an open environment for learning and growth and encouraging active communication with HCW to ensure the successful implementation of BSC. Other researchers [49] encouraged senior management commitments to involve non-managerial HCW, promoting clear articulation of benefits and relevancy of BSC to clinicians. This challenge mirrors the findings of another review [50], which realized that the attitude perceived by health care professionals of accreditation was negative and skeptical because of quality concerns regarding services and their cost. Therefore, the authors in the latter study suggested that health care professionals, especially physicians, require more intensive education about the potential benefits of accreditation.

Finally, the quality assessment revealed that many studies had high RoB, which may have affected the impact results. A recommendation for the researchers and managers implementing BSC in the future is to dedicate more focus to raising the quality of implementation and lowering the RoB. Moreover, a better focus on the second and third generations of BSC aspects is essential.

**Strengths and weaknesses**

The current systematic review contains several strengths. To our knowledge, this is the first paper that has analyzed all the studies which measured the impact of BSC on patient satisfaction, HCW satisfaction, and financial performance in HCO. The results and analysis of this systematic review support a positive impact for applying BSC in HCO, especially on patient satisfaction and financial performance. Further, a greater emphasis on the role of HCW is required when implementing BSC since HCW 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 at most BSC studies. Finally, unlike other BSC reviews [8, 16], which included definitions of biobanks, pharmacies, laboratories, radiology, and medical colleges in HCO, this review limited the definition to the primary, secondary, or tertiary health care organizations. This strategy leads to the homogeneity of the resulting studies and leads to more valid comparisons among the results.

Nevertheless, this paper has some limitations. First, it focused on the impact of 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. Secondly, no meta-analysis could be applied to this systematic review resulting from 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. Thirdly, the current review included studies that
measured the impact after at least 1 year of implementation. Fourthly, it is essential to mention that the impact comparability is roughly more rational for patient satisfaction and HCW 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 the BSC implementation in health care databases; consequently, future systematic reviews are recommended to include studies in management and health policy databases.

Conclusions
In conclusion, this systemic review offers evidence to HCO and policymakers on the benefits of implementing 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 risk of bias in BSC implementation in the future. HCO managers are also advised to consider HCW satisfaction and engagement with BSC implementations. Finally, an additional assessment of the BSC impact in HCO during the COVID-19 pandemic is required, as we could not find any.

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s12913-021-07452-7.

Acknowledgments
We would like to thank Dr. Vahideh Moghaddam, Diégo Andrade, and Wang Zhe for their help in translating and data extraction of Persian, Spanish, and Chinese studies, respectively. Also, we would like to thank Dr. Duha Shellah, Dr. Mohanad Khattabeh, Dr. Ibrahim A. Shilbaya, Mr. Jason Titman for revising the manuscript.

Authors’ contributions
This systematic review was planned and designed by FA, S.L., S.H., I.B & D.E. Specifically, FA & S.H performed the processes of searching for eligible studies and data extraction. F. A & S. H performed the interpretation of the data. F. A completed the current systematic review, and all of the authors participated in revising the manuscript. S.L., I.B & D.E supervised this systematic review. Ultimately, all of the authors approved the final manuscript and agreed to be personally accountable for the author’s contributions and ensure the accuracy and integrity of any part of the work are appropriately investigated and resolved.

Funding
The research was financed by the Thematic Excellence Program 2020—Institutional Excellence Sub-programme / National Excellence Sub-program of the Ministry for Innovation and Technology in Hungary, within the framework of the 2. thematic and Biomedical Engineering programme of the University of Pécs. Code: 2020–4.1.1-TEP2020.

Availability of data and materials
All data generated or analyzed during this study are included in this published article [and its supplementary information files].

Declarations
Ethics approval and consent to participate
Not Applicable since it is a systematic review.

Consent for publication
Not Applicable.

Competing interests
The authors have declared that no competing interests exist.

Author details
1Doctoral School of Health Sciences, Faculty of Health Sciences, University of Pécs, Maria u. 5-7, Pécs H-7621, Hungary. 2Institute for Health Insurance, Faculty of Health Sciences, University of Pécs, Pécs, Hungary. 3Cochrane Hungary, Clinical Center of the University of Pécs, Medical School, University of Pécs, Pécs, Hungary.

Received: 6 September 2021 Accepted: 16 November 2021 Published online: 13 January 2022

References
1. Kaplan R, Norton D. The balanced scorecard—measures that drive performance. Harv Bus Rev. 1992;70(1):71–9.
2. Hannula M, Kulmala H, Suomala P. Total quality management and balanced scorecard-a comparative analysis. In: Werther WJ, et al., editors. Productivity & Quality Management Frontiers - VII. Bradford: MCB University Press; 1999.
3. Amer F, Hammoud S, Khattabeh L, Lohner S, Boncz I. A systematic review: the dimensions and indicators utilized in the performance evaluation of health care organizations- an implication during COVID-19 pandemic. Preprint. 2021. https://doi.org/10.1101/2021.06.26.21259568.
4. Kaplan RS. Conceptual foundations of the balanced scorecard. Handbook Account Res. 2009;3:1253–69.
5. Speckbacher G, Bischof J, Pfeffer T. A descriptive analysis on the implementation of balanced scorecards in German-speaking countries. Manag Account Res. 2003;14:361–88.
6. Meliones J. Saving money, saving lives. Harv Bus Rev. 2000;78(5):62–64 66–7.
7. Duke-Children’s-Hospital. Duke Children’s Hospital Case Abstract. Balanced Scorecard Collaborative Hall of Fame. 1:13. https://thepaladiumgroup.com/download/file/1531102794_11Attachment_dubai_boot_camp_-_March_2019_opt.pdf?3DDubai_Boot_Camp_-_March_2019_opt.pdf. Accessed 15 May 2021.
8. Mcdonald B. A review of the use of the balanced scorecard in healthcare. BMC Consult. 2012;2012:1–32.
9. Amer F, Hammoud S, Farran B, Boncz I, Endrei D. Assessment of Countries’ preparedness and lockdown effectiveness in fighting COVID-19. Dis Med Public Health Prep. 2021;35:e15–22. https://doi.org/10.1017/dmp.2020.217.
10. Lou NM, Montreuil T, Feldman LS, Fried GM, Lavoie-Tremblay M, Bhanji F, et al. Evaluations of Healthcare Providers’ Perceived Support From Personal, Hospital, and System Resources: Implications for Well-Being and
Management in Healthcare in Montreal, Quebec, During COVID-19. Eval Health Prof. 2021;44:319–22. https://doi.org/10.1177/10713632211012742.

11. Ali A, Owied MM, Virginia P, Ruth E, Abdullrahman AT. Using Logit Model and Key Performance Indicators to Construct A Field Hospital for Coronavirus Pandemic Cases. BMC Emerg Med. 2020;12–4. https://doi.org/10.1210/bs.3r.10-10231/1/v1.

12. Alrasawdeh HM, Al-Tammemi AB, Alzawarneh MK, Al-Tammii A, Alkholy M, Al Sirieh F, et al. Occupational burnout and job satisfaction among physicians in times of COVID-19 crisis: a convergent parallel mixed-method study. BMC Public Health. 2021;21:811.

13. Hansen EG, Schlegger S. The sustainability balanced scorecard: a systematic review of architectures. J Bus Ethics. 2016;133:193–221.

14. Ratnamningrum AYA, Setiawan D. Balanced Scorecard: Is It Beneficial Enough? A Literature Review. Asian J Account Perspect. 2020;13:65–84.

15. Bohn V, Lacaille D, Spencer N, Barber CE. Scoping review of balanced scorecards for use in healthcare settings: development and implementation. BMJ Open Qual. 2021;10:1–10. https://doi.org/10.1136/bmjmq-2020-001293.

16. Rabbani F, Jafari SMW, Abbass F, Pappas G, Brommels M, Tomson G. Reviewing the application of the balanced scorecard with implications for low-income health settings. J Healthc Qual. 2007;29:21–34.

17. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med. 2009;6:e1000100.

18. Methley AM, Campbell S, Chew-Graham C, McNally R, Cheraghi-Sohi S, PICO, PICOS and SPIDER: a comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. BMC Health Serv Res. 2014;14:579.

19. Lefebvre C, Glanville J, Briscoe S, Littlewood A, Marshall C, Metzendorf M-I, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ. 2016;355:i4919. https://doi.org/10.1136/bmj.i4919.

20. Higgins JP, Altman DG. Assessing Risk of Bias in Included Studies. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editors. Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019). Cochrane, 2019. Available from www.training.cochrane.org/handbook.

21. Sterne JA, Hernán MA, Reeves BC, Savović J, Ioannidis JP, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ. 2020;365:S9. https://doi.org/10.1136/bmj.s9.

22. Higgins PA, Lane M, Assistant. Assessing a balanced scorecard for a hospital system. J Health Care Finance. 2006;31:280–8. https://doi.org/10.1016/j.hcfn.2013.12.003.

23. Pink GH, McKillop I, Schraa EG, Preyra C, Montgomery C, Baker GR. Creating a balanced scorecard for a hospital system. J Health Care Finance. 2001;27:1–20.

24. Chang W, Tung Y, Huang C. Performance improvement after implementing the balanced scorecard: a large hospital’s experience in Taiwan. Total Qual Manag Bus Excell. 2000;11:943–54.

25. Baker GR, Pink GH. A balanced scorecard for Canadian hospitals. Healthc Manag Forum. 1995;8:7–13.

26. Fields SA, Cohen D. Performance enhancement using a balanced scorecard in a patient-centered medical home. Fam Med. 2011;43:735–9.

27. Gumbus A, Bellhouse DE, Lyons B. A three year journey to organizational and financial health using the balanced scorecard: a case study at a Yale New Haven health system hospital. J Bus Econ Stud. 2003;9:54–64.

28. Harber B. The balanced scorecard-solution at Peel memorial hospital. Healthc Qual. 1998;1:59–63.

29. Abdullah M, Marikmon D, Faizi A, Santi A, Endoskopki P, Cerna S, et al. The use of balanced scorecard to support achievement of business plan in digestive endoscopy center, 2013. Indones J Gastroenterol Hepatol Dig Endosc. 2014;15:93–7.

30. Jones MLH, Filip SJ. Implementation and outcomes of a balanced scorecard model in Women’s Services in an Academic Health Care Institution. Qual Manag Health Care. 2000;8:40–51. https://doi.org/10.1097/00019514-2000080400-00005.

31. Chan GJ, Parca KB, Shihoming ME, Tredwell P, Pourie EJO. Improving health services to displaced persons in Aseh, Indonesia - a balanced scorecard. Bull World Health Organ. 2010;88:709–12.

32. Behrouzi F, Saharoun AM, Ma’aram A. Applications of the balanced scorecard for strategic management and performance measurement in the health sector. Aust Health Rev. 2014;38:208–17.

33. Dyball MC, Cummings L, Yu H. Adoption of the concept of balanced scorecard within NSW health: an exploration of staff attitudes. Financ Account Manag. 2011;27:335–61. https://doi.org/10.1111/j.1468-0408.2011.00528.x.

34. Alkhenizan A, Shaw C. The attitude of health care professionals towards accreditation: a systematic review of the literature. J Fam Community Med. 2012;19:74. https://doi.org/10.4103/2230-8229.98261.