TÉCNICO
LISBOA

On evaluating vertically merged hospitals in terms of quality and access

The Portuguese experience

Guilherme Café Estêvão Pontes Amado

Thesis to obtain the Master of Science Degree in

Biomedical Engineering

Supervisors: Prof. Diogo Filipe Cunha Ferreira
Prof. Alexandre Manuel Martins Morais Nunes

Examination Committee

Chairperson: Prof. Mónica Duarte Correia de Oliveira
Supervisor: Prof. Diogo Filipe Cunha Ferreira
Member of the Committee: Prof. Adalberto Campos Fernandes

December 2020

Declaration

I declare that this document is an original work of my own authorship and that it fulfills all the requirements of the Code of Conduct and Good Practices of the Universidade de Lisboa.

Preface

The work presented in this thesis was performed at Instituto superior Técnico (Lisbon, Portugal), during the period February-December 2020, under the supervision of Prof. Diogo Filipe Cunha Ferreira and Prof. Alexandre Manuel Martins Morais Nunes.

Acknowledgments

To Professor Diogo Cunha Ferreira, not merely for his excellent guidance, availability, and knowledge sharing, but mainly for his friendship and support during this long journey. His vision and drive for academic research have deeply inspired me. It was a great privilege to conduct a research and learn under his guidance.

To Professor Alexandre Morais Nunes, co-supervisor of this thesis, who since the first moment demonstrated a great readiness, availability and support in the realization of the present study. His participation, suggestions, and socio-political inputs were fundamental to the elaboration of this work.

To all my family, but specially my mother, for all the comprehension, patience and support during this arduous yet rewarding journey at Técnico Lisboa.

To my friends and colleagues, for the unconditional support during the realization of this work and for providing me with the best environment I could ask for during these five years. I will take in my memory all the experiences and good moments spent at this university.

To all, accept my endless gratitude.

Abstract

Uncovering the optimal delivery care model to maximize value in health is one of the dominant concerns of health systems. The present study aims at contributing to the serious and contentious discussion regarding the effects of the implementation of the vertical integration model, notably by enriching the literature devoted to establishing the link between healthcare outcomes and vertically integrated healthcare providers, through an exhaustive literature review and a robust performance analysis of Portuguese healthcare providers. Using an output-oriented Data Envelopment Analysis and a Malmquist Index approach, one studied the impact of implementing vertical models (Local Health Units) on quality- and access-related performance, considering the environmental effect. From the study of 39 healthcare providers, between 2015 and 2019, one may conclude that: hospitals included in vertical models exhibit statistically significant higher partial performance than singular hospitals and hospital centers; a significant number of hospitals within vertical models are above the 75th percentile regarding partial performance; hospitals included in vertical models exhibit statistically significant lower overall performance and frontier-shift related performance than singular hospitals and hospital centers, for services availability. The overall consideration is that hospitals within vertical models exhibit slight improvements in quality and access measures when considered the environment in which these are incorporated.

Keywords

Vertical Integration; Data Envelopment Analysis; Malmquist Index approach; Environmental Effect; Quality; Access

Resumo

Encontrar o modelo ideal para maximizar o valor em saúde é uma das dominantes preocupações dos sistemas de saúde. O presente estudo visa contribuir para a significante discussão dos efeitos da implementação do modelo de integração vertical, em especial pelo enriquecimento da literatura dedicada a estabelecer a ligação entre os indicadores de desempenho e os prestadores de saúde verticalmente integrados, através de uma revisão exaustiva da literatura e uma análise de desempenho robusta de prestadores de saúde Portugueses. Usando um modelo de Data Envelopment Analysis, orientado para os outputs, e Índices de Malmquist, estudou-se o impacto da implementação de modelos verticais (Unidades Locais de Saúde) na qualidade e no acesso, tendo em consideração o efeito ambiental. Através do estudo de 39 prestadores de serviços de saúde, entre 2015 e 2019, conclui-se que: hospitais incluídos em modelos verticais apresentam um desempenho parcial estatisticamente superior a hospitais singulares e centros hospitalares; um número significativo de hospitais incluídos em modelos verticais encontra-se acima do percentil 75 em termos de desempenho parcial; hospitais incluídos em modelos verticais exibem um desempenho geral e um desempenho relacionado com o desvio da fronteira estatisticamente inferiores, quando comparados a hospitais singulares e centros hospitalares, especificamente para a dimensão ”disponibilidade de serviços”. A consideração final é que hospitais incluídos em modelos verticais deverão apresentar melhorias na qualidade e no acesso, quando considerado o ambiente externo no qual estão incorporados.

Palavras Chave

Integração Vertical; Data Envelopment Analysis; Índices de Malmquist; Efeito Ambiental; Qualidade; Acesso

Contents

Contents v

List of Figures viii

List of Tables viii

1 Introduction 1

1.1 The health sector in Portugal: The context of the problem . . . 2

1.2 Defining Vertical Integration . . . 4

1.2.1 Vertical integration in healthcare . . . 4

1.3 Local Health Units . . . 5

1.4 Objectives of the thesis . . . 6

1.5 Structure of the thesis . . . 7

2 Literature review 9

2.1 Conceptual framework . . . 10

2.2 Methodology . . . 11

2.2.1 Searching strategy . . . 11

2.2.2 Screening . . . 12

2.2.3 Post-screening/pre-basic statistics . . . 13

2.2.4 Basic statistics over the collected studies . . . 14

2.2.5 Meta-analysis . . . 15

2.3 Main results . . . 15

2.3.1 Basic statistics over the collected studies . . . 29

2.3.2 Meta-analysis . . . 30

2.4 Discussion . . . 30

2.4.1 Impact of vertical integration on costs . . . 31

2.4.2 Impact of vertical integration on prices of care . . . 33

2.4.3 Impact of vertical integration on efficiency . . . 34

2.4.4 Impact of vertical integration on quality of care . . . 35

2.4.5 Impact of vertical integration on access . . . 37

2.5 Concluding remarks and directions for future work 38

3 Case study **41**

3.1 Data collection and methodological issues 42

3.1.1 Sample 42

3.1.2 Desirable and undesirable process variables 42

3.1.3 Environmental variables 45

3.1.4 Models and scenarios 47

3.1.4.A Quality- and access related performance 47

3.1.4.B Outcome-specific performance 47

3.1.4.C Scenarios 48

3.2 Two clusters of healthcare providers: Basic statistics 48

3.3 Methodology 51

3.3.1 Pre-processing 51

3.3.1.A Outliers deletion 51

3.3.1.B Principal Component Analysis 52

3.3.2 Data Envelopment Analysis 53

3.3.3 Consideration of the environmental variables 55

3.3.4 Malmquist Indices 55

3.4 Overview of concepts in analysis 58

4 Results and discussions **59**

4.1 Data Envelopment Analysis: Cluster scores 60

4.1.1 Quality- and access-related cluster scores 61

4.1.2 Quality-related cluster scores 62

4.1.3 Access-related cluster scores 63

4.2 Data Envelopment Analysis: Discriminated scores 64

4.2.1 Quality- and access-related discriminated scores 64

4.2.2 Quality-related discriminated scores 65

4.2.3 Access-related discriminated scores 66

4.3 Malmquist Indices 67

4.3.1 Quality- and access-related indices 67

4.3.2 Quality-related indices 67

4.3.3 Access-related indices 68

4.4 Findings discussion 69

4.4.1 Partial performance 69

4.4.1.A Quality- and access-related partial performance 69

4.4.1.B Quality-related partial performance 71
4.4.1.C Access-related partial performance 73
4.4.2 Discriminated partial performance 75
4.4.3 Overall performance, partial performance spread, and frontier-shift related performance 78

5 Conclusions **81**
5.1 Final considerations 82
5.2 Research limitations and future work 84

6 Bibliography **87**

A Appendix A **A-1**
A.1 Basic statistics A-2
A.2 Meta-analysis A-5
A.3 Impact of vertical integration on efficiency A-7
A.4 Impact of vertical integration on quality of care A-8

List of Figures

2.1 Conceptual framework for the performance assessment of vertical integration in healthcare. 10
2.2 Screening diagram. 13

3.1 Process variables. 45
3.2 Environmental variables. 46

A.1 Which methods are the most used when analyzing vertical integration? A-2
A.2 What is the trend regarding the number of studies published per year? A-3
A.3 Which type of vertical integration is the most studied? A-3
A.4 Which countries do most studies belong to? A-4

List of Tables

2.1 Search strategy. 11
2.2 Vertically integrated providers assessment: favourable evidence. 16
2.3 Vertically integrated providers assessment: unfavourable evidence. 22
2.4 Vertically integrated providers assessment: inconsistent or statistically insubstantial evidence. 26
2.5 Empirical evidence on costs. 31
2.6 Empirical evidence on prices of care. 33
2.7 Empirical evidence on efficiency. 34

2.8 Empirical evidence on quality of care. 36
2.9 Empirical evidence on access. 38

3.1 Process variables used in the literature. 43
3.2 Environmental variables used in the literature. 46
3.3 Overview of models I to III. 47
3.4 Overview of models IV to IX. 48
3.5 Process variables: Basic statistics. 49
3.6 Environmental variables: Basic statistics. 50
3.7 Overview of concepts in analysis. 58

4.1 Quality- and access-related cluster scores: Statistics. 61
4.2 Quality-related cluster scores: Statistics. 62
4.3 Access-related cluster scores: Statistics. 63
4.4 Quality- and access-related discriminated scores. 64
4.5 Quality-related discriminated scores. 65
4.6 Access-related discriminated scores. 66
4.7 Quality- and access-related productivity indexes. 67
4.8 Quality-related productivity indexes. 68
4.9 Access-related productivity indexes. 68
4.10 Quality- and access-related partial performance (Top 10 healthcare providers). 76
4.11 Quality-related partial performance (Top 10 healthcare providers). 76
4.12 Access-related partial performance (Top 10 healthcare providers). 76

A.1 Which methods are the most used when analyzing each one of the outcomes regarding vertical integration? A-2
A.2 What are the journals in which there have been more studies published? A-4
A.3 Which of the selected articles are most cited (Top 10)? Source: Google Scholar A-5
A.4 Methods vs. Type of vertical integration. A-5
A.5 Type of vertical integration vs. Costs A-6
A.6 Methods vs. Costs A-6
A.7 Methods vs. Efficiency A-6
A.8 Explicit analysis of the efficiency-related studies included in the sample. A-7
A.9 Length of stay: An indicator of efficiency. A-8
A.10 Overview of the main quality measures analyzed in the sample of studies. A-8

Acronyms

ACS	Ambulatory care-sensitive
AMC	Academic medical centers
AMI	Acute myocardial infarction
ANOVA	Analysis of variance
ASA	Aspirin at
BoD	Benefit of Doubt
CHF	Chronic heart failure
CMI	Case-mix index
CRS	Constant returns to scale
CT	Computed tomography
DEA	Data envelopment analysis
DID	Difference-in-difference
DMU	Decision-making unit
DRG	Diagnosis-related group
ED	Emergency department
FTE	Full time equivalent
GEE	Generalized estimating equations
GI	Gastrointestinal
GLMM	Generalized linear mixed model
HC	Hospital Center
HF	Heart failure
HiAP	Health in All Policies
HMO	Health Maintenance Organization

HR	Human resources
HUC	Hospital and University Center
KPI	Key performance indicator
LHU	Local Health Unit
LPN	Licensed practical nurses
LTC	Long-term care
LTCF	Long-term care facility
MCO	Managed care organization
MI	Malmquist index
NHS	National Health Service
OLS	Ordinary least squares
OR	Operating room
PCA	Principal component analysis
PN	Pneumonia
PO	Physician organization
PPE	Public enterprise entity
PPP	Public-private partnership
RN	Registered nurses
SEM	Structural equation modeling
SFA	Stochastic frontier analysis
SH	Singular Hospital
SNF	Skilled nursing facility
TE	Technical efficiency
TUCP	Total urgent care procedures
US	United States
VRS	Variable returns to scale
VI	Vertical integration

1

Introduction

1.1 The health sector in Portugal: The context of the problem

Contrary to what one might think, it was only in the twentieth century that several reforms were implemented towards a better healthcare for the Portuguese population. The creation of the Portuguese National Health Service (NHS) in 1979 was the major turning point for the health sector in Portugal, by establishing a universal health system, free at the point of use. The NHS Law followed other critical moments, such as: the declaration of the first act of public health legislation, known as the Ricardo Jorge reform (1901), the creation of the Ministry of Health and Assistance (1958), the acknowledgment of the state as responsible for health policy and implementation (1971), the creation of first-generation primary care centres, through the medium of the Gonçalves Ferreira reform (1971), and the inclusion of the citizens' right to healthcare in the Portuguese Constitution (1976) (Barros et al., 2011).

Since 1979, many additional steps in the Portuguese health policy may be identified. In the early 1980s, an alternative to the public service was developed. Then, in the late 1980s and early 1990s, market mechanisms were promoted, and several policies that drifted away from the market-driven healthcare provision were introduced. At the turn of the century, the NHS became a mixed system - enabling the interaction between the public and the private sectors - with an eye on the promotion of efficiency (Simões et al., 2017).

From an organizational point-of-view, the Portuguese health system is composed of the NHS, private voluntary health insurance schemes, and health subsystems associated with the labour market (Ferreira et al., 2018). More precisely, the healthcare delivery system incorporates a diverse range of healthcare providers (primary care facilities, hospitals, long-term care (LTC) networks, and pharmacies), some of which are public (not-for-profit) and others private. Each provider is coupled to the population and the Ministry of Health and its institutions in its own way.

Regarding physical resources, Portugal had, in 2019, 238 hospitals, presenting approximately a half-and-half distribution regarding public and private hospitals, allowing a total capacity of 36,913 beds, according to *PORDATA*. Concerning primary care facilities and pharmacies, the numbers are around 1,772 (as of 2015, according to Simões et al. (2017)) and 2,924 (as of 2019, according to *PORDATA*), respectively. Geographically, the distribution of healthcare providers, and, therefore, of health professionals, is not even, giving rise to some inequalities in access to care. In fact, the Northern, Lisbon, and Tagus Valley accumulate more than 70% of the health workforce (Ferreira et al., 2018). This problem is further aggravated in some more isolated interior regions of Portugal. Several facilitated recruitment processes were created in those regions to counterbalance this situation, like increases in salaries or even enhancement of conditions for participation in research (Ferreira et al., 2018). Nevertheless, Portugal still struggles today with inequalities of access to care for geographic motives.

As aforementioned, the Portuguese NHS is universal, equitable, and tendentiously free (co-payments are charged to some patients). The financial sustainability behind this system is derived from the implementation of the Beveridge model. In this model, the central government employs the collected funds from citizens' tax payments into the public health sector. The allocation of financial resources for each of the providers relies on various traits, e.g., size, scope of services provided, patient's complexity, cost-efficiency, among others (Ferreira and Marques, 2018).

Total health expenditure per inhabitant has, in general, been increasing in the last few decades, mainly due to demography changes, healthcare advances, and technology evolution (Ferreira and Marques, 2018). Notwithstanding, total healthcare expenditure had a significant break between 2010 and 2014, motivated by the economic and financial crisis lived in Portugal. In the light of the Economic and Financial Adjustment Programme, several austerity-based policies were formulated to reduce costs and waste of public funds, with an eye on the improvement of efficiency and effectiveness of healthcare providers (Nunes et al., 2019). Nevertheless, it is known that financial crises commonly lift several severe concerns and that one was not an exception. Several authors refer to loss of employment, greater difficulty entering the labour market, or even increased mental disorders as relevant adverse effects of that crisis (Nunes et al., 2019). Naturally, it was a situation that changed the health sector in Portugal permanently by creating deficiencies in infrastructures and further aggravating the access to healthcare services.

Whether for austerity reasons or because of the constant search for improvements on cost efficiency, quality, or access of healthcare providers, to name a few, hospital management is a tremendous challenge for the Ministry of Health. Several significant reforms were implemented in hospital management models throughout the years, including corporatization, vertical and horizontal merging of public healthcare providers, and public-private partnership (PPP) contracts (Ferreira and Marques, 2018).

In addition to management, monitoring healthcare providers - one other task of the Ministry of Health - is equally important. Portugal has a large-scale information infrastructure that performs the leading role in monitoring health system performance. It encompasses almost all levels of care and stimulates quality improvements by enabling analyzing the gathered data. Data sources encompass hospitals or even primary care facilities, providing a data infrastructure built with EHRs' information, which is then stored and presented in the Portuguese Health Data Platform (Plataforma de Dados da Saúde, PDS) (Simões et al., 2017). This massive information infrastructure powerfully enables improvements in the health sector, even if several challenges regarding, for example, patient privacy protection and legality towards patient data connections, are still around.

By positioning Portugal in the European context, regarding health expenditure and other outcomes (e.g., life expectancy at birth, total life expectancy), it is easy to perceive that the Portuguese health system is among the most efficient in Europe.

1.2 Defining Vertical Integration

The economic definition of vertical integration relies on the common ownership of different segments in the vertical chain of production (Robinson, 2001). Alternatively stated, a vertically integrated company comprises shared ownership of two or more organizations, in which the output of one of them operates as the input of the subsequent.

Every reorganization strategy of this kind merges an upstream firm with a downstream firm. In supply chain management terminology, an upstream firm manufactures an intermediary product and sells it to the downstream firm. When a company integrates with its product supplier (upstream firm), we are dealing with backward integration. Instead, when a company integrates with its distributor (downstream firm), we are in the presence of forward integration (Romme, 1990).

The increase in benefits to the organization's stakeholders is the main justification for vertical integration. For a standard firm, those benefits correspond to profits, which may increase as a response to efficiency improvements within the company (Byrne and Ashton, 1999).

1.2.1 Vertical integration in healthcare

The integration of services from different levels of care (primary care, acute care, and post-acute care) is designated as vertical integration, as opposed to the horizontal expansion, created when hospitals (acute care) associate between themselves (Szostak, 2015). With the assimilation of healthcare providers of different levels into a single unit, fragmentation of care is prevented.

An additional variety of changes in the healthcare sector may also be categorized as vertical integration: integration of pharmacists' organizations and hospitals (Calvert et al., 2012), cooperation between hospitals and a Health Maintenance Organization (HMO) (insurance firm) (Town and Vistnes, 2001), integration of academic medical centers (AMCs) with clinical practice facilities (Van den Abbeele et al., 2016) or even affiliation of hospitals with laboratories (Forsman, 1996). Nevertheless, in the majority of the literature, the definition is associated with affiliations between primary care and hospitals (Lopes et al., 2017; Short and Ho, 2019) or with the combination of hospitals and post-acute care (Konetzka et al., 2018; Gupta et al., 2019).

In hospital-physician affiliations, physicians are abstractly seen as the upstream firm and hospitals as the downstream firms. Physicians have consultations with patients (primary care), making potential referrals for additional care. Hospitals appear as the downstream firms at the next level in the chain, receiving the patient - the intermediate product - and providing the next level of healthcare services (Post et al., 2017). Within the hospital-physician affiliation concept, many forms of integration may be implemented: the employment of physicians (the tightest form); comanagement; and independent affiliations (the loosest form), which happens when hospitals purchase physician administrative time through medical directorships (Sowers et al., 2013).

Regarding integration of acute care with post-acute care, hospitals may be seen as the upstream firms and post-acute care facilities as the downstream firms. In the same fashion as before, the patient is considered the intermediate product, carried from one level of the healthcare production process to another (Post et al., 2017).

Generally, patient-centered vertically integrated health systems may present two distinct components: structural integration and functional integration. The former implies the ownership of a range of healthcare services necessary for the patient flow through the vertical chain. The latter is introduced when effective coordination between those services is established. Understandably, functional integration may be harder to attain than structural integration, meaning the challenge mainly concentrates on achieving and measuring functional integration (Byrne and Ashton, 1999). Indeed, several integration strategies include only the structural integration component to facilitate its implementation (Huckman, 2005; Capps et al., 2018).

In healthcare, the stakeholders' concept includes not only individuals with financial interests but also the patients themselves. This circumstance expands the standard strategy goal to increases in efficiency, quality of care, access, among others, in addition to the standard increments in profits (when talking specifically about a private healthcare provider) (Byrne and Ashton, 1999).

1.3 Local Health Units

As perceived in Subchapter 1.1, new organizational arrangements involving public hospital institutions have been implemented to find a solution to the Portuguese health system's weaknesses. From those arrangements, the vertical merging of healthcare providers should be highlighted.

Local Health Unit (LHU) was the chosen term to designate the vertical merging between hospitals and health centers in Portugal. The first one, Matosinhos LHU, was created in 1999, through the Decree-Law No. 207/99, of 9th June. It emerged from the integration of the Pedro Hispano Hospital with the Health Centers of Matosinhos, Senhora da Hora, São Mamede, and Leça da Palmeira (Lourenço et al., 2010). This new approach premised two main objectives, which were set out in the preamble of the same Decree-Law:

> ”The improvement in the provision of healthcare by the NHS is based, in part, on the creation of conditions that enable the better management of its institutions and the better articulation of these institutions among themselves and with other institutions in the same geographical area.”

It was only in 2007 that the second LHU was created. Alto Alentejo PPE LHU integrated Portalegre and Elvas Hospitals and the Health Centers of Portalegre. In the following year, three additional PPE LHUs were initiated: Northern Minho, Southern Alentejo, and Guarda PPE LHUs (Lourenço et al., 2010).

In 2009, the Regional Health Administration of Lisbon and Tagus Valley (Administração Regional de Saúde de Lisboa e Vale do Tejo, ARSLVT) took the initiative to convene a meeting for the discussion of LHUs' performance in the preceding years. Several conclusions were reached, which stands out the lack of both existing evidence regarding the impacts of the creation of LHUs and monitoring indicators to objectively assess LHUs (Lourenço et al., 2010). Still in 2009, an LHU was established in Castelo Branco. Since then, Northeast/Bragança, in 2011, and Coastal Alentejo/Santiago do Cacém, in 2012, also greeted vertical mergers (Simões et al., 2017). It was then that Portugal reached the still present-day number of eight LHUs.

In 2015, the Portuguese Health Regulatory Agency (Entidade Reguladora da Saúde, ERS) carried out a study on the Portuguese LHUs, reaching no evidence of significant performance improvements when comparing LHUs to non-integrated health providers (Entidade Reguladora da Saúde, 2015).

Ultimately, relevant principles, rules, and goals have been revised, and others added to the LHU legal regime in the Decree-Law No. 18/2017, of 20th February, of which the following stand out: a) LHU financing is carried out by adjusted capitation for the risk calculated based on the environment in which each one is inserted (Chapter I, Section III, Article 25.º, 5); b) organization and functioning guidelines for LHUs should be defined both in the clinical and non-clinical areas (Chapter II, Section I, Article 7º, 1.c); c) processes to monitor and evaluate LHUs should be further developed, focusing on efficiency, quality of care, and economic and financial sustainability (Chapter II, Section I, Article 7º, 1.m); d) action plans should be coordinated across both levels of care, to create the LHU global action plan (Chapter II, Section I, Article 9º, a).

1.4 Objectives of the thesis

Health systems around the globe have been struggling with questions regarding the optimal delivery care model to better treat patients. This current study aims to contribute for the serious and contentious discussion regarding the effects of implementing the vertical integration model, notably by enriching the literature devoted to establish the link between healthcare outcomes and vertically integrated healthcare providers, through an exhaustive systematic review and a study of the Portuguese experience.

More precisely, the systematic review aims at: a) collecting a large sample of papers that analyze vertically integrated healthcare providers, regarding prices of care, costs, efficiency, quality and/or access; b) obtaining statistical information regarding the studies' sample (as the most studied country, most used methodology or even the most analyzed outcomes, to name a few); c) determining how vertical integration impacts each one of the outcomes. Additionally, the Portuguese case study seeks to:

d) discover the impact of the implementation of LHUs on quality and access of the hospital-settings component, using robust methodologies, implemented in MATLAB®, and considering the environmental effect; e) understand the political repercussions of the results obtained.

1.5 Structure of the thesis

This work is divided into five chapters. In the first chapter, Portugal's health sector and particular introductory concepts concerning health economics were presented. The second chapter briefly unveils the systematic review, revealing the conceptual framework, the methodology (the searching strategy, the screening process, and the rationale regarding the statistical procedures), the main results (from the exhaustive literature review and the meta-analysis), discussions, and conclusions and remarks for future work. The Portuguese case study is defined in the third chapter. This chapter describes the gathering of data, the study's sample, the desirable and undesirable variables, the environmental variables, and the distinct models and scenarios used. It also presents a fundamental statistical analysis to get the first insight into the performance of the hospitals. Ultimately, the chapter unfolds all the pre-processing procedures (outliers deletion, dimensionality reduction using Principal Component Analysis and product kernel approach to introduce the effect of the environmental variables) and the methodologies used to implement the performance analysis (Data Envelopment Analysis and Malmquist Indices). The fourth chapter provides the results and respective discussions. Finally, the fifth chapter summarizes findings and hypothesizes about future work and possible political repercussions of the findings.

2

Literature review

The following chapter is adapted from an article submitted to a peer-reviewed journal. The research provides an exhaustive analysis of a large sample of studies regarding five policy-related outcomes and a meta-analysis.

2.1 Conceptual framework

The concept of vertical integration described in Subchapter 1.2 utterly involves in practice many different strategies. However, as previously stated, vertical integration is most often defined within the delivery care system. In this literature review, the vertical integration assessment only includes systems integrating primary care with acute care or acute care with post-acute care, *i.e.*, hospitals exhibiting backward or forward integration.

To thoroughly analyze vertical models' performance, one might incorporate several policy-relevant outcomes in the conceptual framework: costs, prices of care, efficiency, quality of care, and access.

Figure 2.1 provides a clarifying schematic of integration strategies of interest as well as the relevant outcomes. The figure outlines a representative scheme of backward ("bckw") and forward ("fwd") integration strategies within the dashed lines, enumerating potential healthcare providers included in those strategies.

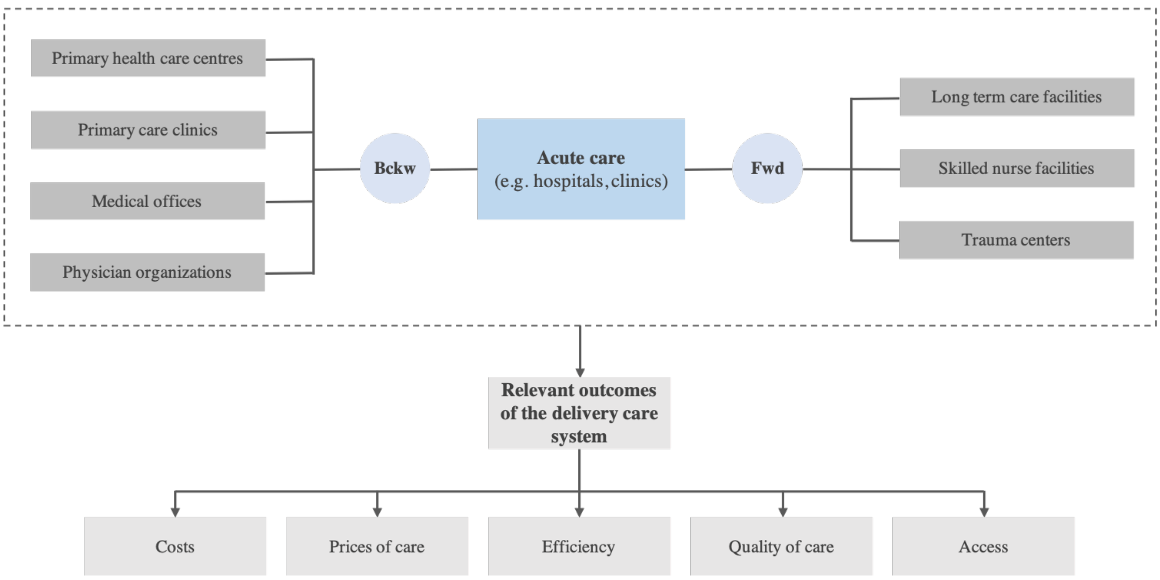


Figure 2.1. Conceptual framework for the performance assessment of vertical integration in healthcare.

2.2 Methodology

2.2.1 Searching strategy

One conducted a systematic literature search in three central databases – Science Direct, PubMed, and PMC – for quantitative research articles regarding vertical integration within the delivery care system.

As previously stated, in the literature, vertical integration is mainly associated with integrating inpatient care and care after discharge and with affiliations between physicians and hospitals. Both cases are commonly referred to as vertical integration. However, the second case also involves terms like ”hospital-physician integration” or even ”hospital-physician affiliation”. One considered this situation when developing the search strategy, as proceeded by Machta et al. (2019). The set of terms ”delivery care integration” was additionally considered to eventually incorporate further studies that do not use the terms above. Contrary to the previously announced terms, this one had not yet been used for database search.

Table 2.1 presents each database's search conditions (words within the main text, keywords, and type of article). Respecting the conceptual framework (Figure 2.1), ”prices”, ”costs”, ”efficiency”, ”quality”, and ”access” were included in the query search.

Because of their relevance, the literature review included seventeen additional studies identified through other sources. Machta et al. (2019) provided thirteen of the additional studies. The remaining four studies (Acerete et al., 2011; Alonso et al., 2014; Caballer-Tarazona and Vivas-Consuelo, 2016; Comendeiro-Maaløe et al., 2019) assessed the Alzira model, a Spanish model that respects the conceptual framework, yet it is only referred, in the literature, as a PPP, despite being as well a vertically integrated system (Rechel et al., 2009a,b).

Table 2.1. Search strategy.

Science Direct	
Articles with these terms	((”vertical integration” AND hospital) OR (”hospital physician” AND (affiliation OR integration)) OR (”delivery care” AND integration)) AND (prices OR costs OR efficiency OR access OR quality)
Limits	Article types: Research articles
PubMed and PMC	
Builder	((”vertical integration” AND hospital) OR (”hospital physician” AND (affiliation OR integration)) OR (”delivery care” AND integration)) AND (prices OR costs OR efficiency OR access OR quality)

2.2.2 Screening

Screening of papers to analyze in the next steps followed the original PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2009 Flow Diagram. This four-phase flow diagram (identification; screening; eligibility; included), widely used in the literature in systematic reviews on healthcare quality and efficiency (Fong et al., 2016; Lear et al., 2017; Machta et al., 2019; Niñerola et al., 2020), facilitates not only the conduction of the systematic review but also the report of the assessment.

Figure 2.2 presents the screening diagram. One might set two phases in the screening and selection proceedings of this literature review: a review of the title and abstract of the 3,079 papers, followed by a full-text assessment of the eligible papers. These phases are included in the structure of the original PRISMA 2009 Flow Diagram.

First, one scrutinized the titles and abstracts of the articles obtained. The course of action was consistent with the conceptual framework. Nevertheless, it was over-inclusive in situations where the title and the abstract were not enough to objectively identify the definition included in that paper or the outcomes used. Additionally, only quantitative research articles or articles in which the type of study was inconclusive (by only reading the abstract) passed the screening. From the original 3,079 articles, one discarded 2,820 in this phase of screening.

Then, 259 full-text articles were assessed for eligibility. Papers were excluded based on the following exclusion criteria: (a) critical appraisal studies (n = 106); (b) different outcomes from those in the conceptual framework (n = 53); (c) qualitative researches (n = 21); (d) different definition from those in the conceptual framework (n = 9); and (e) quantitative researches with no comparison analysis (n = 6). Different outcomes from those in the conceptual framework include subjective concepts as ”patient's satisfaction”, ”patient's perception of care”, and ”practitioner's satisfaction”. Different definitions from those in the conceptual framework include the definitions previously discussed (see Subchapter 1.2.1).

Finally, a set of 64 studies resulted from the screening process.

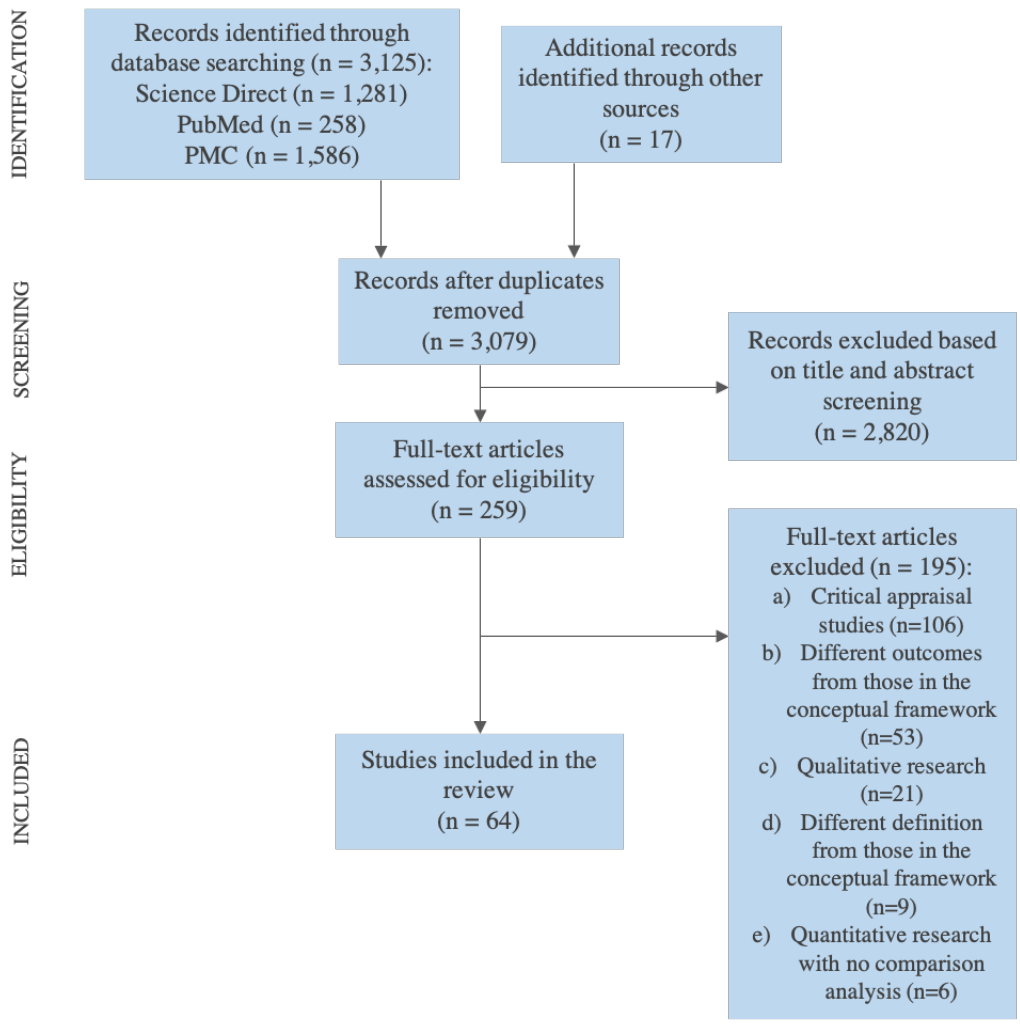


Figure 2.2. Screening diagram.

2.2.3 Post-screening/pre-basic statistics

After the screening process, one characterized each of the 64 studies regarding the following fields:

1. Author(s);
2. Type of vertical integration (backward, forward, or full integration) analysed by the author(s);
3. Methodology used by the author(s) to compare vertically integrated healthcare providers with others and to draw conclusions;
4. The sample brief description (country, type of healthcare providers under analysis, sample size, and data range);
5. Outcome(s) analyzed and respective variables;

6. Main conclusions drawn;

7. Major strengths and weaknesses.

Based on these data, one constructed a table providing an overview of the collected studies. The previously presented seven fields' choice aims to facilitate potential comparisons with other systematic reviews or other studies in the literature.

However, other relevant data were also collected but not displayed to avoid a too large manuscript. These data included:

1. Journal in which the study was published;

2. Number of citations per study;

2.2.4 Basic statistics over the collected studies

An initial selection of questions was elaborated to proceed with a fundamental statistic analysis over the collected studies:

1. Which methods are the most used when analysing vertical integration? Which methods are the most used when analysing each one of the outcomes regarding vertical integration?

2. Which variables are the most used when analysing each one of the outcomes regarding vertical integration?

3. What is the trend regarding the number of studies published per year?

4. Which type of vertical integration is the most studied?

5. Which countries do most studies belong to?

6. What are the journals in which there have been more studies published?

7. Which of the selected articles are most cited (Top 10)?

With the assist of some statistical tools, like pie charts and bar charts, and tables, this initial analysis was accomplished. It is important to refer that evidence concerning the health sector in which the healthcare providers' sample belongs (public, private, or PPP) was also searched for. However, that evidence was not gathered, as most studies are from the US, and those studies use undifferentiated samples straight from the insurers (*e.g.*, Medicaid, Medicare, to name a few), with no distinction between the health sector in the analysis.

2.2.5 Meta-analysis

Considering the large sample of studies included in the analysis, a complex statistic tool was needed to improve the results' validity. For that reason, one has developed a meta-analysis. This statistical procedure helps to establish statistical significance across the included studies and to introduce greater statistical power and more ability to extrapolate to the general population (Borenstein et al., 2009). A set of questions was gathered:

1. Is the methodology used related to the type of vertical integration?
2. Is the evidence regarding each one of the outcomes related to the type of vertical integration?
3. Is the evidence regarding each one of the outcomes related to the methodology used?
4. Is the methodology used related to the journal?
5. Is the type of vertical integration related to the journal?

SPSS Statistics (version 26) was used to analyze the association between several studies' characteristics: the type of vertical integration analyzed in each study, the methodology used, and the evidence regarding each outcome (whether there were increases, decreases, or none). With that purpose, the statistical analysis involved a chi-square test of independence and the measure of association Phi. The chi-square test of independence determines whether two categorical variables in a single sample are associated, using the following hypotheses: H_0 - the variables of interest are independent; H_1 - the variables of interest are associated. A significant test rejecting the null hypothesis (p-value < 0.05 in this study) would suggest that within the sample, the variables being analyzed are associated with each other (Franke et al., 2012). The Phi coefficient is additionally computed to understand whether the association is positive (most of the data falls along in the diagonal cells of the contingency table) or negative (most of the data falls off the diagonal cells of the contingency table) (Franke et al., 2012).

2.3 Main results

The following phase involved reviewing each of the 64 selected studies for its type of vertical integration, methodology, country, sample, data range, main conclusions, and strong/weak points. Table 2.2 provides studies that reached favourable evidence, Table 2.3, unfavourable evidence, and Table 2.4, inconsistent or statistically insubstantial evidence.

Table 2.2. Vertically integrated providers assessment: favourable evidence.

Study	VI: Backward or Forward	Methodology	Country, sample and data range	Outcome(s) analyzed	Main conclusions	Strong points	Weak points
Haddad et al. (2020)	Backward	Multivariable logistic regression and time-series analysis	US; unspecified number of healthcare providers (146,677 surgeons); 2007-17	Access (medicaid acceptance).	The affiliation between primary surgical and procedural practices and hospitals is connected with an increased access to surgical care for Medicaid patients, when comparing to traditional models.	Large number of observations (surgeons).	Only 1 access-related variable used.
Gupta et al. (2019)	Forward	GEE models	US; unspecified number of healthcare providers (35,935 hospital-year observations); 2007-12	Quality (readmission rate).	Integration between hospitals and SNFs reveal lower readmission rates for AMI and PN than non-integrated models.	—	Only 1 quality-related variable used.
Short and Ho (2019)	Backward	Fractional probit regression models	US; 4,438 hospitals (around 20,000 hospital-year observations[a]); 2008-15	Quality (29 variables: given aspirin at discharge; given most appropriate initial antibiotic(s); AMI readmission rates; having surgery and got the right kind of antibiotic; given an evaluation of left ventricular systolic function; ...).	Vertical integration may be significantly associated with gains on PN readmission rates and on a process of care quality measure (continuation of beta blockers for surgical patients), when compared to the traditional model.	4 different dimensions of backward vertical integration analyzed; large sample of hospitals; 29 quality-related variables analyzed.	—
Modi et al. (2019)	Backward	GLMMs	US; unspecified number of healthcare providers (6,381 urologists and 35,929 patients); 2011-14	Costs.	Hospitals or AMCs that acquired urologists practices show significantly lower spendings, when comparing to other urologists practices structures.	—	—
Leleu et al. (2017)	Backward	DEA	US; 1,847 hospitals; 2013	Efficiency (inputs: CMI; beds; post admission days; ...; outputs: nº of admissions, mortality rate; readmission rate; ...).	Hospitals with fully integrated physicians are more efficient than traditional hospitals.	Robust research methodology; large sample of hospitals.	One-year sample.
Lopes et al. (2017)	Backward	DID estimation	Portugal; 12 hospitals (1,597,159 admissions); 2004-13	Quality (readmission rate).	Hospitals integrated with primary care providers are connected with reduced readmissions, when comparing to similar non-integrated hospitals.	Large number of observations (admissions).	Small sample of hospitals; only 1 quality-related variable used.

Table 2.2. Continued.

Study	VI: Backward or Forward	Methodology	Country, sample and data range	Outcome(s) analyzed	Main conclusions	Strong points	Weak points
Herrel et al. (2017)	Backward	Multivariable logistic regression models	US; 40 health markets (21,085 patients); 2007-11	Quality (12 variables: received pretreatment counseling by a urologist and radiation oncologist; avoided imaging in patients with low risk prostate cancer; treated by a high volume provider; avoided treatment when life expectancy $<$10 years; experienced GI toxicity related to treatment; ...).	The integration between hospitals and oncologists practices exhibits better process of care quality measures (*e.g.*, multiple hospitalizations rate), when compared to non-integrated systems.	Large sample of health markets; 12 quality-related variables analyzed.	—
Crespin et al. (2016)	Full integration	Fixed effects regression model	US; 661 ambulatory clinics; 2007-13	Quality (5 variables: hemoglobin A1c percent; blood pressure percent; LDL cholesterol percent; mean of the daily aspirin; documented tobacco-free status indicators).	Ambulatory clinics (that offer primary or specialty ambulatory care) acquired by health systems exhibits improvements on clinic-level quality of care measurements regarding diabetes (*e.g.*, blood pressure percent, hemoglobin A1c percent), when compared to non-acquired ambulatory clinics.	—	—
Lanese (2016)	Backward	Paired sample t-test	US; 1 hospital; 2010-13	Quality (4 variables: CHF 30-day readmission rate; heart failure order set usage; aspirin at discharge; statin at discharge).	The co-management hospital improves quality across every single metric analyzed (*e.g.*, heart failure readmission rate, ASA at discharge, statin at discharge), as compared to baseline data.	—	Limited research methodology; small sample of hospitals.
Al-Amin (2016)	Backward	OLS regression model	US; 1,756 hospitals; 2013	Quality (readmission rate).	Both the hospitals with a fully integrated physician model and the hospitals exhibiting physician ownership are associated with lower readmission rates than hospitals without affiliation models.	2 different dimensions of backward vertical integration analyzed; large sample of hospitals.	Only 1 quality-related variable used; one-year sample.

Table 2.2. Continued.

Study	VI: Backward or Forward	Methodology	Country, sample and data range	Outcome(s) analyzed	Main conclusions	Strong points	Weak points
Rahman et al. (2016)	Forward	Regression models	US; 3,173 hospitals and 14,374 SNFs (827,541 Medicare beneficiaries); 2007	Costs; quality (2 variables: mortality; hospital readmissions).	Hospitals integrated with SNFs exhibit lower spendings than non-integrated systems, despite not showing significant effect on mortality or hospital readmissions.	Large sample of hospitals and SNFs; large number of observations (patients); 2 outcomes analyzed (costs and quality).	One-year sample.
Caballer-Tarazona and Vivas-Consuelo (2016)	Backward	DEA	Spain; 24 hospitals; 2009-10	Costs; efficiency (inputs: HR costs; nº of beds; nº of ORs; ...; outputs: adjusted surgical patients; adjusted admissions; adjusted outpatients; ...); quality (management agreement score obtained from 95 quality indicators).	The Alzira model, a PPP that integrates primary care and acute care, improves efficiency and some quality indicators (such as reduced delays in waiting lists), when comparing to public non-integrated systems, showing, however, unsubstantial evidence regarding costs.	Robust research methodology; 3 outcomes analyzed (costs, efficiency and quality).	Small sample of hospitals.
Carlin et al. (2015)	Not specified	DID estimation	US; 12 clinic systems (796,962 person-years of data); 2006-11	Quality (5 variables: preventive screening for breast, colorectal, and cervical cancer; inpatient admissions; ACS admissions; admissions followed by a readmission within 30 days; ED visits).	The integration between clinic systems and hospitals exhibit increases rates of colorectal and cervical cancer screening and more appropriate ED use, when comparing to non-integrated clinic systems.	Large number of observations.	Small sample of clinic systems.
Rhoads et al. (2015)	Full integration	Pearson's correlation coefficient analysis	US; 348 hospitals (33,593 patient records); 2001-06	Quality (mortality).	The integration among the three levels of care reveals lower 5-year mortality, when compared to non-integrated delivery systems.	—	Only 1 quality-related variable used.
Liepert et al. (2014)	Forward	Prospective cohort study	US; 2 trauma centers; 2009-10	Quality (rate of duplicated CTs).	Integration between hospitals and trauma centers (post-acute care) shows better quality (lower rate of duplicated CTs), when comparing to a traditional referral system.	—	Limited research methodology; small sample of trauma centers; only 1 quality-related variable used.

Table 2.2. Continued.

Study	VI: Backward or Forward	Methodology	Country, sample and data range	Outcome(s) analyzed	Main conclusions	Strong points	Weak points
Cho et al. (2014)	Backward	DEA and 2-stage least squares	US; 2,173 hospitals; 2010	Efficiency (inputs: nº of beds; service mix; FTE employees; ...; outputs: CMI admissions; outpatient visits).	Hospital-physician integration shows statistically significant gains on efficiency, when comparing to non-integrated systems.	Robust research methodology; large sample of hospitals.	One-year sample.
David et al. (2011)	Forward	OLS regression model	US; 2,571 hospitals; 2005	Quality (readmission rate).	The integration between hospitals and SNFs reveal lower hospital readmissions rates, when comparing to non-integrated systems.	Large sample of hospitals.	One-year sample; only 1 quality-related variable used.
Acerete et al. (2011)	Backward	Ratio analysis	Spain; 2 PPP contracts and Valencian public sector data; 1998-2008	Costs.	After a failed contract that lasted until 2002, a new PPP contract was developed, exhibiting lower costs than the traditional system in the Valencian region.	—	Limited research methodology; small sample of healthcare contracts.
Falces et al. (2011)	Backward	Software SPSS	Spain; 1 hospital and 7 primary care providers (3,194 patients); 2008-09	Quality (18 variables: indication of anti-aggregating treatment; beta blockers; renin-angiotensin system inhibitors and statins; performing cardiac catheterization and/or revascularization; performing echocardiogram; ...).	The integration of primary and secondary care within cardiology delivery care improves process of care quality measures regarding ischemic heart disease (*e.g.*, cholesterol control), heart failure (*e.g.*, prescription of beta blockers), and atrial fibrillation (*e.g.*, etiological study followed by echocardiography), when comparing to usual cardiology care.	18 quality-related variables analyzed.	Small sample of hospitals and primary care providers; small number of observations (patients).
Leibert (2011)	Not specified	ANOVA	US; 50 integrated systems and 50 non-system hospitals; 2009	Costs; quality (18 variables: not specified).	Integrated delivery systems exhibit higher performance regarding some process of care quality measures related with AMI, PN and HF, and similar costs, when compared to non-integrated systems.	2 outcomes analyzed (costs and quality); 18 quality-related variables analyzed.	One-year sample.

Table 2.2. Continued.

Study	VI: Backward or Forward	Methodology	Country, sample and data range	Outcome(s) analyzed	Main conclusions	Strong points	Weak points
Weeks et al. (2010)	Backward	Fixed-effects regression model	US; 20 medical groups (741,448 patients); 2005-06	Costs; quality (6 variables: ACS admissions; mammographic screening for breast cancer for women ages 65–69; annual hemoglobin A1c testing among diabetics ages 65–75; annual lipid testing among diabetics ages 65–75; annual fundoscopy (examination of the retina) by an oph-thalmologist among diabetics ages 65–75; ...).	The affiliation of hospitals and/or health plans with physicians shows better quality performance (*e.g.*, ACS admission rate) and lower costs, when compared to non-integrated models.	Large number of observations (patients); 2 outcomes analyzed (costs and quality).	Small sample of medical groups.
Rosko et al. (2007)	Full integration	SFA	US; 1,144 hospitals; 2001	Efficiency (inputs: prices; outputs: outpatient visits; inpatient admissions).	The full integration of the delivery care system reveals decreased cost inneficiency, when compared to independent hospital systems.	Robust research methodology; large sample of hospitals.	One-year sample.
Ciliberto and Dranove (2006)	Backward	Regression model	US; 320 hospitals (2,541,086 patients); 1994–2001	Price of care.	In general, hospital-physician affiliation is not associated with significant changes in the price of care, when compared to non-affiliated systems. Yet, a small subset of hospitals (rural hospitals) shows significant decreases in the price of care.	Large sample of observations (patients).	—
Gillies et al. (2006)	Backward	Multivariate cross-sectional regression model	US; 272 health plans; 2003	Quality (12 variables: breast cancer screening; adolescent immunization screening; controlling high blood pressure; LDL-C level (comprehensive diabetes care); LDL-C level (cholesterol management after acute events); ...).	The integration between hospitals and physicians practices show better results on process of care quality measures (*e.g.*, breast cancer screening, adolescent immunization screening, controlling high blood pressure), when compared to non-integrated models.	12 quality-related variables analyzed.	One-year sample.

Table 2.2. Continued.

Study	VI: Backward or Forward	Methodology	Country, sample and data range	Outcome(s) analyzed	Main conclusions	Strong points	Weak points
Ugolini and Nobilio (2003)	Full integration	Logistic regression model	Italy; 1 hospital; 1997-2001	Quality (mortality).	The cardiovascular delivery care system that integrates the three levels of care exhibits lower mortality rate, when compared to data before integration.	—	Small sample of hospitals; only 1 quality-related variable used.
Chu et al. (2003)	Backward	DEA	Taiwan; 1 hospital; 1995-97	Efficiency (inputs: personnel, medicine and depreciation costs; outputs: total revenue).	The hospital-physician integration strategy analyzed (creation of responsibility centers) improves the hospital's revenue efficiency, when compared to pre-strategy data.	Robust research methodology.	Small sample of hospitals.
Chu et al. (2002)	Backward	DEA	Taiwan; 90 hospitals; 1994-96	Efficiency (inputs: physicians; nurses; hospital beds; ...; outputs: ambulatory and emergency visits; inpatient days; inpatient visits).	The hospital-physician integration strategies analyzed (responsibility centers, total quality management and physician fee programs) improves the hospital's efficiency, when compared to non-integrated hospitals.	3 different dimensions of backward vertical integration analyzed; robust research methodology.	Small sample of hospitals.
Wang et al. (2001)	Backward and Forward	SEM technique	US; 363 hospitals; 1994	Efficiency.	The integration of hospitals with LTCFs is positively and substantially related to hospital efficiency of production, as indicated by *e.g.* adjusted admissions per bed. The integration between hospitals and physician practices is statistically less likely to improve efficiency of production.	The 2 types of vertical integration are assessed independently.	One-year sample.
Brickman et al. (1998)	Full integration	Prospective cohort study	US; 5-hospital MCO; 1994-96	Costs; quality (not specified).	The creation of a MCO (including 5 hospitals, physician practices, LTCFs, ...) improved process of care quality measures regarding multiple conditions (*e.g.*, PN, stroke) and decreased costs of care, when comparing to previous data.	2 outcomes analyzed (costs and quality).	Limited research methodology; small sample of organizations.

[a] the number changes slightly throughout the study years.

Table 2.3. Vertically integrated providers assessment: unfavourable evidence.

Study	VI: Backward or Forward	Methodology	Country, sample and data range	Outcome(s) analyzed	Main conclusions	Strong points	Weak points
Ho et al. (2020)	Backward	Regression models	US; unspecified number of healthcare providers (11,444 physicians); 2014–16	Costs.	Hospitals with the ownership of physician practices exhibits higher annual spending per patient than traditional hospitals.	—	—
Jung et al. (2019)	Backward	DID estimation	US; unspecified number of healthcare providers (81,899 Medicare beneficiaries); 2009-13	Costs.	The affiliation between hospitals and oncology physician practices exhibits higher spendings regarding drugs and chemotherapy administration (the act of injection), when compared to pre-integration data.	—	—
Ho et al. (2019)	Backward	Regression models	US; around 1,900 POs[a] (approximately 600,000 patients[a]); 2014-16	Costs; quality (3 variables: readmission within 30 days of discharge for hospitalized patients; appropriate care for diabetic patients; screening mammography for women ages 50–64).	The integration between hospitals and physician practices are associated with higher per patient spendings, yet no changes are observed regarding quality measurements (*e.g.*, readmission rates), when compared to non-integrated physician practices.	Large sample of POs; large number of observations (patients); 2 outcomes analyzed (costs and quality).	—
Capps et al. (2018)	Backward	DID estimation	US; unspecified number of healthcare providers (around 200,000 physician-year observations and 3.3 million enrollee-year observations); 2007-13	Costs; price of care.	Hospitals that acquire physician practices are associated with higher total per-enrolee spending and higher care prices, when comparing to pre-acquisition data.	Large number of observations; 2 outcomes analyzed (costs and prices).	—
Cho et al. (2018)	Backward	OLS regression model	US; 154 hospitals (468 hospital-year observations); 2002-06	Costs.	Expenditure is higher in hospital-physician affiliations than in other independent arrangements.	—	Small number of observations.

Table 2.3. Continued.

Study	VI: Backward or Forward	Methodology	Country, sample and data range	Outcome(s) analyzed	Main conclusions	Strong points	Weak points
Chunn et al. (2018)	Backward	Regression models	US; unspecified number of healthcare providers (4,830 cardiologists and 13,642 physician-year observations); 2010-14	Costs; efficiency.	The acquisition of cardiologists practices by hospitals increased spendings (related with increments in physicians compensation) and decreased the clinical efficiency of production.	2 outcomes analyzed (costs and efficiency)	—
Rossiter (2018)	Backward	Regression models	US; unspecified number of healthcare providers (around 2 million patients[a]); 2013-15	Costs; quality (8 variables: emergent/emergency department care needed/preventable/ avoidable; avoidance of antibiotic treatment in adults with acute bronchitis; adolescent well-care visits; chlamydia screening in women; eye examination; ...).	Hospital-affiliated physicians exhibit substantially higher per patient spendings, yet similar process of care quality measures, when compared to independent physicians.	Large number of observations (patients); 2 outcomes analyzed (costs and quality).	—
Pesko et al. (2018)	Backward	Regression models	US; unspecified number of healthcare providers (3,010 physicians and 282,372 beneficiaries); 2008	Costs; quality (3 variables: emergency department visits; index admissions; readmissions).	Hospital-physician affiliations show higher spendings and similar quality-related measures (*e.g.*, readmission rates), when compared to independent physicians.	Large number of observations (patients); 2 outcomes analyzed (costs and quality).	One-year sample.
Koch et al. (2017)	Backward	DID estimation	US; 27 mergers (2.5 million Medicare beneficiaries); 2005-10	Costs.	Hospitals acquiring physician practices exhibit an increase in spendings, when comparing to pre-acquisition data.	Large number of observations (patients)	Small sample of mergers.
Chukmaitov et al. (2015)	Backward	Regression models	US; 189 hospitals (5.4 million discharges); 2006-09	Costs; quality (mortality).	Integrated hospital-physician arrangements manifest increased mortality, yet no statistically significant changes in costs, when comparing to other types of arrangement	Large number of observations (discharges); 2 outcomes analyzed (costs and quality).	Only 1 quality-related variable used.

Table 2.3. Continued.

Study	VI: Backward or Forward	Methodology	Country, sample and data range	Outcome(s) analyzed	Main conclusions	Strong points	Weak points
Neprash et al. (2015)	Backward	Regression models	US; unspecified number of healthcare providers (7,391,335 enrollees); 2008-12	Price of care.	Hospital-physician affiliation exhibits higher outpatient care prices and no price differences for inpatient care, compared to data preceding the affiliation.	Large number of observations (patients).	—
Casalino et al. (2014)	Backward	Regression models	US; unspecified number of healthcare providers (284,401 beneficiaries); 2007-09	Quality (preventable admissions).	Hospital-owned practices show higher preventable admissions than physician-owned practices.	Large number of observations (patients).	Only 1 quality-related variable used.
Baker et al. (2014)	Backward	Regression models	US; unspecified number of healthcare providers (2.1 million hospital claims); 2001-07	Costs; price of care.	Hospital ownership of physician practices leads to higher levels of hospital spendings and care prices, when compared traditional arrangements.	Large number of observations (claims); 2 outcomes analyzed (costs and prices).	—
Robinson and Miller (2014)	Backward	Regression models	US; 158 medical groups (4.5 million patients); 2009-12	Costs.	Hospital-physician affiliations exhibits higher expenditures per patient than independent physician practices.	Large number of observations (patients).	—
Kralewski et al. (2013)	Backward	ED and ACS algorithms (from Agency for healthcare Research and Quality)	US; 212 medical group practices (615,400 Medicare enrollees); 2009	Quality (3 variables: nonemergent ED rate; ED rate; ACS hospital rate).	The integration between hospitals and physicians practices have higher nonemergent and emergent primary care treatable ED rates and higher ACS hospitalization rates, when compared to non-integrated systems.	Large number of observations (patients).	One-year sample.
McWilliams et al. (2013)	Backward	Linear regression models	US; unspecified number of healthcare providers (4.29 million beneficiaries); 2009	Costs; quality (7 variables: 30-day readmission; screening mammography; LDL cholesterol testing; hemoglobin A1c testing; retinal exam; ...).	Hospital-based primary care physician groups reveals higher spendings, higher readmission rates and similar performance on 4 quality process of care quality measures, when compared to independent primary care physician groups.	Large number of observations (patients); 2 outcomes analyzed (costs and quality).	One-year sample.

Table 2.3. Continued.

Study	VI: Backward or Forward	Methodology	Country, sample and data range	Outcome(s) analyzed	Main conclusions	Strong points	Weak points
Kralewski et al. (2012)	Backward	Regression models	US; 234 medical group practices (133,703 Medicare enrollees); 2009	Costs; quality (4 variables: patients with diabetes who had an LDL lab test during the past year; patients with diabetes who had a kidney function lab test during the past year; proportion of inappropriate ED visits; ambulatory care–sensitive hospital rate).	Hospitals integrated with physician practices have significantly higher costs and similar performance regarding diabetes quality measures (*e.g.*, proportion of inappropriate ED visits, ambulatory care–sensitive hospital rate), when compared to independent physician practices.	Large number of observations (patients); 2 outcomes analyzed (costs and quality).	One-year sample.
Cuellar and Gertler (2006)	Backward	SFA	US; around 1,000 hospitals[b]; 1994-98	Price of care; efficiency (inputs: prices.; outputs: nº of hospital admissions; average length of stay; nº of outpatient visits; ...).	Affiliations between physician practices and hospitals shows substantial increases in prices of care and no significant changes in efficiency, when compared to other delivery care models.	3 different dimensions of backward vertical integration analyzed; robust research methodology; large sample of hospitals; 2 outcomes analyzed (prices and efficiency).	—
Alexander and Morrisey (1988)	Backward	OLS regression model	US; 3,027 hospitals; 1982	Costs.	4 of the 5 dimensions of hospital-physician integration studied are associated with higher hospital costs, when compared to the traditional delivery care system.	5 dimensions of hospital-physician affiliation analyzed; large sample of hospitals.	One-year sample.

[a] the number changes slightly throughout the study years.

[b] the number changes slightly depending on the outcomes.

Table 2.4. Vertically integrated providers assessment: inconsistent or statistically insubstantial evidence.

Study	VI: Backward or Forward	Methodology	Country, sample and data range	Outcome(s) analyzed	Main conclusions	Strong points	Weak points
Yuan et al. (2019)	Backward	Multiple linear regression models and ANOVA	China; unspecified number of healthcare providers (1,118 patients); 2016-18	Quality (primary care quality scores derived from surveys).	Primary care, when integrated with hospitals, showed no significant difference on quality of care to loose collaboration systems.	—	Small number of observations (patients); quality measures derived from surveys.
Comendeiro-Maaløe et al. (2019)	Backward	Logistic multivariate multilevel regression models and SFA	Spain; 67 hospitals; 2003 and 2015	Efficiency (inputs: functioning beds; FTE physicians; FTE nursing staff; ...; outputs: discharges weighted by DRG ; outpatient activity.); quality (9 variables: potentially Avoidable Hospitalisations; low value surgical procedures; orthopaedics; AMI rate; percutaneous coronary rate).	The integration between primary care providers and hospitals (through PPP contracts) exhibits contradictory performance levels regarding both quality and efficiency measures, when comparing to similar public providers.	Robust research methodology; 2 outcomes analyzed (efficiency and quality).	Small sample of hospitals.
Henke et al. (2018)	Backward	Retrospective longitudinal regression model	US; 3,957 hospitals (99 million discharges); 2010-12	Costs; quality (5 variables: in-hospital mortality; admission for AMI; admission for HF; admission for acute stroke; admission for PN; uncomplicated cesarean delivery rate).	The integration between hospitals and physician practices exhibit higher costs, better levels of AMI, heart failure and PN mortality, when compared to non-integrated models.	Large sample of hospitals; large number of observations (discharges); 2 outcomes analyzed (costs and quality).	—
Konetzka et al. (2018)	Forward	Regression models	US; 4 011 hospitals, 16,251 SNFs and 12,720 HHAs; 2005-13	Quality (readmission rates).	vertical integration between hospitals and different post-acute care providers shows contradictory results in terms of readmission rates, when comparing to traditional models.	Large sample of healthcare providers.	—
Casalino et al. (2018)	Backward	Cross-sectional observational analysis	US; unspecified number of healthcare providers (31,888 physicians and 868,213 Medicare beneficiaries); 2012-13	Costs; quality (2 variables: readmissions; ACS admissions).	Hospital-owned practices exhibit similar total spending and quality-measures (*e.g.*, readmissions) to physician-owned practices.	Large number of observations (patients); 2 outcomes analyzed (costs and quality).	Limited research methodology.

Table 2.4. Continued.

Study	VI: Backward or Forward	Methodology	Country, sample and data range	Outcome(s) analyzed	Main conclusions	Strong points	Weak points
Li et al. (2018)	Full integration	GEE model	US; unspecified number of healthcare providers (380,053 enrollees); 2011-13	Quality (4 variables: readmission rates; 30-day mortality; surgical complications; failure to rescue).	Hospitals within an integrated delivery system shows similar performance on quality-related measures (*e.g.*, 30-day mortality), when compared to traditional systems.	Large number of observations (patients).	—
West et al. (2017)	Backward	Linear model regression	US; 6,006 hospitals; 2006-13	Costs; quality (2 variables: readmission rates; mortality).	Hospital-physician affiliations exhibit substantial improvements on mortality and readmissions, in parallel to higher operating costs, when comparing to other hospitals.	Large sample of hospitals; 2 outcomes analyzed (costs and quality).	—
Scott et al. (2017)	Backward	Regression models	US; 2,888 hospitals; 2002-13	Quality (2 variables: readmission rates; mortality).	Physician employment alone do not substantially improve quality of care measurements (*e.g.* mortality-related metrics and readmission rates), when compared to other delivery care models.	2 different dimensions of backward vertical integration analyzed; large sample of hospitals.	—
Alonso et al. (2014)	Backward	DEA	Spain; 25 hospitals; 2009	Efficiency (inputs: nº of beds; nº of FTE physicians; nº of FTE nursing staff.; outputs: nº of discharges; nº of outpatient visits).	Hospitals integrated with primary care facilities do not show significant statistical differences from traditional hospitals regarding efficiency.	Robust research methodology.	Small sample of hospitals; one-year sample.
Curry et al. (2013)	Full integration	SPSS software	England; 1 health system (2 hospitals, 3 community healthcare service providers, ...); 2011-12	Quality (4 variables: HbA1c control; cholesterol control; blood pressure control; dementia-case finding).	The full integration across all the levels of care (mostly directed to elderly) shows similar process of care quality measures (*e.g.*, blood pressure control, dementia-case finding) to England baseline data.	—	Small sample of health systems.
Chukmaitov et al. (2009)	Full integration	Logistic regression models	US; 1,100 hospitals; 1995-2000	Quality (mortality).	Integrated hospital systems exhibit contradictory results regarding AMI, CHF and PN mortality, when compared to independent hospital systems.	Large sample of hospitals.	Only 1 quality-related variable used.
Huckman (2005)	Backward	Regression models	US; unspecified number of healthcare providers (2.3 million discharges); 1992-2000	Costs; quality (mortality).	The acquisition of cardiologists practices by hospitals have relatively little impact on average cost and quality of major cardiac procedures, when compared to traditional hospitals.	Large number of observations (discharges); 2 outcomes analyzed (costs and quality).	—

Table 2.4. Continued.

Study	VI: Backward or Forward	Methodology	Country, sample and data range	Outcome(s) analyzed	Main conclusions	Strong points	Weak points
Madison (2004)	Backward	Multivariate regression analysis	US; approximately 4,000 hospitals[a]; 1994-98	Quality (2 variables: mortality; readmission rates).	The affiliation between physician practices and hospitals do not provide significant changes in terms of patient outcomes, specifically regarding mortality, when compared to other delivery care models.	3 different dimensions of backward vertical integration analyzed; large sample of hospitals.	—
Carey (2003)	Backward	SFA	US; 1,209 hospitals; 1998	Efficiency.	Hospital-physician affiliations exhibits similar efficiency to traditional systems.	Robust research methodology; large sample of hospitals.	One-year sample.
Kralewski et al. (2000)	Backward	Two-stage regression model	US; 86 clinics; 1995	Costs.	The control on clinic physicians' referrals shows no effect on per member per year adjusted patient costs.	—	Small sample of clinics; one-year sample.
Goes and Zhan (1995)	Backward	Longitudinal regression models	US; 300 hospitals; 1981-1990	Costs.	The three kinds of hospital-physician integration strategy studied yielded contradictory financial performance results, namely regarding spendings, when comparing to hospitals that operate independently from physicians.	3 dimensions of hospital-physician affiliation analyzed.	—

[a] the number changes slightly throughout the study years.

2.3.1 Basic statistics over the collected studies

The fundamental statistical analysis of the collected studies was elaborated and accompanied by charts, graphs, and tables.

Using frequentist inference, one verified that the most used methodology to analyze vertical integration is the regression model, followed by DEA and Difference-in-difference (DID) estimation (Figure A.1). These three methods account for approximately 65% of the sample of methodologies (which consists of 68 methodologies, as there are studies that use more than one).

When analyzing each outcome separately, as visualized in Table A.1, one might encounter the same result for four of the five outcomes: costs, prices of care, quality of care, and access. However, concerning efficiency, DEA is the most used, followed by SFA, which is consistent with Cylus et al. (2016). It seems logical that, as it consists of a more complex outcome, more robust methodologies are required.

One additional subject that should be addressed is the trend regarding the number of studies published per year. In the graph of Figure A.2, one might observe the cumulative number of published studies (from the set of 64) since 1988, the year of publishing of the first study included in the analysis. By exploring Figure A.2a, one might observe that from 2011 to 2019, there is an apparent increase in the number of studies published per year. This means that the topic of vertical integration of healthcare providers is gaining significance in the literature. In Figure A.2b, one presents the cumulative number of studies until 2020. Using a MATLAB function (polyfit), one may hypothesize that in around 15 years, the sample would include approximately 100 articles.

When examining the type of vertical integration across all the studies, one may realize that backward integration is the most studied strategy, followed by full integration and then forward integration (Figure A.3). Curiously, the integration between primary care facilities and hospitals represents around 77% of the sample.

Furthermore, Figure A.4 confirms the US as the most analyzed country. The European countries (Spain, Italy, Portugal, and England) and the Asian countries (Taiwan and China) represent only 12.5% and 4.7% of the sample, respectively, showing small representativeness.

Table A.2 exhibits the journals that published more studies from the sample. The journals ”Health Services Research”, ”Journal of Health Economics”, and ”Health Affairs” were the ones that most published, representing around one-third of the sample. All of the three journals have substantial impact factors (2019-20) and Scimago H-indexes.

By searching on the Google Scholar platform, it was possible to gather information about the citations of each one of the articles included in the sample. Table A.3 exhibits the top 10 most cited articles in descending order. Interestingly, all of those articles examined backward integration, which reinforces the results of Figure A.3 that backward integration is the most studied type of vertical integration, and, potentially, the most used in healthcare models.

2.3.2 Meta-analysis

As previously announced, a chi-square test of independence and the measure of association Phi are used to answer the questions presented in Subchapter 2.2.5. Tables A.4-A.7 display the significant correlation results, *i.e.* with p-value < 0.05. Significant associations between variables are signaled with an "X", referring to the respective p-values and phi coefficients. Comparisons for every variable were computed. Nonetheless, tables with no statistically significant results for the significance level of 5% were omitted.

By observing tables A.4-A.7, the questions enunciated in Subchapter 2.2.5 may be answered as follows:

1. **Is the methodology used related to the type of vertical integration?** Yes. More specifically, regression models appear to be positively associated with backward integration, DID estimations with articles that do not specify the type of vertical integration, and the aggregate of other methodologies appears to be negatively correlated with backward integration.

2. **Is the evidence regarding each one of the outcomes related to the type of vertical integration?** Yes. Backward integration appears to be positively associated with cost increases (and negatively correlated with cost decreases), and both forward integration and full integration appear to be positively associated with costs decreases.

3. **Is the evidence regarding each one of the outcomes related to the methodology used?** Yes. DEA is positively associated with efficiency increases, SFA with efficiency unchanged, and the aggregate of other methodologies with costs decreases (and negatively associated with costs increases).

4. **Is the methodology used related to the journal?** No, none of the methodologies appears to be related to the journal (only the three journals with most studies were considered), the reason why no tables were presented regarding this question.

5. **Is the type of vertical integration related to the journal?** No, none of the types of vertical integration appears to be related to the journal (only the three journals with most studies were considered), the reason why no tables were presented regarding this question.

2.4 Discussion

Sixty-four quantitative studies that analyze the performance of vertically integrated systems, when compared to non-integrated healthcare providers, were analyzed.

A recent systematic review (Machta et al., 2019), using a sample of twenty-nine papers, concluded that vertically integrated providers might have "*both positive and negative effects on policy-relevant outcomes*". These results may suggest that the assessment of vertical integration may be more complex than expected. Thus, its implementation must be thought carefully and must consider each integrated healthcare provider's environmental circumstances.

Throughout this Subchapter, the effects of vertical integration on costs, prices of care, efficiency, quality of care, and access are scrutinized.

2.4.1 Impact of vertical integration on costs

Twenty eight studies described the effects of vertical integration on costs, sixteen of which exhibited increments, five, decrements, and seven, unchanged costs (see Table 2.5).

Table 2.5. Empirical evidence on costs.

Increases (n = 16)	Decreases (n = 5)	Unchanged (n = 7)
Ho et al. (2020); Jung et al. (2019); Ho et al. (2019); Capps et al. (2018); Henke et al. (2018); Cho et al. (2018); Chunn et al. (2018); Rossiter (2018); Pesko et al. (2018); Koch et al. (2017); West et al. (2017); Baker et al. (2014); Robinson and Miller (2014); McWilliams et al. (2013); Kralewski et al. (2012); Alexander and Morrisey (1988).	Modi et al. (2019); Rahman et al. (2016); Acerete et al. (2011); Weeks et al. (2010); Brickman et al. (1998).	Casalino et al. (2018); Caballer-Tarazona and Vivas-Consuelo (2016); Chukmaitov et al. (2015); Leibert (2011); Huckman (2005); Kralewski et al. (2000); Goes and Zhan (1995).

Within this outcome, one might make a slight distinction. Patient-level costs and hospital-level costs may be analyzed independently. With this, one may inspect both hospital-related effects and the impact on the care payments (which can be done either by the patients or by health insurance programs).

From the sample shown in Table 2.5, a total of ten studies analyzed hospital-costs changes. Five reached conclusions towards costs increases for the hospitals. The studies reached those conclusions by analyzing operating costs (Chunn et al., 2018; West et al., 2017), and even in a more peculiar way, hospital cost per discharge (Henke et al., 2018; Alexander and Morrisey, 1988) and hospital expenditure per patient (Cho et al., 2018). Several justifications for those increases appear in the literature: a mixture of higher investments in infrastructure, technology, and clinical staff (Henke et al., 2018), increments in physician compensation, mainly to counterbalance the transition of the billing process from physicians' offices to hospitals (Chunn et al., 2018; West et al., 2017), or even over-employment of physicians in administrative boards (Cho et al., 2018; Alexander and Morrisey, 1988). On the other pole, Acerete et al.

(2011) came across reduced labor costs. Nevertheless, by scrutinizing this last study, one may affirm that the financial savings might be derived from an excessive focus on more lucrative specialties, lacking in others. Finally, four papers noticed similar hospital-level costs when comparing vertically integrated systems with traditional ones. Chukmaitov et al. (2015) and Leibert (2011) used the cost per discharge measure to reach those results, while Caballer-Tarazona and Vivas-Consuelo (2016) analyzed staff, medical and pharmacy costs, and Goes and Zhan (1995), operating expenses.

The remaining eighteen papers looked over patient-level costs. Eleven came across costs increases (Ho et al., 2020; Jung et al., 2019; Ho et al., 2019; Capps et al., 2018; Rossiter, 2018; Pesko et al., 2018; Koch et al., 2017; Baker et al., 2014; Robinson and Miller, 2014; McWilliams et al., 2013; Kralewski et al., 2012). Patient-level expenses are a combination of "utilization" and prices of care. Several researchers conjectured exhaustively about those results. In some cases, prices of care may be the trigger of that increase (Capps et al., 2018; Baker et al., 2014), while in other cases, the degree of care utilization may be the main factor (Jung et al., 2019; Ho et al., 2019; Koch et al., 2017). While one discusses prices of care increases in more detail in Subchapter 2.4.2, one may focus now on the degree of utilization. It is straightforward that most models that integrate primary care physicians with hospitals convert physicians into hospitals' stakeholders. This reality changes physicians' treatment patterns (recognized by Jung et al. (2019)), making them tendentious to use more material resources and many treatments. This fact should not be generalized to every single case. Nevertheless, several studies glanced over physicians' stake in the hospital's financial stability, such as Ho et al. (2020, 2019), Pesko et al. (2018), and Baker et al. (2014), to name a few. Apart from that, one other study observed higher spendings in the sequence of the consolidation of specialists with fee-for-service payments. This model is known for being slightly costlier (McWilliams et al., 2013). Contrarily, Modi et al. (2019), Rahman et al. (2016), Weeks et al. (2010), and Brickman et al. (1998) uncovered lower patient-level spendings. Not every single study was able to determine the cause of those savings. However, Rahman et al. (2016) state that the increased coordination between hospitals and skilled nursing facilities (SNFs) decreases SNF-related costs, creating overall savings. Finally, three papers (Casalino et al., 2018; Huckman, 2005; Kralewski et al., 2000) noticed similar patient-level costs when comparing vertically integrated systems with traditional ones.

A general trend to higher total spending on care, which may be more substantial at the patient-level than at the hospital-level, is pictured. As previously stated, a variety of backward vertical integration models (mainly the ones enabling a tighter form of integration) adds physicians to the stakeholders' pot. This reality may explain increases in care utilization due to changes in treatment patterns, which increases patient-level costs (Jung et al., 2019). Additionally, by integrating physicians in administrative boards, physician compensation may also increase. Whether for that reason or just for financial incentives to adapt to a new system, the truth is that, in this sample, hospitals' costs related to labor also exhibit an increasing trend.

The previous assertion refers to backward integration (most of the sample, which can be confirmed in Table A.5). However, in opposition, forward integration appears to be positively correlated with cost decreases by observing the same table. By looking over Tables 2.2-2.4, it is easy to confirm that only one study analyzes costs in a forward integration model, meaning the sample is not significant. However, although it cannot be generalized, that study, developed by Rahman and his colleagues, reveals one might achieve savings via lower spendings on a hospital-based SNF patient in the 30 days following their original hospital discharge. Therefore, there may be a positive impact of forward integration in costs of care, which can be confirmed with further future studies.

2.4.2 Impact of vertical integration on prices of care

Five studies described the effects of vertical integration on prices of care, four of which exhibited increments and one, decrements (Table 2.6).

Table 2.6. Empirical evidence on prices of care.

Increases (n = 4)	Decreases (n = 1)	Unchanged (n = 0)
Capps et al. (2018); Neprash et al. (2015); Baker et al. (2014); Cuellar and Gertler (2006).	Ciliberto and Dranove (2006).	—

As stated in Subchapter 2.4.1, some studies reveal prices of care increments in the analyzed sample when in the presence of vertical integration. The price of care appears to be one potential cause of increments in patient-level expenses. Nonetheless, it is vital to understand further the whole context of the increase in care prices and potential causes.

According to Table 2.6, four studies came across increases in prices of care. Capps et al. (2018) found out that facility fees may be the primary precursor of that increase. Those hospital fees are not billed, evidently, under care delivered by independent physician practices. In turn, Neprash et al. (2015), Baker et al. (2014), and Cuellar and Gertler (2006) agreed that hospitals might take advantage of their market power to increase prices. Baker et al. (2014) further add that vertically integrated hospitals may even see their market power increased. Hospital-physician affiliations enable hospitals to gather control over a wider variety of services or even deprive competitors of a "source of" or "destination for" referrals. Whether market power is an objective of integration or just an unintentional consequence of it, it is, nonetheless, a possible cause for the care prices increases. Contrariwise, Ciliberto and Dranove (2006) verified price decreases, which were substantially larger among rural hospitals, possibly induced by a double marginalization effect.

This relatively small sample uncovers a prevailing tendency to higher prices of care when in the presence of vertically integrated delivery systems. The exploitation by the integrated delivery systems of their market power stands out as the most probable cause of this increase.

2.4.3 Impact of vertical integration on efficiency

Twelve studies described the effects of vertical integration on efficiency, seven of which exhibited increments, one, decrements, and four, unchanged efficiency (Table 2.7).

Table 2.7. Empirical evidence on efficiency.

Increases (n = 7)	Decreases (n = 1)	Unchanged (n = 4)
Leleu et al. (2017); Caballer-Tarazona and Vivas-Consuelo (2016); Cho et al. (2014); Rosko et al. (2007); Chu et al. (2003); Chu et al. (2002); Wang et al. (2001).	Chunn et al. (2018).	Comendeiro-Maaløe et al. (2019); Alonso et al. (2014); Cuellar and Gertler (2006); Carey (2003).

By observing Table A.8 (an additional table that scrutinizes the studies included in the present subchapter), one might notice that from the twelve studies that study efficiency, ten analyzed technical efficiency, and two analyzed the efficiency of production. As already mentioned, the study of hospitals' technical efficiency implies using a set of inputs (*e.g.*, costs, prices, number of beds) and outputs (*e.g.*, number of admissions, emergency department visits). In turn, the efficiency of production, also known as productivity, is a relatively broader concept, which consists of the combination of technical efficiency, "doing things right", and effectiveness, "doing the right things" (Fried et al., 2008). It is even possible, for that reason, to analyze hospitals' productivity through the measurement of technical efficiency. Due to the substantial difference between concepts and their legitimacy in analyzing healthcare providers' performance, one should analyze them independently.

Regarding technical efficiency, six studies noticed increments (Leleu et al., 2017; Caballer-Tarazona and Vivas-Consuelo, 2016; Cho et al., 2014; Rosko et al., 2007; Chu et al., 2003, 2002), and four came across unsubstantial changes (Comendeiro-Maaløe et al., 2019; Alonso et al., 2014; Cuellar and Gertler, 2006; Carey, 2003). In every case, researchers analyzed efficiency using the most frequent techniques in benchmarking (DEA or SFA). Due to the robustness frequently associated with these methods, there may be strong support for potential conclusions about technical efficiency variations upon vertical integration. The study of technical efficiency was largely diversified between researchers. Leleu et al. (2017), for example, stood out by considering both "good" outputs (*e.g.*, number of admissions, number of inpatient surgeries) and "bad" outputs (*e.g.*, mortality rates) to "escape" upward biased conclusions. Moreover, Chu et al. (2003) studied a pure revenue definition of efficiency (by using costs as inputs and total revenue as a single output), establishing, for that reason, a way of additionally analyzing financial performance. Chu and his colleagues also emerge, both in 2002 and 2003, for specifically analyzing hospital-physician affiliation strategies, as responsibility centers systems (2002 and 2003), total quality management (2002), and physician fee programs (2002). In a general way, studies verified that tighter forms of vertical integration are those who exhibit higher efficiency, whether by frequently letting physicians influence about 80% of total healthcare costs (Leleu et al., 2017) or by making it easier to acknowledge and communicate the truly extraordinary best practices (Rosko et al., 2007).

The study of the broader measure of hospital's performance, the efficiency of production, provided a research with positive evidence (Wang et al., 2001) and one other with negative (Chunn et al., 2018). Unlike in the case of technical efficiency, none of the studies used non-parametric methodology. Wang et al. (2001) used structural equation modeling (SEM) techniques to analyze indicators of productivity (adjusted admissions per bed and adjusted admissions per full-time employee). The researchers analyzed backward and forward integration independently, reaching conclusions on positive changes only on the latter. Backward integration showed unsubstantial changes regarding productivity. In turn, Chunn et al. (2018) came across productivity decrements in the presence of cardiologists' integration with hospitals. Chunn and his colleagues reached those conclusions using work relative value units (survey data) as inputs of regression models. Despite these results, the authors argue that the study may have been done on an early timeline and that the results could be different after a more prolonged adaptation phase.

In addition to the efficiency scores from DEA or SFA, one additional indicator, length of stay, is often used to draw conclusions about efficiency (Jones, 2009). By emphasizing this indicator, one may get an idea of the hospital's functionality. This indicator may also influence costs, as a shorter stay represents a faster shift from inpatient settings to post-acute settings, which are less costlier (Organisation for Economic Co-operation and Development, 2019). By observing Table A.9, one may verify contradictory results regarding this indicator. Nevertheless, it seems that a lower length of stay is out of the question.

Through this analysis, one might state that decreases in technical efficiency upon vertical integration are practically out of the question. Still, there is no assurance nor a substantial inclination to visualize a positive influence. Regarding the efficiency of production, given the small sample and the contradictory results, there is no room for meaningful direct conclusions, even though technical efficiency is within the concept of efficiency of production.

2.4.4 Impact of vertical integration on quality of care

Thirty eight studies described the effects of vertical integration on quality of care, twenty of which exhibited increments, four, decrements, and fourteen, unchanged quality of care (Table 2.8).

Table 2.8. Empirical evidence on quality of care.

Increases (n = 20)	Decreases (n = 4)	Unchanged (n = 14)
Short and Ho (2019); Gupta et al. (2019); Henke et al. (2018); Lopes et al. (2017); Herrel et al. (2017); West et al. (2017); Caballer-Tarazona and Vivas-Consuelo (2016); Al-Amin (2016); Crespin et al. (2016); Lanese (2016); Carlin et al. (2015); Rhoads et al. (2015); Liepert et al. (2014); David et al. (2011); Falces et al. (2011); Leibert (2011); Weeks et al. (2010); Gillies et al. (2006); Ugolini and Nobilio (2003); Brickman et al. (1998).	Chukmaitov et al. (2015); Casalino et al. (2014); Kralewski et al. (2013); McWilliams et al. (2013).	Ho et al. (2019); Yuan et al. (2019); Comendeiro-Maaløe et al. (2019); Rossiter (2018); Pesko et al. (2018); ; Konetzka et al. (2018); Casalino et al. (2018); Li et al. (2018); Scott et al. (2017); Curry et al. (2013); Kralewski et al. (2012); Chukmaitov et al. (2009); Huckman (2005); Madison (2004).

A diverse range of variables/metrics may be used to approximate the levels of healthcare quality. The studies in the sample used several measures, of which stand out for their wide use, three: mortality, readmissions rate, and process of care measures. These three measures are, for that reason, analyzed independently. With that purpose, an auxiliary table that scrutinizes the quality of care analysis was elaborated (Table A.10).

Four studies came across better quality concerning mortality-related measures, one, worst, and five, unchanged. The worsening of mortality may not be the case. Nevertheless, no proper conclusions may be made regarding these slightly contradictory results. Something important to mention is that some out-of-sample studies suggest that mortality may have a tenuous relationship with quality of care (Shahian et al., 2012; Holloway and Quill, 2007), mainly due to the heterogeneity of in/outpatients in hospitals. Following Ferreira and Marques (2018) beliefs, researchers should integrate some severity index, like the case-mix index (CMI), when in the presence of heterogeneity of in/outpatients. Only three of the ten studies that studied mortality used indexes to adjust mortality. Henke et al. (2018) used a set of All Patient Refined Diagnosis Related Groups (APR-DRG) severity indicators, Ugolini and Nobilio (2003), the Charlson index, and Chukmaitov et al. (2009), the CMI. The first two came across decreased mortality and the other, unchanged. Despite small, this set of three papers that analyzed adjusted mortality may incline towards quality improvements regarding this measure.

Regarding readmission rates, five studies came across better quality, one, worst, and seven, unchanged. Just like with mortality, the fair distribution between positive and neutral studies limits conjectures.

Process of care quality measures may be related to surgical patients, inpatients or outpatients, and include continuation of beta-blockers for surgical patients, cholesterol control, breast cancer screening, adolescent immunization screening and, controlling high blood pressure, to name a few (Short and Ho, 2019; Carlin et al., 2015). These measures usually indicate how rapidly patients obtain treatment and how well hospitals deliver preventive services (Short and Ho, 2019). A total of nine studies analyzed these measures. Six studies came across better quality, and three, unchanged. Unlike the other two measures, this one may exhibit strong evidence of quality improvement. In fact, from all the measures, the process of care quality measures are the ones with more potential to benefit, at least directly, from better coordination of care (Short and Ho, 2019), a direct consequence of vertical integration. Better coordination of care may include enhancements regarding communication between providers and increments of access to resources, such as technology. Therefore, processes of care that require both providers and resources of different levels of care may exhibit improvements.

Other measures were studied, yet, not enough to reach substantial conclusions. Casalino et al. (2014) and Comendeiro-Maaløe et al. (2019), for example, analyzed preventable hospital admissions, reaching negative and non-significant changes, respectively. Liepert et al. (2014) analyzed duplicated computerized tomography (CT) scans, verifying a lower rate. Alternatively, Caballer-Tarazona and Vivas-Consuelo (2016) came across reduced delays in waiting lists.

Overall, the results spotted in Table 2.8 do not provide sufficient grounds for taking absolute conclusions regarding the impact of vertical integration on quality of care. Nevertheless, almost certainly, decrements of quality of care upon vertical integration are out of the question.

After the meticulous analysis of the three measures, one might state, nevertheless, that process of care measures seemed to be the ones that would have the most significant and positive differences. This independent analysis was essential to spot the measures where most variations exist. However, it is also important to refer that significant changes in a single measure are insufficient to draw conclusions about the overall quality of care, as increments in a certain quality measure may be achieved by decreasing other quality measures. That is why scrutinizing Table 2.8 is essential, even though a broad set of studies included only one quality measure, not analyzing, for that reason, a generalized effect.

2.4.5 Impact of vertical integration on access

A single study described the effects of vertical integration on access, as indicated in Table 2.9, reaching conclusions about a positive connection between the strategy and the outcome at hand.

Table 2.9. Empirical evidence on access.

Increases (n = 1)	Decreases (n = 0)	Unchanged (n = 0)
Haddad et al. (2020).	—	—

Haddad et al. (2020) developed a study on access to surgical care when in the presence of hospital-physician affiliations. By analyzing Medicaid beneficiaries, the researchers found evidence that endorses a substantial increase in access to surgical care for vulnerable, low-income patients over ten years. Access was examined by determining the association of practice Medicaid acceptance and vertical integration, using multivariable logistic regression. The study suggests the possibility of increases in both costs and prices of care following the rise in access.

One additional paper, Caballer-Tarazona and Vivas-Consuelo (2016), came across reduced delays in waiting lists, enabling conjectures about access, despite being considered a quality-related measure by the researchers.

Nevertheless, the lack of available articles on the access outcome somewhat limits the conjecture regarding access in vertically integrated providers.

2.5 Concluding remarks and directions for future work

Many researchers who focus their studies on vertical integration often recognize that this integration strategy mainly aims to provide financial benefits (Cuellar and Gertler, 2006; West et al., 2017). The present literature review calls into question those objectives associated with the implementation of this model. One encountered adverse effects of backward vertical integration towards the cost of care. These adverse effects were even more significant when analyzing patient-level costs. Additionally, vertical integration may also negatively influence prices of care, which may be a precursor of increments in patient-level expenses.

Further, the included studies show no absolute conclusions regarding both efficiency and quality of care. A negative influence of vertical integration on these outcomes is almost out of the question, yet, a positive effect is also not absolute.

Regarding access, very little can be said. Ferreira and Marques (2018)'s study may be again highlighted in the literature since it proposes a panoply of access-related variables relevant for future research on vertical integration or even in other areas related to healthcare.

As noted in the conceptual framework, the literature in the analysis focused on vertical integration within the delivery care system. From all the researches, only six examined the integration between secondary care and continuity of care. Wang et al. (2001), David et al. (2011), Liepert et al. (2014), Rahman et al. (2016), and Gupta et al. (2019) reached favorable evidence, while Konetzka et al. (2018)

showed inconsistent evidence. Despite the sample being quite unbalanced (only 6 out of 64 studies examine the integration between secondary care and continuity of care), the forward model appears to have overall positive effects on hospitals.

As previously stated, vertical integration did not appear routinely in the medical literature until the late 80s. However, since 2011, the number of studies per year regarding vertical integration in healthcare has increased, revealing a promising future on the study of this topic (Figure A.2). Nevertheless, it is important to emphasize again that the present literature review includes only quantitative research articles within the delivery care system. It only consists of a small sample of all the available articles regarding vertical integration in healthcare. However, it is enough to comprehend that health researchers are on the pathway to fully understand vertically integrated providers' performance.

One additional subject that should be addressed is the methodology used to analyze vertical integration, which is a serious issue and is inferred in Figure A.1. By observing the overview presented in that figure, the first thing that stands out is a large number of regression analyses compared to DEA or SFA, which are substantially more robust models, by enabling the analysis of a complete concept of efficiency, the flow of produced inputs and outputs (Wei and Wang, 2017). It is never too much to clarify that these methods are extremely significant and complete, allowing not only to examine efficiency but also indirectly, quality of care, access, or even other patient-oriented outcomes with relevance to policy-makers.

Perhaps one of the study's most significant limitations, a detail that may introduce some bias in the conclusions drawn from the study, is that only 11 out of the 64 included studies were carried outside of the US (information gathered from Figure A.4). By reviewing Tables 2.2 to 2.4 and resembling the screening process developed in this study, there was clear evidence of significant efforts from European and Asian countries to analyze healthcare providers' performance. These efforts are crucial to combat the astronomical difference between healthcare systems around the world. In fact, European healthcare managers should not rely entirely on the effects of vertical integration in US healthcare providers to decide whether to implement it in their countries. Thus, it is undeniable to state a clear need for studies in countries other than the US that are gradually introducing vertical integration into their health systems.

3

Case study

This chapter introduces the case study to determine whether hospitals within vertical models exhibit better performance levels than the remaining public Portuguese hospitals. The study presents and details the sample, the variables, the models, and the methodology used.

3.1 Data collection and methodological issues

This study aims at evaluating the performance of hospitals within LHUs and comparing them with traditional hospitals. Performance is computed using data relative to consumed resources and delivered services of healthcare. In view of that, key-performance indicators (KPIs) on quality and access, and environmental data were collected from the official database `https://benchmarking-acss.min-saude.pt/` and PORDATA (`https://www.pordata.pt/`), supported and developed by the (Portuguese) Central Administration of Health Systems and by the Francisco Manuel dos Santos Foundation, respectively.

3.1.1 Sample

The study's sample comprises all of the existing 8 LHUs and 31 traditional hospitals, composing a total of 39 Portuguese public hospitals. The latter group includes singular hospitals (SHs), hospital centres (HCs) and PPPs. The sample is distributed over five years, from January, FY2015, to December, FY2019. Therefore, the global sample includes 2,340 observations of hospitals, of which 480 correspond to LHUs and 1,860 to traditional models.

Pooling the data over time aims to considerably increase the sample size, improving, as a consequence, the resolution of the results. Nevertheless, it is essential to state that this strategy assumes no substantial changes in technology and hospitals' administration over the study period (Chowdhury and Zelenyuk, 2016), which can be refutable in the case of a broad range of years. The data range chosen in this research should not harm this assumption.

Although public hospitals in Portugal are composed of a mix of single hospitals, hospital centres, LHUs, PPPs, and oncology centres (Instituto Português de Oncologia, IPO), the last were excluded from the analysis. This exclusion was derived from the fact that specialized hospitals like those tipically have their technology of production, providing heterogeneity to the sample. An additional healthcare provider (Oeste Hospital Center, PPE) was disregarded from the analysis due to considerable missing data.

3.1.2 Desirable and undesirable process variables

The choice of process variables follows not only the data availability in the databases for all the sample but also a review of seven scientific articles that use both quality- and access-related variables to analyze

Portuguese healthcare providers. Table 3.1 exhibits the distinct variables observed in the literature review.

Table 3.1. Process variables used in the literature.

Cluster	Indicator	Direction	[1]	[2]	[3]	[4]	[5]	[6]	[7]
[Q] Care appropriateness	Outpatient surgeries on potential outpatient procedures	↗	✓	✓			✓	✓	✓
	Minor surgeries per potential minor surgeries	↗				✓			
	Rate of readmissions within 30 days after discharge	↙	✓	✓	✓	✓	✓	✓	✓
	Rate of inpatients staying more than 30 days	↙	✓	✓				✓	✓
[Q] Clinical safety	Decubitus ulcer rate	↙	✓	✓	✓	✓	✓	✓	✓
	Postoperative pulmonary embolism / deep vein thrombosis rate	↙	✓	✓	✓	✓	✓	✓	✓
	Postoperative septicaemia rate	↙	✓	✓	✓	✓	✓	✓	✓
	Obstetric trauma rate – vaginal delivery with instrument	↙	✓		✓		✓	✓	✓
	Obstetric trauma rate – vaginal delivery without instrument	↙	✓		✓		✓	✓	✓
	In-hospital death rate for low severity levels	↙	✓		✓		✓	✓	✓
	Catheter related bloodstream infections rate	↙		✓		✓			
	Rate of caesarean sections per delivery	↙						✓	✓
	Caesarean sections on TUCP	↙						✓	
	First caesarean sections on TUCP	↙						✓	
	Vaginal deliveries after caesarean sections on TUCP	↙						✓	
[A] Timeliness	Rate of first medical appointments within time	↗	✓	✓			✓	✓	✓
	Rate of surgeries within time	↗		✓	✓		✓	✓	
	Hip fracture surgery in the first 48h	↗	✓	✓	✓		✓	✓	✓
	Waiting time before surgery	↙	✓	✓			✓		
[A] Services availability	Hospital beds per 1,000 inhabitants	↗	✓						
	Hospital beds per 100 standard inpatients	↗		✓	✓				
	Full-time equivalent nurses per 1,000 standard patients	↗	✓		✓				
	Full-time equivalent doctors per 1,000 standard patients	↗	✓		✓				
	Inpatient bed occupancy rate	↗	✓						
	Operating theatre capacity utilization	↗	✓						
	Average delay in inpatient services (days per patient)	↙	✓						

[1] Ferreira and Marques (2018); [2] Ferreira and Marques (2020); [3] Ferreira and Marques (2017); [4] Ferreira et al. (2019a); [5] Ferreira et al. (2018); [6] Ferreira et al. (2020); [7] Ferreira et al. (2019b).

The lower, the better: ↙

The higher, the better: ↗

According to the literature review, both quality and access are genuinely complex concepts when applied to health systems. Following the discussion of Ferreira and Marques (2018), one might divide both quality and access into different categories. Additionally, scrutinizing conjectures explored by several researchers (Donabedian, 2005; Navarro-Espigares and Torres, 2011; Ferrier and Trivitt, 2013), quality may be divided into two categories: care appropriateness and clinical safety. In turn, access may present three dimensions: timeliness of services, services availability, and characteristics of the population at risk (Gulliford et al., 2002; Peters et al., 2008). In the current study, ”characteristics of the population at risk” were not considered an access-related dimension. Nevertheless, all of the variables usually included in that specific dimension were added to the data set of environmental variables.

The complexity inherent in these concepts makes its assessment extremely difficult. To handle this situation, one might use variables/metrics to approximate those concepts. In the sample of articles used for the literature review, quality is mostly approximated with the following indicators: ”rate of readmissions within 30 days after discharge” (readmissions rate may mirror the service quality provided in the first admission), ”decubitus ulcer rate”, ”ostoperative pulmonary embolism / deep vein thrombosis rate”, and ”postoperative septicaemia rate”. The previous rates directly link to how well hospitals succeed in preventing each one of the respective conditions.

Regarding access, the most widely used variables were: ”rate of first medical appointments within time” (there is a legislated maximum guaranteed time for first medical appointments in hospitals: Decree-Law no. 44/2017 of 20th April), ”rate of surgeries within time” (there are maximum defined times for surgeries: Decree-Law no. 153/2017 of 4th May) and ”hip fracture surgery in the first 48h” (hip fractures represent a meaningful cause of morbidity and mortality mostly in elderly patients (Zuckerman, 1996)).

As previously stated, process variables were chosen by scrutinizing studies that analyze the performance of Portuguese healthcare providers on quality and access. This step is crucial as one has to be sure to include in the analysis variables that indeed exist in Portuguese databases. Nevertheless, as it somewhat limits the diversity of analyzed papers, it is relevant to check whether variables in Table 3.1 are used to analyze different health systems. Within clinical safety, ”decubitus ulcer rate” may be highlighted for being broadly studied (Tsang et al., 2008; Lyder et al., 2012; Sullivan and Schoelles, 2013). Within care appropriateness, ”rate of readmissions within 30 days after discharge” stands out for being widely used to analyze Canadian healthcare providers (Allin et al., 2016; Chowdhury and Zelenyuk, 2016), as well as Danish (Dahl and Kongstad, 2017) and American (Khushalani and Ozcan, 2017). Finally, when considering the timeliness of access, ”hip fracture surgery in the first 48h” is also used to analyze the performance of NHS hospitals in England (Bottle and Aylin, 2006).

Ultimately, all the indicators that were used more than once in the analyzed literature (see Table 3.1) were included in the case study (Figure 3.1). In the figure, and henceforward, desirable variables are represented as g^+ and undesirable as g^-. Metrics like the ones used should approximate the performance on healthcare access and quality by providing a robust indicator of what can and should be taken into account in an effective healthcare plan (Fullman et al., 2018).

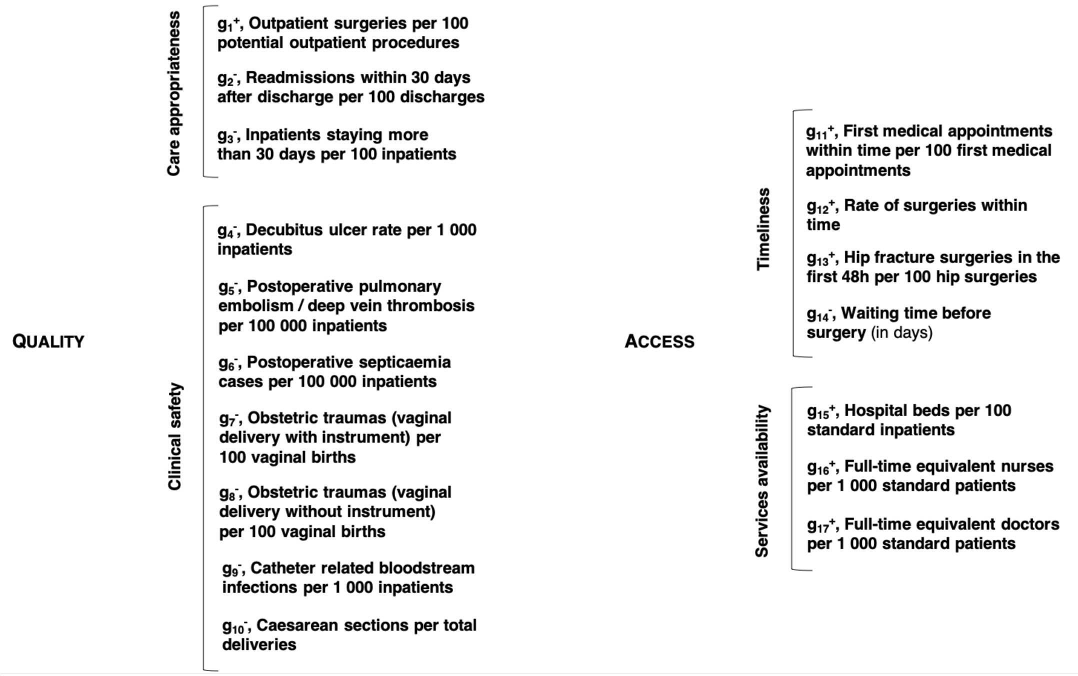


Figure 3.1. Process variables.

3.1.3 Environmental variables

As with process variables, the choice of environmental variables has followed criteria like data availability and frequency of usage in the literature. Table 3.2 identifies the variables adopted in the literature to characterize the environment of hospitals in Portugal.

It is well known that several demographic and epidemiological factors may influence hospital's activity. Indeed, the older population is prone to more severe diseases, thus influencing the quality of the provided services. Additionally, factors such as population density or even purchasing power per capita tend to affect access.

By observing the PORDATA database, one may verify that these variables are council-specific. Nevertheless, populations from more than one council are usually targeted by the same hospital, meaning that a weighted average of targeted councils' data must be used to aggregate information for the same hospital. Weights are the relative population of each one of the councils targeted by that hospital.

Again, all the indicators used more than once in the analyzed literature were included in the case study (Figure 3.2), excluding "stillbirth rate" due to data unavailability. These metrics should provide enough data to take into consideration the exogenous environment in the Portuguese health sector.

Table 3.2. Environmental variables used in the literature.

Indicator	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Population density		✓	✓	✓	✓	✓	✓
Population size					✓	✓	
Purchasing power per capita	✓	✓	✓	✓	✓	✓	✓
Illiteracy rate	✓	✓		✓	✓	✓	✓
Stillbirth rate	✓	✓		✓	✓	✓	✓
Childhood mortality rate	✓	✓		✓	✓	✓	✓
Elderly rate		✓	✓	✓	✓	✓	✓
Youth rate		✓			✓	✓	✓
Death rate		✓			✓	✓	✓
Elderly relative mortality rate		✓			✓		✓
Crude birth rate		✓			✓	✓	✓
Sec./Terc. education rate		✓			✓		✓
Dependence index		✓			✓	✓	✓
Inhabitants per doctor		✓			✓	✓	✓
Inhabitants per pharmacist		✓			✓	✓	✓

[1] Ferreira and Marques (2018); [2] Ferreira and Marques (2020); [3] Ferreira and Marques (2017); [4] Ferreira et al. (2019a); [5] Ferreira et al. (2018); [6] Ferreira et al. (2020); [7] Ferreira et al. (2019b).

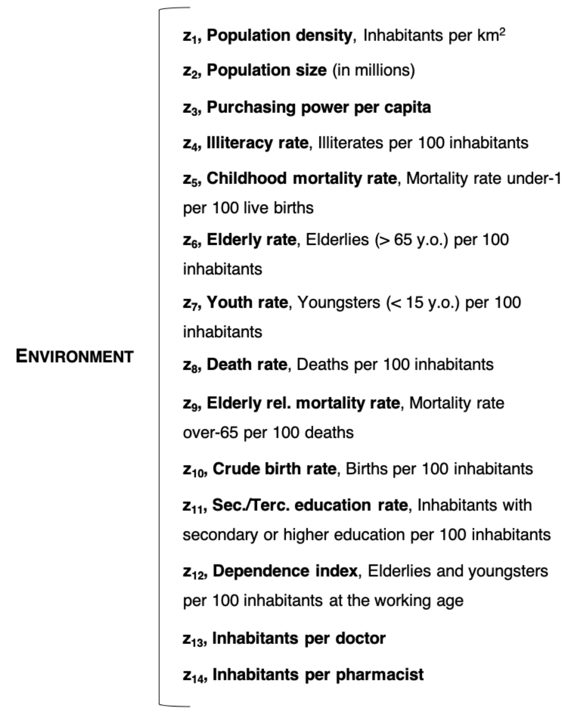


Figure 3.2. Environmental variables.

3.1.4 Models and scenarios

3.1.4.A Quality- and access related performance

To deal with gaps in the data set for certain years and variables, this study introduces three distinct models to simultaneously analyze quality- and access-related performance. In fact, three healthcare providers (Santa Maria Maior Hospital, PPE; Figueira da Foz District Hospital, PPE; Litoral Alentejano LHU, PPE) have substantial missing data for three specific variables (g_7^-, g_8^- and g_{10}^-). To prevent excluding the healthcare providers or the variables from the analysis and, at the same time, introducing bias in the results (due to substantial missing data), the first three models were designed as follows:

Model I. The three variables under consideration are excluded from the analysis.

Model II. The three healthcare providers under consideration are excluded from the analysis.

Model III. Both the healthcare providers and the variables are included in the analysis.

Table 3.3 summarizes the models.

Table 3.3. Overview of models I to III.

Model I		*Model II*		*Model III*	
Hospitals	Variables	Hospitals	Variables	Hospitals	Variables
39	14	36	17	39	17

3.1.4.B Outcome-specific performance

In addition to the overall case study, it is interesting to analyze each one of the outcomes separately and, besides that, each one of the outcomes' dimensions. For that reason, six additional models were adapted from Model I:

Model IV. Only quality-related variables are included in the analysis.

Model V. Only care appropriateness indicators are included in the analysis.

Model VI. Only clinical safety indicators are included in the analysis.

Model VII. Only access-related variables are included in the analysis.

Model VIII. Only timeliness indicators are included in the analysis.

Model IX. Only services availability indicators are included in the analysis.

Table 3.4. Overview of models IV to IX.

Model IV		*Model V*		*Model VI*	
Hospitals	Variables	Hospitals	Variables	Hospitals	Variables
39	7	39	3	39	4

Model VII		*Model VIII*		*Model IX*	
Hospitals	Variables	Hospitals	Variables	Hospitals	Variables
39	7	39	4	39	3

3.1.4.C Scenarios

As certain data values were still missing, although insubstantially in model I and II, an additional step was required. Within each model two scenarios were created:

Scenario I (optimistic). Missing data is replaced by optimistic values, i.e., maximum values for desirable process variables and minimum values for undesirable process variables.

Scenario II (pessimistic). Missing data is replaced by pessimistic values, i.e., minimum values for desirable process variables and maximum values for undesirable process variables.

In light of this, there is a total of 18 models to analyze.

3.2 Two clusters of healthcare providers: Basic statistics

This study introduces two clusters of healthcare providers to be analyzed: cluster A, LHUs, and cluster B, HCs, and SHs. To get an initial perception of each cluster's performance regarding each process variable, some basic statistics may be calculated. A student's t-test for means was elaborated, and the results are presented in Table 3.5, together with the mean, the standard deviation, the coefficient of variation, and the bounds regarding the overall sample and the clusters A and B for each one of the process variables.

In order to better analyze the variables in question, a non-parametric statistical test was additionally carried. A non-parametric test like Kruskal-Wallis test, which considers the medians of the distributions, was applied to deal with possible outliers. The results of this statistical test are also presented in Table 3.5.

The same process was carried out with the environmental variables to analyze potential significant differences between the two clusters. Table 3.6 exhibits the results.

Table 3.5. Process variables: Basic statistics.

	g_1^+	g_2^-	g_3^-	g_4^-	g_5^-	g_6^-	g_7^-	g_8^-	g_9^-	g_{10}^-	g_{11}^+	g_{12}^+	g_{13}^+	g_{14}^-	g_{15}^+	g_{16}^+	g_{17}^+
Overall sample																	
Mean	80.52	8.06	3.49	1.06	171.95	687.59	2.39	0.50	0.10	28.91	72.36	79.56	47.16	0.87	15.53	3.92	6.43
Std. deviation	8.50	1.84	1.21	2.31	258.03	842.82	3.87	0.99	0.37	6.35	13.59	12.84	23.91	0.38	4.43	1.18	1.58
CV(%)	10.56	22.83	34.67	217.93	150.06	122.58	161.93	198.00	370	21.96	18.78	16.14	50.70	43.68	28.53	30.10	24.57
min(g)	0.60	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5.30	0.00	3.30	0.00	0.07	0.00	0.10	0.20
max(g)	100.00	14.33	9.39	26.60	3278.70	8080.80	50.00	10.53	6.10	64.70	100.00	100.00	100.00	4.74	621.66	392.20	51.00
LHU, A																	
Mean	79.78	8.16	3.43	0.59	135.86	566.42	1.49	0.40	0.08	30.64	80.04	79.46	53.61	0.93	15.13	3.06	5.94
Std. deviation	9.38	1.90	1.19	1.48	306.68	886.91	4.02	1.27	0.34	8.52	10.77	12.84	22.43	0.26	3.85	1.18	2.00
CV(%)	11.76	23.28	34.69	250.85	225.73	156.58	269.80	317.50	425.00	27.81	13.46	16.16	41.84	27.96	25.45	38.56	33.67
min(g)	25.90	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5.30	48.00	61.50	0.00	0.21	0.00	0.50	0.80
max(g)	100.00	14.33	9.39	10.40	3278.70	5882.40	33.33	10.53	6.10	64.70	100.00	100.00	100.00	1.96	94.92	9.50	18.50
HC + SH, B																	
Mean	80.71	8.03	3.51	1.18	181.27	718.96	2.60	0.52	0.11	28.49	70.38	79.59	45.49	0.85	15.63	4.16	6.56
Std. deviation	8.25	1.82	1.22	2.47	243.11	828.38	3.81	0.91	0.38	5.62	13.54	11.70	24.00	0.40	4.56	1.08	1.42
CV(%)	10.22	22.67	34.76	209.32	134.12	115.22	146.54	175	345.46	19.73	19.24	14.70	52.76	47.06	29.17	25.96	21.65
min(g)	0.60	0.44	0.27	0.00	0.00	0.00	0.00	0.00	0.00	12.80	0.00	3.30	0.00	0.07	0.77	0.10	0.20
max(g)	100.00	13.58	8.55	26.60	1612.90	8080.80	50.00	5.88	3.80	63.10	99.90	100.00	100.00	4.74	621.66	392.20	51.00
Independent two-sample Student's t-test																	
p-value[a]	0.05	0.16	0.19	0.00	0.00	0.00	0.00	0.03	0.12	0.00	0.00	0.87	0.00	0.00	0.03	0.00	0.00
lower bound	0.01	-0.32	-0.04	0.42	19.57	68.10	0.68	0.01	-0.01	-3.00	-10.81	-1.45	-10.52	-0.11	0.05	0.99	0.43
upper bound	1.86	0.05	0.20	0.76	71.26	236.97	1.55	0.22	0.06	-1.28	-8.52	1.71	-5.73	-0.05	0.96	1.22	0.81
T statistic	1.99	-1.42	1.32	6.67	3.45	3.54	5.05	2.19	1.55	-4.90	-16.57	0.16	-6.65	-5.28	2.20	19.36	6.36
Best performer	∼	∼	∼	A	A	A	A	∼	∼	B	A	∼	A	B	∼	B	B
Kruskal-Wallis test																	
p-value[b]	0.54	0.09	0.02	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.35	0.00	0.00	0.02	0.00	0.00
Best performer	∼	∼	∼	A	A	A	A	A	A	B	A	∼	A	B	∼	B	B

[a] p-value regards the null hypothesis of equal means ($p \leq 0.01$ imply rejecting the null hypothesis).

[b] p-value regards the null hypothesis of similar distributions ($p \leq 0.01$ imply rejecting the null hypothesis).

∼ similar

Table 3.6. Environmental variables: Basic statistics.

	z_1	z_2	z_3	z_4	z_5	z_6	z_7	z_8	z_9	z_{10}	z_{11}	z_{12}	z_{13}	z_{14}
Overall sample														
Mean	549.44	0.65	93.77	6.06	3.18	21.22	13.99	10.67	89.17	7.91	29.00	33.19	398.03	1054.49
Std. deviation	705.20	1.03	14.69	2.90	1.57	4.66	1.73	2.93	2.63	1.61	6.24	9.05	253.19	230.00
CV(%)	128.35	158.46	15.67	47.85	49.37	21.96	12.37	27.46	2.95	20.35	21.52	27.27	63.61	21.81
min(z)	18.39	0.06	71.50	2.50	1.30	15.10	10.43	6.98	82.87	5.80	17.80	21.57	152.70	627.00
max(z)	2436.17	3.60	132.00	11.88	9.36	29.84	16.97	16.40	93.80	11.97	46.80	52.63	1213.82	1598.14
LHU, A														
Mean	270.89	0.16	83.74	9.36	3.68	25.66	12.68	13.54	90.08	7.44	24.31	41.87	691.79	1236.56
Std. deviation	634.25	0.078	12.25	2.92	2.52	4.61	1.45	2.73	2.48	1.26	2.53	9.61	350.84	178.84
CV(%)	234.14	48.75	14.63	31.20	68.48	17.97	11.44	20.16	2.75	16.94	10.41	22.95	50.71	14.46
min(z)	18.39	0.088	73.44	3.45	1.39	15.54	10.91	7.86	84.65	5.97	22.60	22.45	203.08	1041.90
max(z)	1836.18	0.32	110.47	11.88	9.36	29.84	15.22	15.97	92.19	9.68	30.42	52.63	1213.82	1598.14
HC + SH, B														
Mean	621.33	0.78	96.35	5.20	3.05	20.08	14.33	9.94	88.93	8.03	30.21	30.95	322.22	1007.51
Std. deviation	714.16	1.12	14.31	2.24	1.25	3.99	1.64	2.52	2.66	1.68	6.36	7.54	153.40	220.05
CV(%)	114.94	143.59	14.85	43.08	40.98	19.87	11.44	25.35	2.99	20.92	21.05	24.36	47.61	21.84
min(z)	18.40	0.06	71.50	2.50	1.30	15.10	10.43	6.98	82.87	5.80	17.80	21.57	152.70	627.00
max(z)	2436.17	3.60	132.00	11.00	7.20	29.27	16.97	16.40	93.80	11.97	46.80	48.90	797.62	1516.60
Independent two-sample Student's t-test														
p-value	0.22	0.01	0.03	0.00	0.32	0.00	0.01	0.00	0.28	0.36	0.00	0.00	0.02	0.01
lower bound	-211.80	0.20	1.41	-6.07	-1.89	-8.88	0.36	-5.66	-3.26	-0.70	1.21	-17.33	-664.56	-400.09
upper bound	912.68	1.03	23.81	-2.24	0.63	-2.27	2.95	-1.54	0.96	1.89	10.59	-4.52	-74.59	-58.01
T statistic	1.26	3.02	2.28	-4.40	-0.69	-3.42	2.59	-3.55	-1.10	0.93	2.55	-3.46	-2.91	-2.71
Decision	∼	DIF	∼	DIF	∼	DIF	DIF	DIF	∼	∼	DIF	DIF	∼	DIF
Kruskal-Wallis test														
p-value	0.01	0.05	0.01	0.00	0.75	0.01	0.02	0.01	0.144	0.365	0.00	0.01	0.00	0.01
Decision	DIF	∼	DIF	DIF	∼	DIF	∼	DIF	∼	∼	DIF	DIF	DIF	DIF

[a] p-value regards the null hypothesis of equal means ($p \leq 0.01$ imply rejecting the null hypothesis).

[b] p-value regards the null hypothesis of similar distributions ($p \leq 0.01$ imply rejecting the null hypothesis).

∼ similar

DIF different

The independent two-sample Student's t-test employed to compare the performance of the two clusters, achieved contradictory findings. In fact, by observing Table 3.5, it becomes evident that:

- The two clusters have similar performance, when measuring outpatient surgeries on potential outpatient procedures (g_1^+), rate of readmissions within 30 days after discharge (g_2^-), rate of inpatients staying more than 30 days (g_3^-), obstetric trauma rate – vaginal delivery without instrument (g_8^-), catheter related bloodstream infections per 1,000 inpatients (g_9^-), rate of surgeries within time (g_{12}^+) and hospital beds per 100 standard inpatients (g_{15}^+).
- Hospitals within LHUs exhibit better performance when measuring decubitus ulcer rate (g_4^-), postoperative pulmonary embolism / deep vein thrombosis rate (g_5^-), postoperative septicaemia rate (g_6^-), obstetric trauma rate – vaginal delivery with instrument (g_7^-), rate of first medical appointments within time (g_{11}^+) and hip fracture surgery in the first 48h (g_{13}^+).
- The non-LHUs cluster exhibits better performance when measuring rate of caesarean sections per delivery (g_{10}^-), waiting time before surgery (g_{14}^-), full-time equivalent nurses per 1,000 standard patients (g_{16}^+), and full-time equivalent doctors per 1,000 standard patients (g_{17}^+).

By confronting the results of the independent two-sample Student's t-test with the ones from the Kruskal-Willis test, one may verify different conclusions regarding obstetric trauma rate – vaginal delivery without instrument (g_8^-) and catheter related bloodstream infections per 1,000 inpatients (g_9^-). The parametric test suggests similar performance between the two clusters regarding those variables, while the non-parametric test suggests the LHU cluster as the best performer.

The divergent results within and between the statistical tests prove the necessity of a more complex benchmarking model to optimise the weights associated with each quality and access variable to obtain a composite index that allows unveiling which cluster is the best performer.

Ultimately, the results from Table 3.6 reveal significant environmental differences between the two clusters. These circumstances, which were expected, add one other premise to the benchmarking model: the effect of the environment must be taken into account.

3.3 Methodology

3.3.1 Pre-processing

3.3.1.A Outliers deletion

Before replacing the models' missing values according to the scenarios, an initial search for outliers was conducted. All the values outside of the range $[\mu_i - 2\sigma_i, \mu_i + 2\sigma_i]$, which correspond to observations of variable *i* within the first two standard deviations, were analyzed. According to the empirical rule, 95% of the observations fall within the range used, which showed to be the best one to analyze the outliers. In fact, analyzing outliers outside the range $[\mu_i - 3\sigma_i, \mu_i + 3\sigma_i]$ appeared to be over-exclusive for outliers.

After being signaled, the values that looked like a typo or appeared illogical were eliminated. Subsequently, missing values were replaced according to the scenarios.

3.3.1.B Principal Component Analysis

There are several pitfalls that researchers may confront that could have a significant influence on the application of the data envelopment analysis (DEA), which is described in Subchapter 3.3.3. First, when using a high number of output variables, some variables may be correlated, increasing the database's complexity unnecessarily. Second, and also when using a high number of output variables, the number of decision-making units (DMUs) (concept further explored in Subchapter 3.3.2) may be occasionally close to the number of inputs and outputs, which could yield a large number of efficient units (Amirteimoori et al., 2014). This last drawback is further worsened in the present study, as by including the environment effect in the analysis, only similar DMUs are used to make up the frontier for a certain DMU, meaning the number of DMUs used in the model may decrease dramatically, getting frequently closer to the number of output variables.

Considering the aforementioned reasonings, it is important to verify (for each model used) if it is possible to reduce the dimensionality of the problem. This is where the principal component analysis (PCA) comes into the scene. PCA is an orthogonal linear transformation that converts the data to a new coordinate system. By causing a change of basis on the data, some principal components may be disregarded, thus reducing the dimensionality of the problem (Adler and Golany, 2007).

As for every upside there is a downside, PCA has two problems that have to be discussed. First, when the number of principal components is lower than the number of original variables, a part of the original information is lost as a variation factor (Ueda and Hoshiai, 1997). In this analysis, the components included were the components that together represented at least 85% of the data's variance. A number between 80-90% is frequently used, despite the information loss inherent to the process (Adler and Golany, 2007). Second, in this method, many weights for the variables that define principal components take negative values. This situation has to be dealt with, as most DEA theorems assume that data values are positive or at least semi-positive, and the model used is not an exception. One way to achieve this is by adding a sufficiently large positive constant to the input or output values with the non-positive number (Pastor and Ruiz, 2007). This procedure was applied to each one of the principal components that presented negative values.

3.3.2 Data Envelopment Analysis

Data Envelopment Analysis (DEA) is a non-parametric “data-oriented” approach for evaluating the performance of a group of entities named decision-making units (DMUs), which theoretically ”transform” multiple inputs into multiple outputs (Thanassoulis, 2001; Cooper et al., 2011).

Precisely, DEA directs its analysis towards frontiers rather than central tendencies. The latter is common in several methodologies, as, for example, regression models. Moreover, the present methodology stands out for uncovering relationships that would remain hidden with other methods (Cooper et al., 2011). This is a consequence of the fact that one does not require to explicitly formulate assumptions of weights or specify formal relationships between inputs and outputs, which makes DEA the most suitable methodology for the present research work.

One may recall from Chapter 2 that, in addition to DEA, Stochastic-Frontier Analysis (SFA) is widely used in the literature. Nevertheless, although it also analyzes frontiers, it requires strong functional assumptions, and it does not allow multiple outputs to be analyzed simultaneously (Jacobs et al., 2006), which makes its use in the present study impractical. However, despite these DEA’s advantages compared to SFA, it is essential to recognize that DEA also has disadvantages. Disadvantages as not enabling the distinction between efficiency variation noise, being vulnerable to outliers, and presenting endogeneity problems (Jacobs et al., 2006).

One of the most important considerations when applying a DEA model is whether to assume constant or variable returns to scale. The constant returns to scale (CRS) assumption was proposed by Charnes et al. (1978) in the original DEA paper. This framework is suitable when all units function at the optimal scale, which is difficult to assume in healthcare, for various reasons: imperfect competition, limitations on finance, mergers, to name a few (Jacobs et al., 2006). Banker et al. (1984) extended the model to be appropriate for a sub-optimal scale, which encompasses the creation of the variable returns to scale (VRS) model.

Even though one intends to analyze healthcare DMUs, the choice of CRS or VRS commonly hangs on the circumstances and the motivation of the study. One may recall that every single process variable included in the present case study incorporates a ratio. This kind of data implies the usage of the CRS model since any information regarding DMUs’ proportions is neglected in the construction of a ratio (Jacobs et al., 2006).

Taking into account the previous considerations, the formulation of a DEA CRS model may be illustrated. A DMU may be denoted by k, which is characterised by a set of m inputs, $x_k = \{x_{1k}, ..., x_{ik}, ..., x_{mk}\}$ and s outputs, $y_k = \{y_{1k}, ..., y_{rk}, ..., y_{sk}\}$. A DMU may be classified as efficient, exhibiting a score of 1 (100%), if and only if no other DMU present inputs or outputs that can be improved without worsening some of its other inputs or outputs (Cooper et al., 2011). The present concept is referred to as relative efficiency. This kind of efficiency, which may also be designated as technical efficiency (TE), is formulated as presented in Equation 3.1, as in Huguenin (2012).

$$TE_k = \frac{\sum_{r=1}^{s} u_r y_{rk}}{\sum_{i=1}^{m} v_i x_{ik}}, \tag{3.1}$$

being u_r and v_i, the weights of output r and input i, respectively.

The technical efficiency of a unit k is maximized under certain constraints, requiring, for that reason, the usage of linear programming. The problem can be considered by following two different approaches: input-oriented or output-oriented. This case study follows the latter approach, in which the weighted sums of outputs are maximized, holding inputs constant (Jacobs et al., 2006). For an output-oriented model, a frontier is identified based on the DMUs achieving the highest output mix given their inputs. The efficient DMUs form a piecewise linear envelope of surfaces in multidimensional space. Then, each DMU is assigned an efficiency score by comparing its output/input ratio to that of efficient DMUs (Jacobs et al., 2006).

Using linear programming notation, one is facing the problem of optimizing a linear objective function subject to a set of constraints. The dual equations (or equations in the multiplier form) for an output-oriented DEA CRS model are presented in Equation 3.2, as in Huguenin (2012).

$$\begin{aligned}
&\text{Minimize} \quad \sum_{i=1}^{m} v_i x_{ik} \\
&\text{Subject to} \\
&\sum_{i=1}^{m} v_i x_{ij} - \sum_{r=1}^{s} u_r y_{rj} \geq 0 \qquad j = 1, ..., n \\
&\sum_{r=1}^{s} u_r y_{rk} = 1 \\
&u_r, v_i > 0 \qquad \forall r = 1, ..., s; i = 1, ..., m
\end{aligned} \tag{3.2}$$

The model equations may be written in one other form, the envelopment form. This form is often recommended to solve the computation as it only involves $s + m$ constraints rather than $n + 1$ constraints in the multiplier form. The equations in the envelopment form (or primal equations) are presented in Equation 3.3, as in Huguenin (2012).

$$\begin{aligned}
&\text{Maximize} \quad \phi_k \\
&\text{Subject to} \\
&\phi_k y_{rk} - \sum_{j=1}^{n} \lambda_j y_{rj} \leq 0 \qquad r = 1, ..., s \\
&x_{ik} - \sum_{j=1}^{n} \lambda_j x_{ij} \geq 0 \qquad i = 1, ..., m \\
&\lambda_j \geq 0 \qquad \forall j = 1, ..., n
\end{aligned} \tag{3.3}$$

being $\frac{1}{\phi_k}$, the technical efficiency of unit k and λ_j, the associated weighting of outputs and inputs of unit j.

Ultimately, it is crucial to refer that this model was fitted in a Benefit of Doubt (BoD) framework. The basic BoD model is on a par with the original DEA CRS model of Charnes et al. (1978), with all KPIs considered as outputs and a dummy input equal to one for all the units. Formally, the routine way to neutralize the impact of inputs in the model is to set $m = 1$ and $x_i = 1$ for all j observations (Van Puyenbroeck, 2017). By adapting the previously defined DEA model to this framework, one constructs a version that exclusively focuses on outputs.

3.3.3 Consideration of the environmental variables

In order to deal with the environmental effect, conditional formulations were employed. In a general sense, in the DEA model, only similar DMUs were used to make up the frontier for a certain DMU. That similarity is analyzed by measuring the global bandwidth, using the dataset of environmental variables, and applying the product kernel approach.

The global bandwidth may be measured using the Silverman's bandwidth, $h_{silverman}$, which is written in Equation 3.4:

$$h_{silverman} = 1.06 \cdot \min\{\sigma, \frac{R_{23}}{1.34}\} \cdot n^{-1/5}, \tag{3.4}$$

being σ, R_{23} and n, the standard deviation, the interquartile range and the sample's size, respectively (Ferreira, 2016).

Following, one has to use the product kernel approach. Let $Z \in \mathbb{R}^s{}_+$ the $n \times s$ matrix of s exogenous variables. The first step is to find a kernel function $K : \mathbb{R}^s{}_+ \rightarrow \mathbb{R}$ that receives the $n \times s$ matrix and returns a $n \times 1$ vector with probabilities (Ferreira, 2016). Using the bandwidth and the kernel function (gaussian), Equation 3.5, the product of the s individual kernel functions may be computed for each DMU. The vector that results from that computation signals similar DMUs with values greater than zero and dissimilar DMUs with values equal to zero.

$$K(x, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2 / 2\sigma^2} \tag{3.5}$$

3.3.4 Malmquist Indices

In parallel to DEA, Malmquist Indices (MI) are used for comparing clusters of DMUs. Caves et al. (1982) developed the MI, although the construction of input quantity indices as ratios of distance functions was introduced much earlier by the Professor Sten Malmquist (Jacobs et al., 2006).

First, it is essential to understand the concept of distance function. We may begin by introducing the concept of technology of production Φ^t, which represents the set of all output vectors, $Y^t \in \mathbb{R}^s{}_+$, which can be produced using the input vectors, $X^t \in \mathbb{R}^m{}_+$ (Camanho and Dyson, 2006). Both the input distance function $D_i(X^t, Y^t)$ or the output distance function $D_o(X^t, Y^t)$ may be used and are presented in Equations 3.6 and 3.7, respectively.

$$D_i(X^t, Y^t) = \max\{\lambda : (\frac{X^t}{\lambda}, Y^t) \in \Phi^t\} \tag{3.6}$$

$$D_o(X^t, Y^t) = \min\{\theta : (\frac{Y^t}{\theta}, X^t) \in \Phi^t\} \tag{3.7}$$

In 1978, researchers showed that the distance function is the reciprocal to Farrell's measure of technical efficiency (Färe and Lovell, 1978). This opened the possibility of using DEA models to compute the MI (Camanho and Dyson, 2006). As one is using a DEA model to solve these distance functions, one may consider the present methodology a non-parametric approach.

Then, in 1994, Färe and his colleagues defined an input-oriented productivity index as the geometric mean of two MIs, one concerning the technology of production at time t and the other at time $t+1$ (Färe et al., 1994). This results in the index $M^{t,t+1}$ (Equation 3.8), a Malmquist-type measure of productivity, as in Camanho and Dyson (2006).

$$M^{t,t+1} = \left[\frac{D^t(X^{t+1}, Y^{t+1})}{D^t(X^t, Y^t)} \cdot \frac{D^{t+1}(X^{t+1}, Y^{t+1})}{D^{t+1}(X^t, Y^t)}\right]^{1/2} \tag{3.8}$$

Following the rationale behind the previous Malmquist-type index, an overall measure for the comparison of performance between two groups of DMUs (group A and B) may be defined (Equation 3.9), as in Camanho and Dyson (2006).

$$I^{AB} = \left[\frac{\left(\prod_{j=1}^{\delta_A} D^A(X_j{}^A, Y_j{}^A)\right)^{1/\delta_A}}{\left(\prod_{j=1}^{\delta_B} D^A(X_j{}^B, Y_j{}^B)\right)^{1/\delta_B}} \cdot \frac{\left(\prod_{j=1}^{\delta_A} D^B(X_j{}^A, Y_j{}^A)\right)^{1/\delta_A}}{\left(\prod_{j=1}^{\delta_B} D^B(X_j{}^B, Y_j{}^B)\right)^{1/\delta_B}}\right]^{1/2}, \tag{3.9}$$

being δ_A and δ_B, the number of DMUs of the groups, $X^A \in \mathbb{R}^m{}_+$ and $X^B \in \mathbb{R}^m{}_+$, the inputs of the groups, and $Y^A \in \mathbb{R}^s{}_+$ and $Y^B \in \mathbb{R}^s{}_+$, the outputs of the groups. $D^B(X_j{}^A, Y_j{}^A)$ may be read as: the input distance function for a DMU in group A with respect to the frontier of group B.

Using the index I^{AB}, one may proceed with a cross-sectional comparison of the performance of two clusters of DMUs operating in different conditions, at a particular moment in time, instead of a basic measure of productivity change between two time periods ($M^{t,t+1}$) (Camanho and Dyson, 2006). More specifically, I^{AB} evaluates the distance of the DMUs to a single reference technology. By observing the formula, it is easily confirmed that the first part of the expression estimates the ratio between the average distance of DMUs from group A to the group A frontier and the average distance of DMUs from group B to the group A frontier. The second portion of the expression expresses the same ratio but using the group B frontier. Therefore, I^{AB} is fundamentally the geometric mean of those two ratios (Camanho and

Dyson, 2006).

The index I^{AB}, the overall performance measure, may be decomposed into two sub-components (which is only possible because of the usage of the geometric mean formula), as in Camanho and Dyson (2006). The decomposition expression is presented in Equation 3.10, and the expressions for the two sub-components, IE^{AB}, and IF^{AB}, are presented in Equations 3.11 and 3.12, respectively.

$$I^{AB} = IE^{AB} \cdot IF^{AB} \tag{3.10}$$

$$IE^{AB} = \frac{\left[\prod_{j=1}^{\delta_A} D^A(X_j^A, Y_j^A)\right]^{1/\delta_A}}{\left[\prod_{j=1}^{\delta_B} D^B(X_j^B, Y_j^B)\right]^{1/\delta_B}} \tag{3.11}$$

$$IF^{AB} = \left[\frac{\left(\prod_{j=1}^{\delta_A} D^B(X_j^A, Y_j^A)\right)^{1/\delta_A}}{\left(\prod_{j=1}^{\delta_A} D^A(X_j^A, Y_j^A)\right)^{1/\delta_A}} \cdot \frac{\left(\prod_{j=1}^{\delta_B} D^B(X_j^B, Y_j^B)\right)^{1/\delta_B}}{\left(\prod_{j=1}^{\delta_B} D^A(X_j^B, Y_j^B)\right)^{1/\delta_B}}\right]^{1/2} \tag{3.12}$$

This decomposition allows for different comparisons. IE^{AB} is used to compare the within-group efficiency spreads, and IF^{AB} expresses the productivity gap between the frontiers of the two groups. This means that a good overall performance may be connected with two elements: less dispersion in the efficiency levels of the DMUs in one group compared to the other (IE^{AB}), or the dominance of the best practice frontier (IF^{AB}) (Camanho and Dyson, 2006).

Ultimately, the interpretation of the three presented measures may be illustrated. It is crucial to recall that the present study uses an output-oriented approach, which causes the interpretations to be precisely the opposite as those suggested by Camanho and Dyson (2006) since the researchers focus on an input-oriented approach. Therefore, the interpretations of an output-oriented approach are the following:

- $I^{AB} > 1$: reveals better performance in group A than in group B;
- $IE^{AB} > 1$: reveals that the efficiency spread is smaller (i.e., there is greater consistency in efficiency levels) in DMUs of group A than in those of group B;
- $IF^{AB} > 1$: reveals greater productivity of the frontier of group A compared to group B.

3.4 Overview of concepts in analysis

As previously mentioned, one implements two distinct models to carry out a performance analysis. In this context, it is crucial to highlight the inherent concepts to each model's outputs. Table 3.7 recalls and standardizes the concepts defined in the preceding subchapters. This step is also crucial to clarify that the present study intends to study the concept of performance regarding quality and access and not necessarily the concept of efficiency, which is usually a broader concept.

Table 3.7. Overview of concepts in analysis.

Performance Analysis				
Models	DEA model	Non-parametric MI approach		
Outputs	Efficiency scores	I^{AB}	IE^{AB}	IF^{AB}
Concepts in analysis	Partial performance	Global performance	Partial performance spread	Frontier-shift related performance

4

Results and discussions

In this chapter, the results from applying the DEA CRS model and the MI approach are both presented and discussed. It is important to refer that both the DEA scores and the MI scores presented for each model consist of the arithmetic mean of both scenarios (optimistic and pessimistic).

4.1 Data Envelopment Analysis: Cluster scores

Tables 4.1 to 4.3 provide some basic statistics concerning the DEA scores for the sample described in Chapter 3. Statistics like the mean, the standard deviation, the coefficient of variation, and the bounds regarding the overall sample and the clusters A and B are presented for each model.

Using the parametric two-sample Student's t-test and the nonparametric Kruskal-Wallis test, one may analyze the statistical significance of the results provided by the DEA CRS output-oriented algorithm. Tables 4.1 to 4.3 were complemented with the p-value concerning the null hypothesis (H_0: clusters A and B exhibit similar partial performance) and some other statistical results from the two tests. This study considers a significance level of 1% (or 0.01), meaning that if the p-values are above that value, there is no statistical evidence supporting the null hypothesis's rejection.

As mentioned earlier, an efficient DMU has an efficiency score of 1, meaning the closer the score of a certain DMU gets to 1, the more efficient it is. By having this in mind, when analyzing the mean efficiency score, in addition to the significance value, the best performer for each model may be discovered and displayed in the respective table.

To simultaneously analyse quality-related and access-related variables, one designed three distinct models (models I to III). Recalling Chapter 3, all of the three models include both quality-related and access-related variables. The differences in design were only to deal with gaps in the data set for certain years and variables, as previously explained. Additionally, quality-related efficiency is analyzed using a subset of seven variables. To analyze quality and the two quality dimensions independently, one designed three distinct models (models IV to VI). Finally, access is also analyzed using a subset of seven variables. To analyze access and the two access dimensions independently, one designed three distinct models (models VII to IX), as previously mentioned.

4.1.1 Quality- and access-related cluster scores

Table 4.1 describes all the statistics behind the concept of quality- and access-related partial performance.

Table 4.1. Quality- and access-related cluster scores: Statistics.

	Model I ϕ	Model II ϕ	Model III ϕ
Overall sample			
Mean	1.029	1.012	1.018
Std. deviation	0.043	0.019	0.028
CV(%)	4.179	1.877	2.750
min(ϕ)	1.000	1.000	1.000
max(ϕ)	1.360	1.123	1.329
LHU, A			
Mean	1.019	1.010	1.013
Std. deviation	0.029	0.017	0.021
CV(%)	2.846	1.683	2.073
min(ϕ)	1.000	1.000	1.000
max(ϕ)	1.159	1.105	1.120
HC + SH, B			
Mean	1.031	1.013	1.019
Std. deviation	0.046	0.019	0.029
CV(%)	4.462	1.876	2.846
min(ϕ)	1.000	1.000	1.000
max(ϕ)	1.360	1.123	1.329
Independent two-sample Student's t-test			
p-value	0.000	0.004	0.000
lower bound	0.009	0.001	0.004
upper bound	0.015	0.005	0.008
T statistic	7.005	2.901	4.982
Best performer	A	A	A
Kruskal-Wallis test			
p-value	0.000	0.004	0.000
Best performer	A	A	A

Consider models I to III. On average, hospitals within LHUs exhibit slightly higher partial performance than the remaining hospitals. This inference is reinforced by the two statistical tests performed. In fact, both tests provide results supporting that the difference between clusters of healthcare providers is significant at the 0.01 level.

Additionally, one may observe that the three models provided notably different DEA scores. This confirms the importance of designing these three different models to deal with gaps in the data set, even though, in this particular case, the best performer would still be cluster A.

4.1.2 Quality-related cluster scores

Table 4.2 describes all the statistics behind the concept of quality-related partial performance.

Table 4.2. Quality-related cluster scores: Statistics.

	Model IV ϕ	Model V ϕ	Model VI ϕ
Overall sample			
Mean	1.021	1.130	1.004
Std. deviation	0.025	0.110	0.014
CV(%)	2.449	2.212	1.394
min(ϕ)	1.000	1.000	1.000
max(ϕ)	1.328	1.880	1.299
LHU, A			
Mean	1.019	1.101	1.001
Std. deviation	0.025	0.081	0.004
CV(%)	2.453	7.357	0.400
min(ϕ)	1.000	1.000	1.000
max(ϕ)	1.328	1.617	1.033
HC + SH, B			
Mean	1.021	1.137	1.005
Std. deviation	0.025	0.115	0.015
CV(%)	2.449	10.114	1.493
min(ϕ)	1.000	1.000	1.000
max(ϕ)	1.265	1.880	1.299
Independent two-sample Student's t-test			
p-value	0.064	0.000	0.000
lower bound	0.000	-0.045	-0.004
upper bound	0.000	-0.027	-0.003
T statistic	1.869	-7.855	-8.747
Best performer	similar	A	A
Kruskal-Wallis test			
p-value	0.025	0.000	0.000
Best performer	similar	A	A

Consider model IV. On average, hospitals within LHUs exhibit slightly higher partial performance than the remaining sample. However, the differences are not meaningful in the statistical sense because the 99% confidence intervals (associated with the two clusters of healthcare providers) overlap within a decent range of efficiencies.

Through models V and VI, results are again favourable for hospitals within LHUs and significant at the 0.01 level. Therefore, one might say that hospitals within LHUs are slightly better performers than traditional models, concerning both care appropriateness (model V) and clinical safety (model VI).

4.1.3 Access-related cluster scores

Table 4.3 describes all the statistics behind the concept of access-related partial performance.

Table 4.3. Access-related cluster scores: Statistics.

	Model VII ϕ	Model VIII ϕ	Model IX ϕ
Overall sample			
Mean	1.115	1.116	1.436
Std. deviation	0.110	0.118	0.356
CV(%)	9.865	10.574	24.791
min(ϕ)	1.000	1.000	1.000
max(ϕ)	1.712	1.953	3.116
LHU, A			
Mean	1.076	1.085	1.382
Std. deviation	0.070	0.083	0.318
CV(%)	6.506	7.650	23.010
min(ϕ)	1.000	1.000	1.000
max(ϕ)	1.429	1.488	2.658
HC + SH, B			
Mean	1.125	1.123	1.450
Std. deviation	0.116	0.124	0.365
CV(%)	10.311	11.042	25.172
min(ϕ)	1.000	1.000	1.000
max(ϕ)	1.712	1.953	3.116
Independent two-sample Student's t-test			
p-value	0.000	0.000	0.000
lower bound	-0.060	-0.048	-0.101
upper bound	-0.041	-0.030	-0.035
T statistic	-11.751	-8.221	-4.055
Best performer	A	A	A
Kruskal-Wallis test			
p-value	0.000	0.000	0.000
Best performer	A	A	A

Once more, statistical evidence allows rejecting the null hypothesis for all the three models, meaning the clusters of hospitals exhibit different distances to the respective efficiency frontier. DEA scores of cluster A are significantly closer to 1, at the 0.01 level, compared with cluster B, for all the three models. Therefore, cluster A exhibits slightly higher partial performance than cluster B, concerning overall access (model VII), but also timeliness (model VIII) and services availability (model IX).

4.2 Data Envelopment Analysis: Discriminated scores

Tables 4.4 to 4.6 provide the DEA scores for each one of the 39 healthcare providers in study.

4.2.1 Quality- and access-related discriminated scores

Table 4.4 presents the discriminated DEA scores for the 39 healthcare providers for models I to III.

Table 4.4. Quality- and access-related discriminated scores.

Healthcare provider	Model I	Model II	Model III	Average	Ranking
Barreiro/Montijo HC, PPE	1.022	1.008	1.016	1.015	16
Leiria HC, PPE	1.024	1.010	1.009	1.014	15
Lisboa Ocidental HC, PPE	1.040	1.017	1.032	1.030	35
Setúbal HC, PPE	1.019	1.008	1.013	1.013	14
Baixo Vouga HC, PPE	1.036	1.008	1.016	1.020	27
Médio Ave HC, PPE	1.023	1.012	1.015	1.017	19
Coimbra HUC, PPE	1.015	1.009	1.006	1.010	9
Entre Douro e Vouga HC, PPE	1.023	1.012	1.012	1.016	17
Médio Tejo HC, PPE	1.025	1.012	1.013	1.016	18
Póvoa de Varzim/Vila do Conde HC, PPE	1.006	1.003	1.005	1.005	1
Tâmega e Sousa HC, PPE	1.033	1.018	1.020	1.024	31
Tondela-Viseu HC, PPE	1.027	1.012	1.012	1.017	20
Trás-os-Montes e Alto Douro HC, PPE	1.036	1.007	1.015	1.020	25
Cova da Beira HUC, PPE	1.032	1.006	1.018	1.019	23
Lisboa Central HUC, PPE	1.038	1.022	1.024	1.028	34
São João HUC, PPE	1.096	1.030	1.071	1.066	38
Algarve HUC, PPE	1.014	1.005	1.009	1.009	7
Porto HUC, PPE	1.050	1.023	1.031	1.035	36
Lisboa Norte HUC, PPE	1.131	1.032	1.051	1.071	39
Vila Nova de Gaia/Espinho HUC, PPE	1.014	1.012	1.010	1.012	12
Senhora da Oliveira Hospital, PPE	1.064	1.017	1.047	1.043	37
Braga Hospital, PPE	1.013	1.004	1.006	1.008	3
Cascais Hospital, PPP	1.012	1.009	1.015	1.012	13
Loures Hospital, PPP	1.014	1.011	1.011	1.012	11
Vila Franca de Xira Hospital, PPP	1.041	1.010	1.028	1.026	33
Figueira da Foz District Hospital, PPE	1.026	NA	1.009	1.018	21
Santarém District Hospital, PPE	1.026	1.014	1.017	1.019	24
Espírito Santo de Évora Hospital, PPE	1.019	1.018	1.025	1.021	29
Fernando Fonseca Hospital, PPE	1.012	1.016	1.008	1.012	10
Garcia de Orta Hospital, PPE	1.010	1.010	1.007	1.009	5
Santa Maria Maior Hospital, PPE	1.030	NA	1.010	1.020	26
Guarda LHU, PPE	1.010	1.004	1.007	1.007	2
Castelo Branco LHU, PPE	1.014	1.005	1.008	1.009	6
Matosinhos LHU, PPE	1.021	1.023	1.011	1.018	22
Alto Minho LHU, PPE	1.030	1.012	1.020	1.021	28
Baixo Alentejo LHU, PPE	1.011	1.012	1.006	1.010	8
Litoral Alentejano LHU, PPE	1.031	NA	1.016	1.024	32
Nordeste LHU, PPE	1.013	1.003	1.010	1.009	4
Norte Alentejano LHU, PPE	1.025	1.013	1.025	1.021	30

By observing Table 4.4, one may verify that the Póvoa do Varzim/Vila do Conde Hospital Center appears to be the best performer when analyzing both quality- and access-related variables. This healthcare provider is followed by Guarda LHU and Braga Hospital.

4.2.2 Quality-related discriminated scores

Table 4.5 presents the discriminated DEA scores for the 39 healthcare providers for models IV to VI.

Table 4.5. Quality-related discriminated scores.

Healthcare provider	Model IV	Model V	Model VI	Average	Ranking
Barreiro/Montijo HC, PPE	1.015	1.120	1.004	1.046	25
Leiria HC, PPE	1.013	1.117	1.000	1.043	22
Lisboa Ocidental HC, PPE	1.018	1.187	1.004	1.070	30
Setúbal HC, PPE	1.026	1.106	1.011	1.048	26
Baixo Vouga HC, PPE	1.019	1.100	1.004	1.041	17
Médio Ave HC, PPE	1.017	1.084	1.001	1.034	10
Coimbra HUC, PPE	1.015	1.111	1.002	1.042	21
Entre Douro e Vouga HC, PPE	1.011	1.062	1.001	1.025	3
Médio Tejo HC, PPE	1.016	1.126	1.002	1.048	27
Póvoa de Varzim/Vila do Conde HC, PPE	1.009	1.098	1.000	1.036	12
Tâmega e Sousa HC, PPE	1.008	1.058	1.002	1.023	2
Tondela-Viseu HC, PPE	1.033	1.131	1.004	1.056	29
Trás-os-Montes e Alto Douro HC, PPE	1.011	1.086	1.002	1.033	9
Cova da Beira HUC, PPE	1.027	1.200	1.002	1.076	33
Lisboa Central HUC, PPE	1.016	1.103	1.007	1.042	19
São João HUC, PPE	1.025	1.315	1.008	1.116	38
Algarve HUC, PPE	1.011	1.076	1.004	1.030	7
Porto HUC, PPE	1.026	1.245	1.006	1.092	37
Lisboa Norte HUC, PPE	1.035	1.226	1.004	1.089	36
Vila Nova de Gaia/Espinho HC, PPE	1.026	1.113	1.010	1.050	28
Senhora da Oliveira Hospital, PPE	1.032	1.184	1.003	1.073	32
Braga Hospital, PPE	1.008	1.035	1.004	1.016	1
Cascais Hospital, PPP	1.045	1.161	1.026	1.077	34
Loures Hospital, PPP	1.021	1.051	1.012	1.028	6
Vila Franca de Xira Hospital, PPP	1.026	1.219	1.007	1.084	35
Figueira da Foz District Hospital, PPE	1.015	1.120	1.001	1.045	24
Santarém District Hospital, PPE	1.021	1.102	1.001	1.041	18
Espírito Santo de Évora Hospital, PPE	1.011	1.092	1.003	1.035	11
Fernando Fonseca Hospital, PPE	1.053	1.393	1.008	1.151	39
Garcia de Orta Hospital, PPE	1.011	1.059	1.005	1.025	4
Santa Maria Maior Hospital, PPE	1.014	1.120	1.000	1.045	23
Guarda LHU, PPE	1.013	1.096	1.001	1.037	13
Castelo Branco LHU, PPE	1.014	1.111	1.001	1.042	20
Matosinhos LHU, PPE	1.013	1.078	1.002	1.031	8
Alto Minho LHU, PPE	1.014	1.098	1.002	1.038	14
Baixo Alentejo LHU, PPE	1.016	1.062	1.002	1.026	5
Litoral Alentejano LHU, PPE	1.052	1.158	1.002	1.071	31
Nordeste LHU, PPE	1.015	1.107	1.001	1.041	16
Norte Alentejano LHU, PPE	1.012	1.100	1.002	1.038	15

By observing Table 4.5, one may verify that the Braga Hospital appears to be the best performer when analyzing quality-related variables. This healthcare provider is followed by Tâmega e Sousa Hospital Center and Entre Douro e Vouga Hospital Center.

4.2.3 Access-related discriminated scores

Table 4.6 presents the discriminated DEA scores for the 39 healthcare providers for models VII to IX.

Table 4.6. Access-related discriminated scores.

Healthcare provider	Model VII	Model VIII	Model IX	Average	Ranking
Barreiro/Montijo HC, PPE	1.093	1.072	1.144	1.103	5
Leiria HC, PPE	1.129	1.185	1.504	1.273	28
Lisboa Ocidental HC, PPE	1.151	1.173	1.756	1.360	34
Setúbal HC, PPE	1.060	1.084	1.488	1.211	21
Baixo Vouga HC, PPE	1.079	1.107	1.331	1.172	16
Médio Ave HC, PPE	1.083	1.091	1.977	1.384	36
Coimbra HUC, PPE	1.056	1.061	1.393	1.170	15
Entre Douro e Vouga HC, PPE	1.102	1.163	1.180	1.148	13
Médio Tejo HC, PPE	1.106	1.156	1.400	1.221	22
Póvoa de Varzim/Vila do Conde HC, PPE	1.031	1.013	1.153	1.066	1
Tâmega e Sousa HC, PPE	1.176	1.129	1.279	1.195	19
Tondela-Viseu HC, PPE	1.175	1.107	1.239	1.174	17
Trás-os-Montes e Alto Douro HC, PPE	1.093	1.067	1.196	1.119	7
Cova da Beira HUC, PPE	1.128	1.136	1.943	1.402	38
Lisboa Central HUC, PPE	1.132	1.114	1.856	1.367	35
São João HUC, PPE	1.217	1.227	1.272	1.239	24
Algarve HUC, PPE	1.081	1.093	1.254	1.143	12
Porto HUC, PPE	1.217	1.146	1.176	1.180	18
Lisboa Norte HUC, PPE	1.205	1.256	1.744	1.402	37
Vila Nova de Gaia/Espinho HC, PPE	1.082	1.064	1.147	1.098	4
Senhora da Oliveira HC, Guimarães, PPE	1.163	1.212	2.225	1.533	39
Braga Hospital, PPE	1.217	1.041	1.135	1.131	9
Cascais Hospital, PPP	1.046	1.049	1.131	1.075	2
Loures Hospital, PPP	1.127	1.198	1.557	1.294	30
Vila Franca de Xira Hospital, PPP	1.176	1.134	1.409	1.239	25
Figueira da Foz District Hospital, PPE	1.109	1.082	1.184	1.125	8
Santarém District Hospital, PPE	1.113	1.104	1.758	1.325	32
Espírito Santo de Évora Hospital, PPE	1.222	1.114	1.472	1.269	26
Fernando Fonseca Hospital, PPE	1.112	1.222	1.338	1.224	23
Garcia de Orta Hospital, PPE	1.158	1.071	1.698	1.309	31
Santa Maria Maior Hospital, PPE	1.122	1.060	1.642	1.275	29
Guarda LHU, PPE	1.046	1.061	1.295	1.134	11
Castelo Branco LHU, PPE	1.043	1.068	1.142	1.084	3
Matosinhos LHU, PPE	1.086	1.059	1.900	1.348	33
Alto Minho LHU, PPE	1.093	1.063	1.160	1.105	6
Baixo Alentejo LHU, PPE	1.101	1.066	1.443	1.203	20
Litoral Alentejano LHU, PPE	1.072	1.159	1.274	1.168	14
Nordeste LHU, PPE	1.049	1.089	1.261	1.133	10
Norte Alentejano LHU, PPE	1.117	1.115	1.581	1.271	27

By observing Table 4.6, one may verify that the Póvoa do Varzim/Vila do Conde Hospital Center appears to be the best performer when analyzing access-related variables. This healthcare provider is followed by Cascais Hospital and Castelo Branco LHU.

4.3 Malmquist Indices

Tables 4.7 to 4.9 provide the productivity indexes for the sample described in Chapter 3. As the results of the regional comparison based on the index I^{AB} should not be interpreted in isolation, the results for the efficiency spread (IE^{AB}), the productivity of the frontier (IF^{AB}), and the overall performance (I^{AB}) are all presented in the same table. Once more, cluster A denotes LHUs and cluster B, non-LHUs.

The results in Tables 4.7 to 4.9 are reported such that a value larger than unity indicates that cluster A (LHUs) has a better performance status than cluster B (non-LHUs) regarding that index. All index values between 0.95 and 1.05 are considered not statistically significant.

4.3.1 Quality- and access-related indices

Table 4.7 reports the results of all the three productivity indexes corresponding to the analysis of both quality- and access-related variables.

Table 4.7. Quality- and access-related productivity indexes.

	IE^{AB}	IF^{AB}	I^{AB}
Model I	1.010	0.994	1.004
Model II	1.003	0.998	1.001
Model III	1.006	0.997	1.003

By observing Table 4.7, one may observe that the clusters appear to exhibit the same relationship for all three models. In fact, through models I to III, the results reveal that hospitals within LHUs appear to have better overall performance ($I^{AB} > 1$), a lower partial performance spread ($IE^{AB} > 1$), yet a lower frontier-shift related performance ($IF^{AB} < 1$), than non-LHUs.

Nevertheless, it is fundamental to remark that all of the productivity index values are between 0.95 and 1.05, meaning that the relationships presented above are not statistically significant.

4.3.2 Quality-related indices

Table 4.8 reports the results of all the three productivity indexes corresponding to the analysis of quality-related variables.

Table 4.8. Quality-related productivity indexes.

	IE^{AB}	IF^{AB}	I^{AB}
Model IV	1.003	0.995	0.998
Model V	1.025	0.975	0.999
Model VI	1.004	0.997	1.001

For models IV and V, cluster A appears to exhibit a lower partial performance spread ($IE^{AB} > 1$), a lower frontier-shift related performance ($IF^{AB} < 1$), and a lower overall performance ($I^{AB} < 1$) when comparing to the cluster of non-LHUs. Regarding model VI, cluster A appears to exhibit similar behaviours on partial performance spread and frontier-shift related performance, yet a higher overall performance ($I^{AB} < 1$), when comparing to cluster B. Nevertheless, it is essential to recall that values between 0.95 and 1.05 are not considered statistically significant, meaning that the previously announced relationships regarding access and its timeliness dimension are not significant from a statistical point of view.

4.3.3 Access-related indices

Table 4.9 reports the results of all the three productivity indexes relating to the comparison of access-related performance between LHUs (cluster A) and non-LHUs (cluster B).

Table 4.9. Access-related productivity indexes.

	IE^{AB}	IF^{AB}	I^{AB}
Model VII	1.035	0.978	1.012
Model VIII	1.037	0.991	1.028
Model IX	1.012	0.919	0.930

Through models VII and VIII, cluster A appears to exhibit lower partial performance spread ($IE^{AB} > 1$), a lower frontier-shift related performance ($IF^{AB} < 1$), and a higher overall performance ($I^{AB} > 1$), when comparing to cluster B. Nevertheless, it is important to recall that values between 0.95 and 1.05 are not considered statistically significant, meaning the previously announced relationships, regarding access and its timeliness dimension, are not significant from a statistical point of view.

Consider model IX. The cluster of LHUs exhibit a statistically significant lower frontier-shift related performance ($IF^{AB} < 1$). In consequence, as IF^{AB} dominates expression 3.10, hospitals within LHUs also exhibit a statistically significant lower overall performance ($I^{AB} < 1$), when analyzing services availability. The results exhibit, nevertheless, statistically insignificant results on partial performance spread.

4.4 Findings discussion

This study proposes a new framework to evaluate the vertical merging of Portuguese healthcare providers. Process variables that describe the quality and the access to healthcare were used to determine DEA scores and productivity indexes by comparing hospitals within LHUs with SHs and HCs.

One started this study by introducing some basic statistics concerning each of the process variables used in the case study (Table 3.5). The statistical tests performed suggest a superior overall performance from the cluster of LHUs. More specifically, assuming the same weight to each of the variables, the cluster of LHUs exhibited similar performance to the cluster of non-LHUs on care appropriateness, superior performance on clinical safety and timeliness, and lower performance on services availability. As discussed earlier, this procedure was essential to obtain an initial perception of each cluster's performance regarding each of the process variables and, consequently, each of the outcome's dimensions. As a matter of fact, some findings were already expected. Consider the services' availability dimension. One may have anticipated that hospitals within LHUs would be the worst performers regarding services availability, as all of the LHU models were implemented in the areas of the interior of Portugal. By analyzing these results in parallel to the variables incorporated in this dimension, particularly FTE nurses per 1,000 standard patients (g_{16}^{+}) and FTE doctors per 1,000 standard patients (g_{17}^{+}), one may recall one serious issue introduced in Subchapter 1.1: there is a lack of health workforce in isolated interior regions of Portugal.

4.4.1 Partial performance

To tackle environmental disparities like the one presented before, a robust method that considers the environment of healthcare providers was implemented. Using a model like so, one may understand whether performance is influenced by the environment in a way that significantly contests the results of the initial statistical analysis.

The first robust methodology, DEA, provided efficiency scores for each of the DMUs, which reached a number of 2,340 in most models. As previously mentioned, one used statistical tests to analyse the statistical significance of the DEA scores.

4.4.1.A Quality- and access-related partial performance

First, by designing models that include both quality- and access-related variables (models I to III), one may draw conclusions on each cluster's overall partial performance. As described in Subchapter 4.1.1, the statistical tests suggest that hospitals within LHUs exhibit a slightly higher partial performance than the cluster of non-LHUs. Indeed, p-values exhibit statistical significance for the three models. Nevertheless, one may also look at the average values of the DEA scores. By directly observing the DEA scores, one may verify not only that all of the average values for each cluster are extremely close to 1, but also that, in absolute terms, the differences between these values are extremely small. This to say that even though the performed tests exhibit statistically significant differences, both clusters are extremely close to the efficiency frontier in absolute terms, with relatively small differences between themselves.

One may also discuss some statistical oversights. It is widely known that the majority of statistical tests require a variety of assumptions. Specifically, the utilization of a Student's t-test requires, among others, the assumption of normality (Neideen and Brasel, 2007). This assumption is strong as it is not clear whether it correctly describes the sample's distribution. Conversely, Kruskal-Wallis test does not assume the normality of the sample, the reason why it was performed to introduce additional veracity to the results. Nevertheless, this last statistical test assumes the residuals' normality (Neideen and Brasel, 2007). This is to say that one may rely on the results of the statistical tests, however, one should not underestimate the assumptions made to perform these tests. This deliberation applies to all the statistical tests performed in this study.

Returning to Table 4.1, it is important to profoundly discuss the closeness of the results to the efficiency frontier. Recalling the results presented in Subchapter 4.1, one may notice that the results regarding outcome-specific performance analysis are not so close to the efficiency frontier. This may suggest that the more variables used in the models, the closer to 1 the results will be. In theory, it is known that a considerable number of DMUs are considered relatively efficient when there is an extravagant number of inputs or outputs relative to the number of units (Adler and Golany, 2007). In the present case study there may be a larger number of used outputs relative to the number of units in some models since it is restricted by the conditional formulations that introduce the environmental effect. The PCA procedure decreased this effect by allowing 7 components in model I (instead of the original 14 variables) and 9 components in model II and III (instead of the original 17 variables). Nevertheless, it looks like this effect was not completely vanished.

Now comparing the previously announced results with the literature available. From all the ten studies that describe the impact of vertical integration on technical efficiency, only one appears to include both quality- and access-related variables in its model. Leleu et al. (2017) uses a DEA model to analyze the technical efficiency of vertically integrated healthcare providers. The researchers incorporate FTE registered and licensed practical nurses, and the number of beds as inputs, and mortality and readmission rates as outputs. In such a way, researchers include the care appropriateness dimension of quality and the services availability dimension of access, which should also be an excellent proxy to evaluate partial performance. For the sake of this comparison, it is essential to recall that the present BoD model only includes variables in the outputs, as dummy inputs are used in the model design. Nevertheless, the usage of KPIs as outputs enables comparisons with standard DEA models. Therefore, one may state that the present results are in line with the ones from the researchers, *i.e.*, backward integration appears to have positive effects on the partial performance of healthcare providers. The American researchers argued the results were expected, as physicians control around 80% of the healthcare costs and, therefore, the majority of the process. This is in line with discussions from Chilingerian and Sherman (1990). Portuguese studies also believe that physicians' association in the vertical integration process is an important condition for the success of this model in Portugal (Lopes et al., 2017).

Ultimately, it should be relevant to compare the present study with societies demographically identical to the Portuguese. In that context, one may recall analyses of the Alzira's model performance. The Spanish literature does not reach a concrete conclusion on the impact of vertical integration on efficiency, as one study reaches positive results, yet two reach no significant changes. One may hypothesize that efficiency decrements could be out of question. Nevertheless, more studies would be needed to properly compare health systems, particularly studies that consider quality and access variables. Lastly, one should also add that the difference in the contractual environment (Alzira is a PPP and the Portuguese LHU model consists of a PPE), could introduce some bias in the potential extrapolation of the Spanish results.

4.4.1.B Quality-related partial performance

Second, quality-related partial performance may be analyzed independently by examining the results from models IV to VI. As confirmed in Subchapter 4.1.2, the statistical tests suggest that hospitals within LHUs exhibit slightly higher partial performance non-LHUs, when analyzing models V and VI. Nevertheless, under model IV, no significant statistical differences between clusters are encountered. These results appear to be contradictory, as differences in care appropriateness and clinical safety (quality dimensions) are significant, but at the overall quality level, they are not. Nevertheless, this occurrence may have a reasonably simple explanation. A well-known practical flaw of DEA models is that, when estimating the scores, variables with zero weights have no influence on the score, despite the alleged importance they may have (Førsund, 2013). In this sense, in the present analysis, some variables may have zero weight in model IV, which may have caused a non-significant result. This situation can be contoured by introducing restrictions on weights, making it impossible for them to reach zero. However, it is also crucial to remember that the DEA CRS model was applied in its envelopment form, the usual recommended form to solve the computation, and weight restrictions may only be incorporated when the model is in its multiplier form. A potential way to confirm the present hypothesis would be to additionally design and run a model in the multiplier form.

From the ten studies collected in Chapter 2, which analyze the impact of vertical integration on efficiency, none appears to be only using quality-related variables in its DEA or SFA model. Nevertheless, quality is a vastly studied outcome in the literature, even though it is analyzed directly and not through efficiency-related analyses. Recalling Chapter 2, a consensus was not reached regarding the impact of vertical models on quality of care, as twenty studies exhibit an increased quality of care, four, decreased, and fourteen, unchanged. The thirty-eight studies were further scrutinized. It was possible to identify a positive effect of this organizational strategy on process of care measures, yet, no well defined conclusions regarding mortality and readmission rates. By taking a further look at the variables presented in Figure 3.1, one may verify that one includes KPIs that consider readmission rates in the care appropriateness dimension, and several process of care KPIs both in the appropriateness and the clinical safety dimension. Therefore, one may state straightaway that the statistically significant results obtained from model VI (clinical safety) for quality-related partial performance appear to be consistent with the literature review elaborated in Chapter 2, as several studies also appear to exhibit a positive effect of this strategy on process of care measures (Brickman et al., 1998; Gillies et al., 2006; Leibert, 2011; Falces et al., 2011; Herrel et al., 2017; Short and Ho, 2019). Thereupon, by focusing on model V (care appropriateness), one may state that the positive results regarding the impact of this strategy on this dimension are not supported by the literature, as there is a significant number of researchers that suggest no impact on readmission rates (Madison, 2004; Scott et al., 2017; Li et al., 2018; Casalino et al., 2018; Konetzka et al., 2018; Pesko et al., 2018; Ho et al., 2019), a measure that one may include in the care appropriateness dimension.

As previously discussed in Chapter 2, some researchers defend that the implementation of vertical integration models in the delivery care system explicitly improves the continuity of care (Carlin et al., 2015). From all the measures, the process of care quality measures are the ones that have more potential to benefit from better coordination of care (Short and Ho, 2019). Therefore, successfully vertically integrated healthcare providers should exhibit enhancements on the communication between providers and increments of access to resources. In such a way, processes of care that require providers and resources of distinct levels of care should exhibit improvements. By recalling the KPIs incorporated in the clinical safety dimension, one may detail several processes of care that should require a strong communication across the first two levels of care. Consider decubitus ulcers (g_4^-). It is known that the process of care for this condition is extremely dependent on the coordination of care between primary and acute care, as it covers two extensive components: the prevention and the management of the ulcer (National Institute for Health and Care Excellence, 2014). Thus, the process implies using devices and treatment strategies provided in both primary care and secondary care settings. Additionally, consider the postoperative conditions, also incorporated in the clinical safety dimension in variables g_6^- and g_7^-. Assuming enhancements in continuity of care, the articulation of preoperative care with the surgery environment will also increase. Indeed, researchers defend that the integration between primary care physicians and specialists may lead to increased efficiencies in the treatment of patients with certain pathologies (Brot-Goldberg and de Vaan, 2018). Thus, in the view of the previous discussion, a positive impact of vertical integration on clinical safety-related partial performance, observed from model VI, is completely understandable when in the presence of, in fact, a successful vertical model.

Additionally, it is fundamental to reaffirm that the majority of the body of evidence included in Chapter 2 is from the US, which is known to have a peculiar healthcare system. More specifically, the US presents a free market system, while most of the European systems, including Portugal, pursue a rigorously regulated ideological framework, hence exhibiting a significantly different environment (Janus and Minvielle, 2017). For this reason, European healthcare managers should not rely entirely on the effects of vertical integration in US healthcare providers to decide whether to implement it in their countries. Looking closely at the sample of 64 studies, one can find a Portuguese study that compares readmission rates in LHUs with non-integrated hospitals, using a DID estimation. This study should be more relevant to compare with the present analysis. Lopes et al. (2017) argue that the Portuguese vertical healthcare model appears to affect positively readmission rates. One may clearly state that the present case study and this study from the literature reached conclusions that point in the same direction, giving potential hope for the success of the Portuguese model. In addition to the backward Portuguese model, some researchers reached positive results when analyzing forward vertical models (Gupta et al., 2019; David et al., 2011), which consists of integrating acute care and post-acute care settings. By reflecting on this aspect, one might affirm that when talking about readmission rates, it seems reasonable that this clinical measure exhibits greater improvements when in the presence of a forward integration model, as post-acute care treatment strategies have significant importance in the potential of readmission of a patient into the delivery care system (Li et al., 2016; Lee et al., 2019).

Without neglecting the statistical significance of the results obtained by model VI, which clearly exhibits enhancements on clinical safety-related partial performance by integrating Portuguese health centers and hospitals, one other Portuguese framework may be lightly discussed. The National Network of Integrated Continuous Care (Rede Nacional de Cuidados Continuados Integrados, RNCCI) is the Portuguese version of a forward vertical integration model. This organizational model intends to fill the gap between hospital discharges and tertiary care, mainly to an older age range (Lourenço et al., 2010). In theory, it should not only increase the access to a age group, that commonly exhibits a large dependence index, but should also be the framework with the greatest potential to decrease readmissions rates to hospital settings. For this reason, in future work, this network could be analyzed in parallel to hospitals within LHUs to hypothesize about the success of vertical integration in Portugal.

4.4.1.C Access-related partial performance

Access-related partial performance may also be analyzed independently. As explained in Subchapter 4.1.3, the statistical tests suggest that hospitals within LHUs exhibit slightly higher partial performance non-LHUs, as the p-values exhibit statistical significance for all the three models (VII to IX). Therefore, one might affirm that LHUs appear to exhibit better partial performance on access and the timeliness and the services availability dimensions.

From the ten efficiency-related analysis studies collected in Chapter 2, six appear to be using services availability variables as their models' inputs. In opposition, the timeliness dimension of access does not appear to be the main focus in the literature, as no study was found to be examining this dimension. Nevertheless, by focusing on services availability, one may affirm that the present results, from model IX, are coherent with three studies from the previously announced body of evidence (Chu et al., 2002; Cho et al., 2014; Caballer-Tarazona and Vivas-Consuelo, 2016). The three remaining articles exhibit no efficiency changes upon implementing this strategy (Carey, 2003; Alonso et al., 2014; Comendeiro-Maaløe et al., 2019). It is important to refer that, even though these variables were considered to be access-related in the conceptual framework of the present case study, the researchers do not directly hypothesize about access. Nevertheless, the highlight of these papers' results is still relevant.

Recalling Subchapter 3.2, when analyzing each of the access-related variables using basic statistics, cluster B appears to outperform cluster A, mainly in the services availability dimension. As previously discussed, this seemed reasonable, as there is a lack of health workforce in isolated interior regions of Portugal, location where most of the LHU models were implemented. As it would not be fair to compare clusters in such different conditions, environmental variables were considered. As perceived in the previous paragraph, by taking into account the environmental effect, hospitals within LHUs exhibit a higher partial performance regarding access and both access dimensions. One may, therefore, hypothesize about the environmental variables that most influenced the results. Consider z_{13} and z_{14}, the number of inhabitants per doctor and the number of inhabitants per pharmacist, respectively. Researchers defend that health resources indicators like these strongly influence the overall measure of access to healthcare (Levesque et al., 2013), and even more specifically, access to healthcare by vulnerable groups (Hendryx et al., 2002). By further analyzing Table 3.6, one may verify that for both variables, the cluster of LHUs exhibits much higher average values. By observing z_{13} individually, the average value of inhabitants per doctor for the cluster of LHUs is more than double of the average value for the cluster of non-LHUs. Thus, this variable exhibits extremely relevant differences. Additionally, consider z_3, purchasing power per capita. By retaking a look at Table 3.6, one may observe a significantly lower average of purchasing power per capita in cluster A. Indeed, researchers suggest that the household purchasing power parity may have the potential to lead to lower utilization of healthcare, more peculiarly, preventive healthcare (Bronchetti et al., 2019).

In practice, it is important to refer that all the environmental variables that exhibit a statistically significant difference between clusters (population density, purchasing power per capita, illiteracy rate, elderly rate, death rate, sec./terc. education rate, dependence index, inhabitants per doctor, and inhabitants per pharmacist) contribute to the environment effect introduced in the DEA model, which is clarified in Subchapter 3.3.2.

When talking about direct analyses of access, Haddad et al. (2020) stands out as the only paper in the literature that analyzes the effect of vertical integration on this outcome. The researchers developed a study on the access to surgical care when in the presence of hospital-physician affiliations, reaching positive conclusions towards implementing this organizational strategy. The results are consistent with the work of the researchers. Nevertheless, it is also important to remember from Chapter 2, that these researchers discuss an eventual increase in costs and prices of care, following the increments in access. In this sense, even if vertical models allow a better access to healthcare for the Portuguese population, costs and prices of care should be analyzed in parallel to make sure cost-containment political measures are not jeopardized.

Ultimately, one may recall that the LHU model was mainly implemented in interior regions of Portugal. As a matter of fact, the idea of vertical integration in rural areas is broadly explored in the literature. Indeed, the concept of integrated health networks is consistently applied in rural areas. Evidence suggests that physician-hospital integration models in rural areas are successful (Morrisey et al., 1990; Moscovice et al., 1998; Stensland and Stinson, 2002). Researchers believe this kind of network detain the potential for enhancing the delivery and guarantee of rural healthcare by perpetuating local access to care in rural areas (Moscovice et al., 1997). Therefore, the results obtained in this study, for the Portuguese vertical model, align with the literature.

4.4.2 Discriminated partial performance

Recalling Subchapter 4.2, Tables 4.4 to 4.6 presented the discriminated DEA scores for the 39 healthcare providers for models I to III, models IV to VI, and models VII to IX, respectively. As it is of a clear interest to mention and describe the best performing healthcare providers from the sample, the ranking of each one was also presented. Tables 4.10 to 4.12 are additionally developed to ease the analysis of the ranking and, therefore, discuss the results. It is important to refer that this ranking allows us to analyze not only each healthcare provider from an individual point of view, but also to understand whether LHUs represent a significant portion of the sample of the 10 best performing healthcare providers in Portugal.

As previously announced, when including both quality- and access-related variables, Póvoa do Varzim/Vila do Conde HC exhibits the highest partial performance, followed by Guarda LHU and Braga Hospital (Table 4.10).

Table 4.10. Quality- and access-related partial performance (Top 10 healthcare providers).

Ranking	Healthcare provider
1	Póvoa de Varzim/Vila do Conde HC, PPE
2	Guarda LHU, PPE
3	Braga Hospital, PPE
4	Nordeste LHU, PPE
5	Garcia de Orta Hospital, PPE
6	Castelo Branco LHU, PPE
7	Algarve HUC, PPE
8	Baixo Alentejo LHU, PPE
9	Coimbra HUC, PPE
10	Fernando Fonseca Hospital, PPE

By observing Table 4.11, one may recall the previously described results. When including quality-related variables, Braga Hospital exhibits the highest partial performance, followed by Tâmega e Sousa HC and Entre Douro e Vouga HC.

Table 4.11. Quality-related partial performance (Top 10 healthcare providers).

Ranking	Healthcare provider
1	Braga Hospital, PPE
2	Tâmega e Sousa HC, PPE
3	Entre Douro e Vouga HC, PPE
4	Nordeste LHU, PPE
5	Garcia de Orta Hospital, PPE
6	Castelo Branco LHU, PPE
7	Algarve HUC, PPE
8	Baixo Alentejo LHU, PPE
9	Trás-os-Montes e Alto Douro HC, PPE
10	Médio Ave HC, PPE

Again, one may recall previously described results, by observing Table 4.12. When including access-related variables, Póvoa do Varzim/Vila do Conde HC exhibits the highest partial performance, followed by Cascais Hospital and Castelo Branco LHU.

Table 4.12. Access-related partial performance (Top 10 healthcare providers).

Ranking	Healthcare provider
1	Póvoa de Varzim/Vila do Conde HC, PPE
2	Cascais Hospital, PPP
3	Castelo Branco LHU, PPE
4	Vila Nova de Gaia/Espinho HC, PPE
5	Barreiro/Montijo HC, PPE
6	Alto Minho LHU, PPE
7	Trás-os-Montes e Alto Douro HC, PPE
8	Figueira da Foz District Hospital, PPE
9	Braga Hospital, PPE
10	Nordeste LHU, PPE

One might observe that 4 out of the 8 LHUs in analysis incorporate the list of the 10 best performing healthcare providers from Table 4.10, and 3 out of the 8 LHUs in analysis incorporate the lists from Tables 4.11 and 4.12. These results were somewhat expected after observing the cluster efficiency scores, presented in Subchapter 4.1. Nevertheless, it is essential to recall that the DEA model allowed fair comparisons according to the environment on which a particular healthcare provider was inserted. In this context, the efficient frontier of a determined healthcare provider was constructed by comparing it with providers in a similar environment. Studies like the ones developed by IASIST, for example, are usually based on the direct observation of indicators, reason why LHUs may not usually be highlighted as excellent performers. Furthermore, there are other healthcare providers for which one could have expected poor results because of their widely known inefficiencies, as Algarve HUC or even Baixo Alentejo LHU, and which actually appear to be among the best performing healthcare providers. In this context, one should be careful when describing the results. For example, when analyzing Table 4.10, one should avoid to firmly state that the hospital within Guarda LHU appears to be "the second best performing healthcare provider in the sample", but instead, one should state that the hospital within Guarda LHU appears to be "the second best performing healthcare provider in the sample when compared to the healthcare providers in similar environments". These results allow us to transpose the success of the Portuguese vertical model, which is usually applied in rural areas, to a potential success when eventually applied in urban areas or other environments.

Furthermore, one can hypothesize about the fact that hospitals like Porto HUC and Santa Maria Maior Hospital are not represented in any of the three tables previously exhibited in the present subchapter. As a matter of fact, both healthcare providers exhibit a better than average technology of production and conditions for appropriate care. Nevertheless, one should also recall that these providers are backup hospitals for all the country, meaning that complex conditions encountered in any Portuguese region can be redirected to these hospitals. As a consequence, these healthcare providers may exhibit lower performance. To consider situations as the one presented, one could have also introduced complexity variables in the model. Nevertheless, the introduction of complexity variables as the CMI was not possible as LHUs management reports lacked information regarding these variables.

Focusing on LHUs, both Nordeste LHU and Castelo Branco LHU consistently appear in the 10 best performing healthcare providers lists exhibited in Tables 4.10 to 4.12. As LHUs should be the ones that, in theory, mostly benefit from an environmental-based comparison, which is not the prevailing methodology in the literature, these healthcare providers do not appear to have much credited success on other studies in the literature. And as a matter of fact, a common procedure is to remove LHUs from the sample when analyzing the performance of Portuguese healthcare providers, as, most of the time, the researchers believe the effect of vertical integration may create bias in the benchmarking process. In the present study that is precisely the effect that one one intended to study. However, from the sample of seven articles used to find out the usual process and environmental variables used in the literature (Subchapter 3.1), six removed LHUs from the study's sample. These circumstances somewhat limit further comparisons of the discriminated DEA scores with the Portuguese literature.

By observing simultaneously Tables 4.10 and 4.12, one may see that Póvoa do Varzim/Vila do Conde Hospital Center is the healthcare provider with the best partial performance when grouping models I to III and models VII to IX. As a matter of fact, this healthcare provider was rewarded in 2016 with an honour attributed by the consulting company IASIST, by integrating the Top 5 best performing Portuguese hospitals, "Top 5'16 – A Excelência dos Hospitais". This study focused on the analysis of quality of care and efficiency indicators, and operating costs. More recently, in 2019, the Shared Services of the Ministry of Health (Serviços Partilhados do Ministério da Saúde, SPMS) also recognized the efficiency-related efforts of the healthcare provider by distinguishing them with the prize "Eficiência@Catálogo". Therefore, the results appear to be coherent with the several distinctions attributed. Additionally, a study that examined technical efficiency using SFA reached similar conclusions, by considering the Póvoa do Varzim/Vila do Conde HC as one of the most efficient providers in Portugal (Menezes et al., 2006). Even though LHUs are not broadly studied in the literature, by comparing the results regarding other more widely studied healthcare providers, as is the case of this hospital center, one may validate the present model, and, thus, the applicability of all the results of this case study.

4.4.3 Overall performance, partial performance spread, and frontier-shift related performance

The second robust methodology, Malmquist Productivity Index Approach, provided productivity scores for each model. As previously mentioned, the overall performance measure I^{AB} may be decomposed into the partial performance spread component, IE^{AB}, and the frontier-shift related performance, IF^{AB}. One considered a result statistically significant when outside of the interval between 0.95 and 1.05.

Recalling Subchapter 4.3, one may emphasize that statistically significant results are only exhibited on model IX, the one that allows us to study the services availability dimension of access. More specifically, the analysis of model IX enables us to state that cluster B outperforms cluster A in the frontier-shift related performance, IF^{AB}, and in the overall performance, I^{AB} (Table 4.9). In this sense, despite the positive from the DEA model, which suggest that hospitals within LHUs exhibit higher partial performance, the Malmquist Productivity Index Approach adds that, at the same time, hospitals within LHUs appear to have lower frontier-shift related performance than the remaining hospitals, in terms of services availability.

First of all, and as previously described in Subchapter 4.4.1, one may recall that the DEA scores obtained from the DEA model are significant at the 0.01 level. For the Malmquist Productivity Index results, one considered 5% significance level to be a balanced value. Using that 5% significance level, the interval of non-significance (0.95 to 1.05) was constructed. It can be observed that the IE^{AB} index scores for all the 9 models appear to be within the defined interval. This means that no model exhibits significant differences in partial performance spread between the two clusters. Nevertheless, one may highlight a significant sensitivity of the results to the choice of the non-significance interval. As a matter of fact, if one proposed a less conservative approach to construct the interval of non-significance, for example with a 1% significance level (0.99 to 1.01), models I, III, V, VII, and VIII would become significant and result in a lower partial performance spread for cluster A, which would be consistent with the DEA's results.

As previously discussed in Chapter 2, productivity is a relatively broader concept than efficiency. Productivity may be regarded as the combination of technical efficiency, "doing things right", and effectiveness, "doing the right things" (Fried et al., 2008). Accordingly, one may hypothesize that hospitals within LHUs appear to be doing "things right" but not "the right things". Thus, the concept of effectiveness may be discussed when in the presence of vertical integration. In the healthcare spectrum, when one talks about "doing the right things", one is debating the usage of the right healthcare resources. Indeed, when implementing a vertical model within the delivery care system, one may observe significant shifts in the process of care and, therefore, in the resources used. This occurrence is usually present as a consequence of the limits imposed. By wanting to increase continuity of care, there are imposed restrictions on the possible healthcare settings for where the patient may be redirected. In this sense, researchers discuss that even if such models improve policy-relevant outcomes, as costs of care, prices of care, efficiency, quality of care, or even access, by improving communication across care settings, they can also increase inappropriate referrals (Baker et al., 2014). Since the main flows in a vertical model, whether of real, financial, or informational nature, are designed to improve continuity of care, many restrictions on the care units used are imposed for a particular process of care, which would not happen otherwise. Thus, concerns regarding the referral model may be one of the reasons for potential effectiveness decrements when in the presence of vertical models. Baker et al. (2014) also add that these unsuitable referrals may be derived from implicit payments, meaning that as this type of model tends to limit patients' flow, some doctors pay for referrals. This last circumstance should not be present in the Portuguese healthcare environment, but in the US, where vertical integration models are implemented in a free market delivery care system.

Again, when talking about frontier-shift related performance, the only significant result was verified in model IX, suggesting lower levels of productivity for hospitals within LHUs. From the literature review elaborated in Chapter 2, one may observe that two studies compare vertically integrated healthcare providers and traditional models regarding this outcome (Wang et al., 2001; Chunn et al., 2018). Wang et al. (2001) reaches increments in productivity upon the integration of healthcare providers, while Chunn et al. (2018) suggest decrements. Therefore, the literature appears to have not yet reached a clear consensus regarding the impact of this strategy on this outcome. Nevertheless, and although Chunn et al. (2018) analyzes clinical measures of productivity, which should be more comparable when analyzing quality dimensions, the results from the researchers and the present study both point in the same direction. Expressly, vertical models appear to have a lower productivity than traditional models. The same researchers defend that a lower clinical productivity may be influenced by some forms of physician compensation. As it is known, this kind of strategy takes away the independence of physicians by integrating them in a fixed-flow model with a specific hospital. Clearly, this circumstance may be an obstacle for several physicians, as these models take away, or somewhat limit, their independence as healthcare practitioners. To convince healthcare practitioners to incorporate this kind of models, healthcare managers may introduce forms of compensation, as higher financial compensation and lower workloads (Chunn et al., 2018). Particularly, the effect of lower workloads may decrease clinical productivity. As a matter of fact, researchers defend that clinical productivity is associated with the tenor of staff contracts (Appleby et al., 2010). More specifically, an analysis by McKinsey & Company for the Department of Health in 2009 found that, even when contracts have the standard number of hours, a lower utilization of the contracted hours of general practitioners causes a lower overall productivity of the healthcare provider (McKinsey & Company, 2009). This evidence proves that some forms of physician compensation that may follow the implementation of vertical models can influence the productivity of healthcare providers.

When talking about financial compensation, the only way to confirm such hypothesis would be to introduce HR-related costs in the DEA model, as performed by Caballer-Tarazona and Vivas-Consuelo (2016). The researchers conclude that these costs appear to increase, yet an association with clinical productivity is not discussed. It is important to state that these compensations are feasible in vertical integration models that also have the potential to increase providers' market power (Baker et al., 2014). Finally, one should clarify again that these discussions are only conjectures to explain the results and not absolute truths. Nevertheless, one believes that hypothesising about including additional variables in the models, as HR-related costs, should help future researchers to develop even more complete models that could increase the veracity and applicability of the obtained results.

5

Conclusions

5.1 Final considerations

Health systems worldwide consistently hypothesize about the ideal delivery care model to optimize the value creation in healthcare. Reevaluating the organization of the current health system is crucial. In Portugal, it is particularly essential as heterogeneities in access and quality cause a significant part of today's inefficiencies in the country's delivery care system. Indeed, creating value in health should be one of the country's primary focus as it influences various sectors' economic performance. In that context, it is also important to remember that health considerations should be integrated into policymaking across all sectors, which is manifested by the well-known approach, Health in All Policies (HiAP).

The monitoring and analysis of healthcare providers' performance are exceedingly relevant insofar as assisting policy creation from evaluating former policies. To apply state-of-the-art performance analysis, one should also highlight the significance of operations research. The consistent quantitative research of mathematical models smooths the decision-making process by policymakers or healthcare managers. Indeed, all healthcare stakeholders should consider robust models in addition to their intuitive approaches when making decisions regarding quantifiable problems, enabling potential increments in decisions' effectiveness.

The present study focused exclusively on analyzing two distinct non-parametric models: the DEA model and the non-parametric MI approach. The case study's emphasis on non-parametric models is explained by the dispensableness of a priori defining a functional form of the efficiency frontier.

Following the literature, the primal form of the output-oriented DEA CRS model was performed, in a BoD framework. When analyzing the partial performance using both quality- and access-related variables (models I to III), one may conclude that hospitals included in LHU models exhibit statistically significant higher levels of partial performance than hospitals incorporated within traditional models. A second step comprised the analysis of outcome-specific partial performances (models IV to IX). Even when analyzed independently, both outcome-specific analysis point in the same direction, exhibiting partial performance improvements in vertically integrated hospitals. Furthermore, the DEA model's use also allowed us to perform a discriminated analysis of the 39 healthcare providers. One concluded that several hospitals included in LHUs are above the 75^{th} percentile, which complements the conclusions from the independent two-sample Student's t-test (analysis of the mean) and confirms the ones from the Kruskal-Wallis test (analysis of the median).

Several explanatory factors have been proposed to explain the previously announced conclusions, among which stand out a) the tighter integration of physicians within the complete delivery care process; b) improvements of prevention and management actions for specific health conditions; c) refinements in the articulation of preoperative care with the surgery environment; d) enhancements in the delivery and guarantee of healthcare, by perpetuating local access to care.

Ultimately, the non-parametric MI approach was employed. Unlike the first model, one reached significant results only for the services availability dimension of access, which corresponds to model IX. For that dimension, hospitals included in vertical models exhibit lower frontier shift-related performance levels and lower overall performance than hospitals within traditional models.

The overall consideration for the performance analysis in the current study should take into account both models. The first model presents a positive relationship between the study strategy and the hospital's performance on quality and access. The second one does not present overall significant changes when employing vertical integration. One may conclude that the Portuguese reality should be somewhere in the middle. Meaning, hospitals within the LHU models should exhibit relative improvements on quality and access measures when considered the environment in which these are incorporated.

The potential gains that an LHU model may provide are a consequence of the close relationship with the health centre, enabling a more closely monitoring of the patient by working in integration. The existence of a common institutional structure for both levels of care enables enhanced interoperability, shared information systems, and a faster referral process, being unnecessary to enter the national referral process. When in the presence of other models, as the horizontal one, even if HCs englobe Health Centres Groups (Agrupamentos de Centros de Saúde, ACES) in their coverage area, the fact that there is no shared institutional structure makes any integrated action plan limited. One extremely relevant example of the vertical integration model's power is when hospitals are faced with an abnormally high number of patients with a particular health condition. In a vertical integration model, that information would be immediately shared with the primary care settings, which would create condition-specific prevention measures to tackle that problem. When in the presence of any traditional model, the hospital would have to communicate with Directorate-General for Health (Direção Geral de Saúde, DGS), which would emit a norm, to be shared with the Regional Health Administration (Administração Regional de Saúde, ARS), that in turn would communicate with the proximity Health Centres. The fundamental conclusion of this work is that the enhancements in the processes' streamlining and interoperability in vertically integrated hospitals appear to be precursors of improvements in quality and access compared to traditional hospitals.

Notwithstanding that the LHU model is not one of the centre of researchers' attention, some academic works focused on analysing LHU performance. Nevertheless, a variety of factors distinguish the present study from other academic works. First of all, this research stands out for non-parametric models' employment, which restricts the concept of efficiency to its robust definition: the flow between healthcare inputs and the outputs produced. Moreover, this research takes under consideration the environmental effect, which is crucial, bearing in mind LHUs were predominantly implemented in the interior region of Portugal. Ultimately, the recent data range used, FY2015 to FY2019, allowed for an analysis of a further matured LHU model, unlike former studies, that were performed during the first steps of these entities.

The results of the present study come in the same direction as reforms implemented for the integration of healthcare. The Decree-Law No. 18/2017, of 20th February, appear to establish relevant reforms that should have influenced the observed improvements on the vertical model performance. The aforementioned Decree-Law established LHU-related reforms as the adjusted capitation for risk, the organization of practical guidelines and action plans to consolidate the communication between clinical and non-clinical areas, and the improvement of monitoring and evaluation processes to analyze the performance of LHUs on efficiency, quality, and access.

Finally, and in the context of the results presented, it is possible to denote some implications for healthcare managers and policymakers. Something fairly discussed in the international literature, and disclosed in the literature review in Chapter 2, is the potential deterioration of some policy-relevant measures in vertical models, particularly costs, even when enhancements on quality and access are present. This situation further complicates the decision-making of healthcare stakeholders. In Portugal, there are two essential systems within the hospital care external contractualization that should offset potential cost increments. First, there is a system of incentives for hospital performance, aiming to stimulate continuous improvement and find operational management levers. The fulfillment of this system's objectives represents 10% of the value of the program contract. Additionally, a system of penalties aims to prevent healthcare inadequacies' systematic occurrence, corresponding to a penalty of up to 3% of the value of the program contract. For both systems, the set of objectives considered focuses on access, quality, and efficiency. The financial compensation for enhancements in these metrics introduces flexibility in healthcare managers' and policymakers' decision-making process, enabling the introduction of models like the one in analysis from their frame of reference. Ultimately, it is not clear yet whether this integration strategy increases costs, as there is an evident lack of Portuguese studies on this metric's analysis. Nevertheless, one may still conclude that the strategy appears to be connected with evidence of value creation in Health.

5.2 Research limitations and future work

Undoubtedly, the present study enriches the literature devoted to establish the link between healthcare outcomes to the vertically integrated healthcare providers. Nevertheless, some research limitations should be summarized and potential future work should be highlighted.

First, one may recall that a significant number of papers that study vertical models suggest potential decrements of other measures, particularly costs, even in the presence of enhancements on quality and access. In the present analysis one could not confirm the veracity of this circumstance in the Portuguese context, as cost-related variables were not included in the analysis. Nevertheless, one should recall that the current study focuses on the analysis of the hospital environment within the LHU model. In that context, and considering that real costs are not unbundled in primary and hospital environment in the contract-program, costs could not have been introduced in the model. When researchers have access to the bundled costs, future work may include costs in non-parametric models to reach complete and robust results regarding the success of vertical models in Portugal.

The dominant research limitation of the present study is, indeed, the inclusion of process variables uniquely related with hospital-settings. This circumstance limits conjectures on the success of LHUs on the complete delivery care process (across both levels of care). However, one should denote that the ACSS database only provides process variables in hospital settings. Additionally, even when trying to access public information for the Health Centres Groups (Agrupamentos de Centros de Saúde, ACES), data always appears to be unavailable or not recent. In this context, future reforms may include the enhancement of the data gathering process and of the public data infrastructure, in order to enable a complete analysis of the Portuguese healthcare environment. In such circumstances, future researchers should be able to conclude about the success of vertical model across primary and acute care settings in Portugal.

As previously discussed, from an operational point of view, one could alternatively implement the output-oriented DEA CRS model in its multiplier form, to prevent process variables to have a zero weight. Additionally, future work could also include the implementation of a parametric model, SFA. However, one has to remember SFA's implementation would imply an extra concern with the choice of process variables, since multicollinearity between variables is undesirable in that context.

Ultimately, as several studies in the literature performed for other health systems, future researchers could analyze the complete concept of vertical integration (backward and forward integration) within the Portuguese delivery care system. By including data on the National Network of Integrated Continuous Care (Rede Nacional de Cuidados Continuados Integrados, RNCCI), studies could assist policy-makers to identify potential best practices, or in the other direction, possible flaws, in the continuity of care process in the Portuguese public health system.

6

Bibliography

B Acerete, A Stafford, and P Stapleton. Spanish healthcare public private partnerships: The ‘Alzira model’. *Critical Perspectives on Accounting*, 22(6):533–549, 2011.

N Adler and B Golany. PCA-DEA: Reducing the curse of dimensionality. In: Modeling problem structure and data varieties using Data Envelopment Analysis: A problem-solving handbook (pp. 139-153). *Springer*, 2007.

M Al-Amin. Hospital characteristics and 30-day all-cause readmission rates. *Journal of Hospital Medicine*, 11(10):682–687, 2016.

J A Alexander and M A Morrisey. Hospital-physician integration and hospital costs. *Inquiry*, 25(3): 388–401, 1988.

S Allin, M Grignon, and L Wang. The determinants of efficiency in the Canadian health care system. *Health Economics, Policy and Law*, 11(1):39–65, 2016.

J M Alonso, J Clifton, and D Diaz-Fuentes. The impact of new public management on efficiency: An analysis of Madrid’s hospitals. *Health Policy*, 119(3):333–340, 2014.

A Amirteimoori, D K Despotis, and S Kordrostami. Variables reduction in Data Envelopment Analysis. *Optimization*, 63(5):735–745, 2014.

J Appleby, C Ham, C Imison, and M Jennings. Improving NHS productivity. *The Kings Fund*, 2010. Available online at https://www.kingsfund.org.uk/sites/default/files/field/field_publication_file/improving-nhs-productivity-kings-fund-july-2010.pdf (accessed November 24, 2020).

L C Baker, M K Bundorf, and D P Kessler. Vertical integration: Hospital ownership of physician practices is associated with higher prices and spending. *Health Affairs*, 33(5):756–763, 2014.

R D Banker, A Charnes, and W W Cooper. Some models for estimating technical and scale inefficiencies in Data Envelopment Analysis. *Management Science*, 30(9):1078–1092, 1984.

P Barros, S Machado, and J Simões. Portugal: Health system review. *Health Systems in Transition*, 13 (4):1–156, 2011.

M Borenstein, L V Hedges, J P T Higgins, and H R Rothstein. *Introduction to meta-analysis*. Wiley, 2009. ISBN 9780470057247.

A Bottle and P Aylin. Mortality associated with delay in operation after hip fracture: Observational study. *BMJ*, 332(7547):947–951, 2006.

R Brickman, R Axelrod, D Roberson, and C Flanagan. Clinical process improvement as a means of facilitating health care system integration. *Journal on Quality Improvement*, 24(3):143–153, 1998.

E T Bronchetti, G Christensen, and H W Hoynes. Local food prices, snap purchasing power, and child health. *Journal of Health Economics*, 2019. doi: 10.1016/j.jhealeco.2019.102231.

Z C Brot-Goldberg and M de Vaan. Intermediation and vertical integration in the market for surgeons. *Working Paper 2018*, 2018. Available online at https://econ2017.sites.olt.ubc.ca/files/2019/01/Zarek_Brot-Goldberg_JMP.pdf (accessed November 24, 2020).

M M Byrne and C M Ashton. Incentives for vertical integration in healthcare: The effect of reimbursement systems. *Journal of Healthcare Management*, 44(1):34–44, 1999.

M Caballer-Tarazona and D Vivas-Consuelo. A cost and performance comparison of public private partnership and public hospitals in Spain. *Health Economics Review*, 6(1):17, 2016.

S B Calvert, J M Kramer, K J Anstrom, L A Kaltenbach, J A Stafford, and N M Allen LaPointe. Patient-focused intervention to improve long-term adherence to evidence-based medications: A randomized trial. *American Heart Journal*, 163(4):657–665, 2012.

A S Camanho and R G Dyson. Data Envelopment Analysis and Malmquist Indices for measuring group performance. *Journal of Productivity Analysis*, 26:35–49, 2006.

C Capps, D Dranove, and C Ody. The effect of hospital acquisitions of physician practices on prices and spending. *Journal of Health Economics*, 59:139–152, 2018.

K Carey. Hospital cost efficiency and system membership. *Inquiry*, 40:25–38, 2003.

C S Carlin, B Dowd, and R Feldman. Changes in quality of health care delivery after vertical integration. *Health Services Research*, 50(4):1043–1068, 2015.

L P Casalino, M F Pesko, A M Ryan, J L Mendelsohn, K R Copeland, et al. Small primary care physician practices have low rates of preventable hospital admissions. *Health Affairs*, 33(9):1680–1688, 2014.

L P Casalino, P P Ramsay, L C Baker, M F Pesko, and S M Shortell. Medical group characteristics and the cost and quality of care for Medicare beneficiaries. *Health Services Research*, 53(6):4970–4996, 2018.

D W Caves, L R Christensen, and W E Diewert. The economic theory of index numbers and the measurement of input, output, and productivity. *Econometrica*, 50(6):1393–1414, 1982.

A Charnes, W W Cooper, and E Rhodes. Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2:429–444, 1978.

J Chilingerian and H Sherman. Managing physician efficiency and effectiveness in providing hospital services. *Health Services Management Research*, 3:3–15, 1990.

N Cho, S Lee, and J D Lee. Economic evaluation of the impact of physician-hospital integration and physician boards on hospital expenditure per patient: A 5-year longitudinal study. *Medicine*, 97(41), 2018.

N E Cho, J Chang, and B Atems. The effects of health information technology adoption and hospital-physician integration on hospital efficiency. *American Journal of Managed Care*, 20(17):eSP9–eSP15, 2014.

H Chowdhury and V Zelenyuk. Performance of hospital services in Ontario: DEA with truncated regression approach. *Omega*, 63:11–22, 2016.

H Chu, S Liu, J C Romeis, and C Yaung. The initial effects of physician compensation programs in Taiwan hospitals: Implications for staff model HMOs. *Health Care Management Science*, 6(1):17–26, 2003.

H L Chu, S Z Liu, and J C Romeis. Does the implementation of responsibility centers, total quality management, and physician fee programs improve hospital efficiency? Evidence from Taiwan hospitals. *Medical Care*, 40(12):1223–1237, 2002.

A Chukmaitov, D W Harless, G J Bazzoli, H J Carreta, and U Siangphoe. Delivery system characteristics and their association with quality and costs of care: Implications for accountable care organizations. *Health Care Management Review*, 40(2):92–103, 2015.

A S Chukmaitov, G J Bazzoli, D W Harless, R E Hurley, K J Devers, and M Zhao. Variations in inpatient mortality among hospitals in different system types, 1995 to 2000. *Medical Care*, 47(4):466–473, 2009.

V M Chunn, B Sen, S J O'Connor, W F Jessee, J Sasson, and A Y Landry. Integration of cardiologists with hospitals: Effects on physician compensation and productivity. *Health Care Management Review*, 45(4):342–352, 2018.

F Ciliberto and D Dranove. The effect of physician–hospital affiliations on hospital prices in California. *Journal of Health Economics*, 25(1):29–38, 2006.

M Comendeiro-Maaløe, M Ridao-López, S Gorgemans, and E Bernal-Delgado. A comparative performance analysis of a renowned public private partnership for health care provision in Spain between 2003 and 2015. *Health Policy*, 123(4):412–418, 2019.

W Cooper, L Seiford, and J Zhu. *Handbook on data envelopment analysis*. 2011. ISBN 9781441961501.

D J Crespin, J B Christianson, J S McCullough, and M D Finch. Health system consolidation and diabetes care performance at ambulatory clinics. *Health Services Research*, 51(5):1772–1795, 2016.

A E Cuellar and P J Gertler. Strategic integration of hospitals and physicians. *Journal of Health Economics*, 25(1):1–28, 2006.

N Curry, M Harris, L H Gunn, Y Pappas, I Blunt, et al. Integrated care pilot in north-west London: A mixed methods evaluation. *International Journal of Integrated Care*, 13:e027, 2013.

J Cylus, I Papanicolas, and P C Smith. Health system efficiency: How to make measurement matter for policy and management. *European Observatory on Health Systems and Policies*, 2016. Available online at https://www.euro.who.int/__data/assets/pdf_file/0004/324283/Health-System-Efficiency-How-make-measurement-matter-policy-management.pdf (accessed November 24, 2020).

C M Dahl and L P Kongstad. The costs of acute readmissions to a different hospital: Does the effect vary across provider types? *Social Science Medicine*, 183:116–125, 2017.

G David, E Rawley, and D Polsky. Integration and task allocation: Evidence from patient care. *NBER Working Paper Series*, 2011. doi: 10.3386/w17419.

V Donabedian. Evaluating the quality of medical care. *The Milbank Quarterly*, 83(4):691–729, 2005.

Entidade Reguladora da Saúde. Estudo sobre o desempenho das unidades locais de saúde (Portuguese) [On local health units' performance]. *Entidade Reguladora da Saúde*, 2015. Available online at https://www.ers.pt/uploads/writer_file/document/1297/Estudo_sobre_o_Desempenho_das_ULS_27_01_2015.pdf (accessed November 24, 2020).

C Falces, R Andrea, M Heras, C Vehí, M Sorribes, et al. Integración entre cardiología y atención primaria: Impacto sobre la práctica clínica (Spanish) [On integration between cardiology services and primary care]. *Revista Española de Cardiologia*, 64(7):564–571, 2011.

D C Ferreira. Efficiency and productivity under conditional frameworks. In *Seminar*, Instituto Superior Técnico, 2016.

D C Ferreira and R C Marques. A step forward on order-α robust nonparametric method: Inclusion of weight restrictions, convexity and non-variable returns to scale. *Operational Research - An International Journal*, 20(2), 2017.

D C Ferreira and R C Marques. Do quality and access to hospital services impact on their technical efficiency? *Omega*, 86:218–236, 2018.

D C Ferreira and R C Marques. Public-private partnerships in health care services: Do they outperform public hospitals regarding quality and access? Evidence from Portugal. *Socio-Economic Planning Sciences*, 2020. doi: 10.1016/j.seps.2020.100798.

D C Ferreira, A M Nunes, and R C Marques. Doctors, nurses, and the optimal scale size in the Portuguese public hospitals. *Health Policy*, 122(10):1093–1100, 2018.

D C Ferreira, R C Marques, and A M Nunes. Pay for performance in health care: A new best practice tariff-based tool using a log-linear piecewise frontier function and a dual–primal approach for unique solutions. *Operational Research*, 2019a. doi: 10.1007/s12351-019-00502-3.

D C Ferreira, A M Nunes, and R C Marques. Optimizing payments based on efficiency, quality, complexity, and heterogeneity: The case of hospital funding. *International Transactions in Operational Research*, 2019b. doi: 10.1111/itor.12713.

D C Ferreira, A M Nunes, and R C Marques. Operational efficiency vs clinical safety, care appropriateness, timeliness, and access to health care: The case of Portuguese public hospitals. *Journal of Productivity Analysis*, 2020. doi: 10.1007/s11123-020-00578-6.

G D Ferrier and J S Trivitt. Incorporating quality into the measurement of hospital efficiency: A double DEA approach. *Journal of Productivity Analysis*, 40:337–355, 2013.

A J Fong, M Smith, and A Langerman. Efficiency improvement in the operating room. *Journal of Surgical Research*, 204(2):371–383, 2016.

R W Forsman. Why is the laboratory an afterthought for managed care organizations? *Clinical Chemistry*, 42(5):813–816, 1996.

T M Franke, T Ho, and C A Christie. Length of stay efficiency. *American Journal of Evaluation*, 33(3): 448–458, 2012.

H O Fried, C A K Lovell, and S S Schmidt. *The measurement of productive efficiency and productivity growth*. Oxford University Press, 2008. ISBN 9780198040507.

N Fullman, J Yearwood, S M Abay, C Abbafati, F. Abd-Allah, et al. Measuring performance on the healthcare access and quality index for 195 countries and territories and selected subnational locations: A systematic analysis from the global burden of disease study 2016. *The Lancet*, 391(10136): 2236–2271, 2018.

R Färe and C Lovell. Measuring the technical efficiency of production. *Journal of Economic Theory*, 19: 150–162, 1978.

R Färe, S Grosskopf, B Lindgren, and P Roos. Productivity developments in Swedish hospitals: A malmquist output index approach. In: Data envelopment analysis: Theory, methodology, and applications (pp. 253-272). *Springer*, 1994.

F R Førsund. Weight restrictions in DEA: Misplaced emphasis? *Journal of Productivity Analysis*, 40: 271–283, 2013.

R R Gillies, K E Chenok, S M Shortell, G Pawlson, and J J Wimbush. The impact of health plan delivery system organization on clinical quality and patient satisfaction. *Health Services Research*, 41(4 Pt 1): 1181–1191, 2006.

J B Goes and C Zhan. The effects of hospital-physician integration strategies on hospital financial performance. *Health Services Research*, 30(4):507–530, 1995.

M Gulliford, J Figueroa-Munoz, M Morgan, D Hughes, B Gibson, R Beech, and M Hudson. What does 'access to health care' mean? *Journal of Health Services Research Policy*, 7(3):186–188, 2002.

S Gupta, F D Zengul, G K Davlyatov, and R Weech-Maldonado. Reduction in hospitals' readmission rates: Role of hospital-based skilled nursing facilities. *The Journal of Health Care Organization, Provision, and Financing*, 20:1–12, 2019.

D N Haddad, M J Resnick, and S S Nikpay. Does vertical integration improve access to surgical care for Medicaid beneficiaries? *Journal of the American College of Surgeons*, 230(1):130–135, 2020.

M S Hendryx, M M Ahern, N P Lovrich, and A H McCurdy. Access to health care and community social capital. *Health Services Research*, 37(1):85–101, 2002.

R M Henke, Z Karaca, B Moore, E Cutler, H Liu, W D Marder, and H S Wong. Impact of health system affiliation on hospital resource use intensity and quality of care. *Health Services Research*, 53(1): 63–86, 2018.

L A Herrel, S R Kaufman, P Yan, D C Miller, F R Schroeck, et al. Healthcare integration and quality among men with prostate cancer. *Journal of Urology*, 197(1):55–60, 2017.

V Ho, L Metcalfe, L Vu, M Short, and R Morrow. Annual spending per patient and quality in hospital-owned versus physician-owned organizations: An observational study. *Journal of General Internal Medicine*, 35(3):649–655, 2019.

V Ho, S Tapaneeyakul, L Metcalfe, L Vu, and M Short. Using Medicare data to measure vertical integration of hospitals and physicians. *Health Services and Outcomes Research Methodology*, 2020. doi: 10.1007/s10742-020-00207-7.

R G Holloway and T E Quill. Mortality as a measure of quality: Implications for palliative and end-of-life care. *Journal of the American Medical Association*, 298(7):802–804, 2007.

R S Huckman. Hospital integration and vertical consolidation: An analysis of acquisitions in New York state. *NBER Working Paper Series*, 25(1):58–80, 2005.

J M Huguenin. *Data envelopment analysis (DEA): A pedagogical guide for decision makers in the public sector*. IDHEAP: Chair of Public Finance, 2012. ISBN 2940390541.

R Jacobs, P C Smith, and A Street. *Measuring Efficiency in Health Care: Analytic Techniques and Health Policy*. Cambridge University Press, 2006. ISBN 9780521851442.

K Janus and E Minvielle. Rethinking health care delivery: What European and United States health care systems can learn from one another. *Health Affairs Blog*, 2017. doi: 10.1377/hblog20171214.835155.

R Jones. Length of stay efficiency. *British Journal of Healthcare Management*, 15:563–564, 2009.

J Jung, R Feldman, and Y Kalidindi. The impact of integration on outpatient chemotherapy use and spending in Medicare. *Health Economics*, 2019. doi: 10.1002/hec.3860.

J Khushalani and Y A Ozcan. Are hospitals producing quality care efficiently? An analysis using Dynamic Network Data Envelopment Analysis (DEA). *Socio-Economic Planning Sciences*, 60:15–23, 2017.

T G Koch, B W Wendling, and N E Wilson. How vertical integration affects the quantity and cost of care for Medicare beneficiaries. *Journal of Health Economics*, 52:19–32, 2017.

R T Konetzka, E A Stuart, and R M Werner. The effect of integration of hospitals and post-acute care providers on Medicare payment and patient outcomes. *Journal of Health Economics*, 61:244–258, 2018.

J Kralewski, B Dowd, D Knutson, M Savage, and J Tong. Medical group practice characteristics influencing inappropriate emergency department and avoidable hospitalization rates. *The Journal of Ambulatory Care Management*, 36(4):286–291, 2013.

J E Kralewski, E C Rich, R Feldman, B E Dowd, T Bernhardt, C Johnson, and W Gold. The effects of medical group practice and physician payment methods on costs of care. *Health Services Research*, 35(3):591–613, 2000.

J E Kralewski, B E Dowd, and Y W Xu. Medical groups can reduce costs by investing in improved quality of care for patients with diabetes. *Health Affairs*, 31(8), 2012.

B Lanese. Implementation of the affordable care act: A case study of a service line co-management company. *Journal of Health Organization and Management*, 30(6):818–835, 2016.

R Lear, A D Godfrey, C Riga, C Norton, C Vincent, and C D Bicknell. The impact of system factors on quality and safety in arterial surgery: A systematic review. *European Journal of Vascular and Endovascular Surgery*, 54(1):79–93, 2017.

A J Lee, X Liu, T Borza, Y Qin, B Y Li, K L Urish, et al. Role of post-acute care on hospital readmission after high-risk surgery. *Journal of Surgical Research*, 234:116–122, 2019.

M Leibert. Performance of integrated delivery systems: Quality, service and cost implications. *Leadership in Health Services*, 24(3), 2011.

H Leleu, M Al-Amin, M Rosko, and V G Valdmanis. A robust analysis of hospital efficiency and factors affecting variability. *Health Services Management Research*, 31(1):33–42, 2017.

J Levesque, M F Harris, and G Russell. Patient-centred access to health care: Conceptualising access at the interface of health systems and populations. *International Journal for Equity in Health*, 12(18), 2013.

J Li, M C Daldalian, C Erdmann, K G Hunter, B Sutariya, and R Leung. Hospital readmission after post-acute care at different settings: Estimation using the propensity score matching method. *Journal of Community Medicine Health Education*, 6(2), 2016.

J Li, Z Ye, J M Dupree, B K Hollenbeck, H S Min, et al. Association of delivery system integration and outcomes for major cancer surgery. *Annals of Surgical Oncology*, 25(4):856–863, 2018.

A E Liepert, J Bledsoe, M H Stevens, and A Cochran. Protecting trauma patients from duplicated computed tomography scans: The relevance of integrated care systems. *The American Journal of Surgery*, 208(4):511–516, 2014.

S Lopes, O B Fernandes, A P Marques, B Moita, J Sarmento, and R Santana. Can vertical integration reduce hospital readmissions? A difference-in-differences approach. *Medical Care*, 55(5):506–513, 2017.

A Lourenço, A Dias, A Escoval, A S Ferreira, A T Paquete, et al. *30 anos do Serviço Nacional de Saúde: Um percurso comentado (Portuguese) [On the 30 years of the Portuguese National Health Service]*. Almedina, 2010. ISBN 9789724041100.

C H Lyder, Y Wang, M Metersky, M Curry, R Kliman, N R Verzier, and D R Hunt. Hospital-acquired pressure ulcers: Results from the national medicare patient safety monitoring system study. *Journal of the American Geriatrics Society*, 60(9):1603–8, 2012.

R M Machta, K A Maurer, D J Jones, M F Furukawa, and E C Rich. A systematic review of vertical integration and quality of care, efficiency, and patient-centered outcomes. *Health Care Management Review*, 44(2):159–173, 2019.

K Madison. Hospital–physician affiliations and patient treatments, expenditures, and outcomes. *Health Services Research*, 39(2):257–278, 2004.

McKinsey & Company. Achieving world class productivity in the NHS 2009/10 - 2013/14: Detailing the size of the opportunity. 2009. Available online at `http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/documents/digitalasset/dh_116521.pdf` (accessed November 24, 2020).

J M McWilliams, M E Chernew, A M Zaslavsky, P Hamed, and B E Landon. Delivery system integration and health care spending and quality for Medicare beneficiaries. *JAMA Internal Medicine*, 173(15): 1447–1456, 2013.

A Menezes, M F Rendeiro, and J C Vieira. Eficiência técnica dos hospitais Portugueses 1997-2004: Uma análise (regional) com base num modelo de fronteira estocástica Portuguese [On the technical efficiency of portuguese hospitals]. *Revista Portuguesa De Estudos Regionais*, 12:53–75, 2006.

P K Modi, L A Herrel, S R Kaufman, P Yan, T Borza, et al. Urologist practice structure and spending for prostate cancer care. *Health Services Research*, 130:65–71, 2019.

M Morrisey, J Alexander, and R Ohsfedt. Physician integration strategies and hospital output: A comparison of rural and urban institutions. *Medical Care*, 28:586–603, 1990.

I Moscovice, A Wellever, J Christianson, M Casey, B Yawn, and D Hartley. Understanding integrated rural health networks. *The Milbank Quarterly*, 75(4), 1997.

I Moscovice, M Casey, and S Krein. Expanding rural managed care: Enrollment patterns and prospects. *Health Affairs*, 17:172–179, 1998.

National Institute for Health and Care Excellence. The prevention and management of pressure ulcers in primary and secondary care. *National Institute for Health and Care Excellence Clinical Guidelines*, (179), 2014.

J L Navarro-Espigares and E H Torres. Efficiency and quality in health services: A crucial link. *The Service Industries Journal*, 31(3):385–403, 2011.

T Neideen and K Brasel. Understanding statistical tests. *Journal of Surgical Education*, 64(2), 2007.

H T Neprash, M E Chernew, A L Hicks, T Gibson, and J M McWilliams. Association of financial integration between physicians and hospitals with commercial health care prices. *JAMA Internal Medicine*, 175(12):1932–1939, 2015.

A Niñerola, M Sánchez-Rebull, and A Hernandéz-Lara. Quality improvement in healthcare: Six sigma systematic review. *Health Policy*, 124(4):438–445, 2020.

A M Nunes, D C Ferreira, and A C Fernandes. Financial crisis in Portugal: Effects in the health care sector. *International Journal of Health Services*, 49(2):237–259, 2019.

Organisation for Economic Co-operation and Development. Health at a glance 2019. *Organisation for Economic Co-operation and Development (OECD)*, 2019. Available online at `https://www.oecd-ilibrary.org/docserver/4dd50c09-en.pdf?expires=1606356853&id=id&accname=guest&checksum=89DD5D1D28F011499F37D2F5A34518D9` (accessed November 24, 2020).

J T Pastor and J L Ruiz. Variables with negative values in DEA. In: Modeling data irregularities and structural complexities in data envelopment analysis. *Springer*, pages 63–84, 2007.

M F Pesko, A M Ryan, S M Shortell, K R Copeland, P P Ramsay, et al. Spending per Medicare beneficiary is higher in hospital-owned small- and medium-sized physician practices. *Health Services Research*, 53(4):2133–2146, 2018.

D H Peters, A Garg, G Bloom, D G Walker, W R Brieger, and M H Rahman. Poverty and access to health care in developing countries. *Annals of the New York Academy of Sciences*, 1136:161–171, 2008.

B Post, T Buchmueller, and A M Ryan. Vertical integration of hospitals and physicians: Economic theory and empirical evidence on spending and quality. *Medical Care Research and Review*, 75(4):399–433, 2017.

M Rahman, E C Norton, and D C Grabowski. Do hospital-owned skilled nursing facilities provide better post-acute care quality? *Journal of Health Economics*, 50:36–46, 2016.

B Rechel, J Erskine, B Dowdeswell, S Wright, and M McKee. *Capital Investment for Health: Case studies from Europe*. European Observatory on Health Systems and Policies, 2009a. ISBN 9789289041782.

B Rechel, Steve Wright, Nigel Edwards, B Dowdeswell, and M Mckee. *Investing in hospitals of the future*. European Observatory on Health Systems and Policies, 2009b. ISBN 9789289043045.

K F Rhoads, M I Patel, Y Ma, and L A Schmidt. How do integrated health care systems address racial and ethnic disparities in colon cancer? *Journal of Clinical Oncology*, 33(8):854–860, 2015.

J C Robinson. Organizational economics and health care markets. *Health Services Research*, 36(1 Pt 2):177–189, 2001.

J C Robinson and K Miller. Total expenditures per patient in hospital-owned and physician-owned physician organizations in California. *Journal of the American Medical Association*, 312(16):1663–1669, 2014.

A G L Romme. Vertical integration as organizational strategy formation. *Organization Studies*, 11(2): 239–260, 1990.

M D Rosko, J Proenca, J S Zinn, and G J Bazzoli. Hospital inefficiency: What is the impact of membership in different types of systems? *Inquiry*, 44:335–349, 2007.

L F Rossiter. Expenditures and quality: Hospital- and health system–affiliated versus independent physicians in Virginia. *Southern Medical Journal*, 111(10):597–600, 2018.

K W Scott, E J Orav, D M Cutler, and A K Jha. Changes in hospital–physician affiliations in U.S. hospitals and their effect on quality of care. *Annals of Internal Medicine*, 166(1):1–8, 2017.

D M Shahian, L I Iezzoni, G S Meyer, L Kirle, and S L Normand. Hospital-wide mortality as a quality metric: Conceptual and methodological challenges. *American Journal of Medical Quality*, 27(2):112–123, 2012.

M N Short and V Ho. Weighing the effects of vertical integration versus market concentration on hospital quality. *Medical Care Research and Review*, 2019. doi: 10.1177/1077558719828938.

J A Simões, G F Augusto, I Fronteira, and C Hernández-Quevedo. Portugal: Health system review. *Health Systems in Transition*, 19(2):1–184, 2017.

K W Sowers, P R Newman, and J C Langdon. Evolution of physician-hospital alignment models: A case study of comanagement. *Clinical Orthopaedics and Related Research*, 471(6):1818–1823, 2013.

J Stensland and T Stinson. Successful physician-hospital integration in rural areas. *Medical Care*, 40 (10):908–917, 2002.

N Sullivan and K M Schoelles. Preventing in-facility pressure ulcers as a patient safety strategy: A systematic review. *Annals of Internal Medicine*, 158:410–416, 2013.

D C Szostak. Vertical integration in health care the regulatory landscape. *DePaul Journal of Health Care Law*, 17(2), 2015.

E Thanassoulis. *Introduction to the Theory and Application of Data Envelopment Analysis: A Foundation Text with Integrated Software*. Springer, 2001. ISBN 9781461355380.

R Town and G Vistnes. Hospital competition in HMO networks. *Journal of Health Economics*, 20(5): 733–753, 2001.

C Tsang, P Aylin, and W Palmer. Patient safety indicators: A systematic review of the literature. *Centre for Patient Safety & Service Quality*, 2008. Available online at `http://www1.imperial.ac.uk/resources/147E3ECA-1FD2-4AF8-BA34-08AAF1FBCAB1/psireportv3.pdf` (accessed November 24, 2020).

T Ueda and Y Hoshiai. Application of principal component analysis for parsimonious summarization of DEA inputs and/or outputs. *Journal of the Operations Research*, 40(4), 1997.

C Ugolini and L Nobilio. Vertical integration and contractual network in the cardiovascular sector: The experience of the Italian region Emilia Romagna. *International Journal of Integrated Care*, 3(2):e20, 2003.

A D Van den Abbeele, K M Krajewski, S H Tirumani, F M Fennessy, P J DiPiro, Q D Nguyen, G J Harris, H A Jacene, G Lefever, and N H Ramaiya. Cancer imaging at the crossroads of precision medicine: Perspective from an academic imaging department in a comprehensive cancer center. *Journal of the American College of Radiology*, 13(4):365–371, 2016.

T Van Puyenbroeck. On the output orientation of the Benefit-of-the-Doubt-Model. *Social Indicators Research*, 139(2):415–431, 2017.

B B Wang, T T H Wan, J Clement, and J Begun. Managed care, vertical integration strategies and hospital performance. *Health Care Management Science*, 4(3):181–191, 2001.

W B Weeks, D J Gottlieb, D J Nyweide, J M Sutherland, J Bynum, et al. Higher health care quality and bigger savings found at large multispecialty medical groups. *Health Affairs*, 29(5):991–997, 2010.

G Wei and J Wang. A comparative study of robust efficiency analysis and data envelopment analysis with imprecise data. *Expert Systems with Applications*, 81:28–38, 2017.

J West, G Johnson, and A K Jha. Trends in acquisitions of physician practices and subsequent clinical integration: A mixed methods study. *Journal of Evaluation in Clinical Practice*, 23(6):1444–1450, 2017.

S Yuan, F Wang, Y Zhao, and Y Liu. Assessing perceived quality of primary care under hospital-township health centre integration: A cross-sectional study in China. *The International Journal of Health Planning and Management*, 35(1):196–209, 2019.

J Zuckerman. Hip fracture. *The New England Journal of Medicine*, 334(23):1519–25, 1996.

A

Appendix A

A.1 Basic statistics

Figures A.1 to A.4 and Tables A.1 to A.3 represent some of the steps of the initial analysis over the collected set of studies, accomplished in Chapter 2.

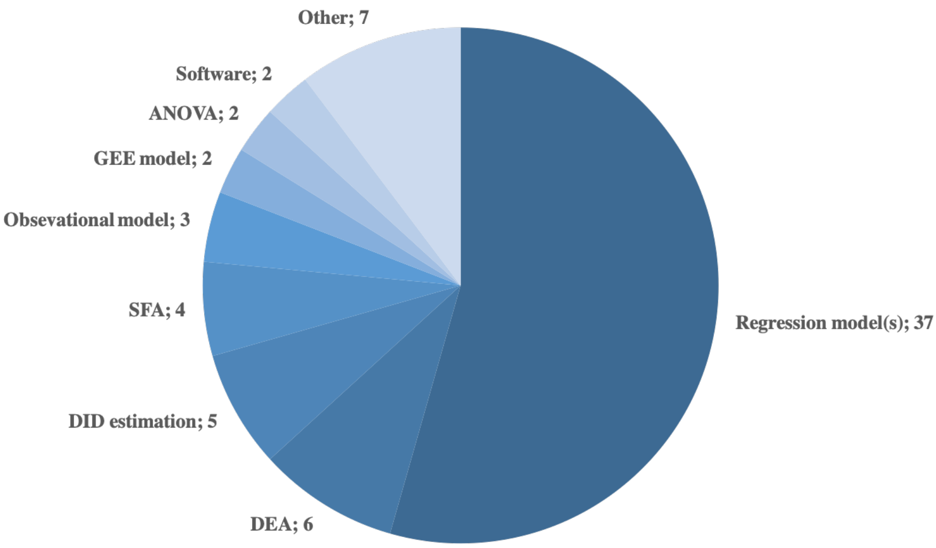
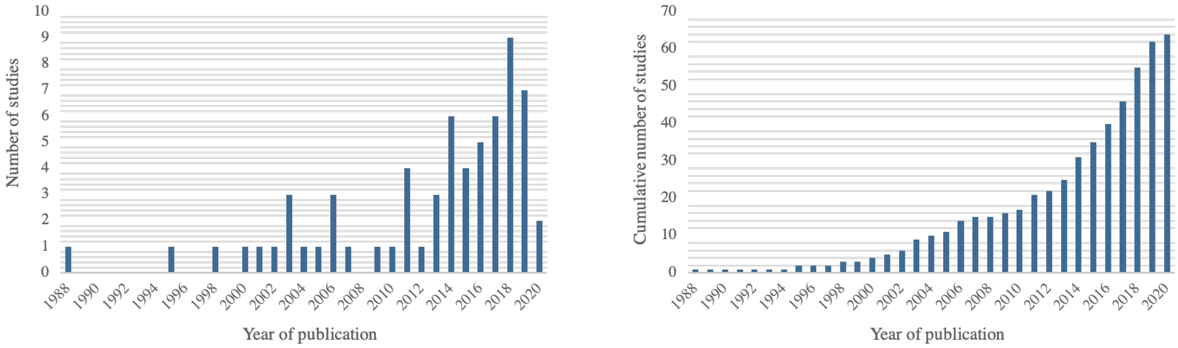


Figure A.1. Which methods are the most used when analyzing vertical integration?

Table A.1. Which methods are the most used when analyzing each one of the outcomes regarding vertical integration?

Costs			Prices of care			Efficiency			Quality of care			Access		
Regression model(s)	19	67%	Regression model(s)	3	60%	DEA	6	50%	Regression model(s)	23	63%	Regression model(s)	1	100%
DID estimation	3	11%	DID estimation	1	20%	SFA	4	34%	Observa-tional	3	8%			
Observa-tional	2	7%	SFA	1	20%	SEM	1	8%	GEE	2	5%			
GLMMs	1	4%				Regression model(s)	1	8%	DID es-timation	2	5%			
ANOVA	1	4%							Software	2	5%			
Other	2	7%							Other	5	14%			



(a) Number of studies per year.

(b) Cumulative number of studies.

Figure A.2. What is the trend regarding the number of studies published per year?

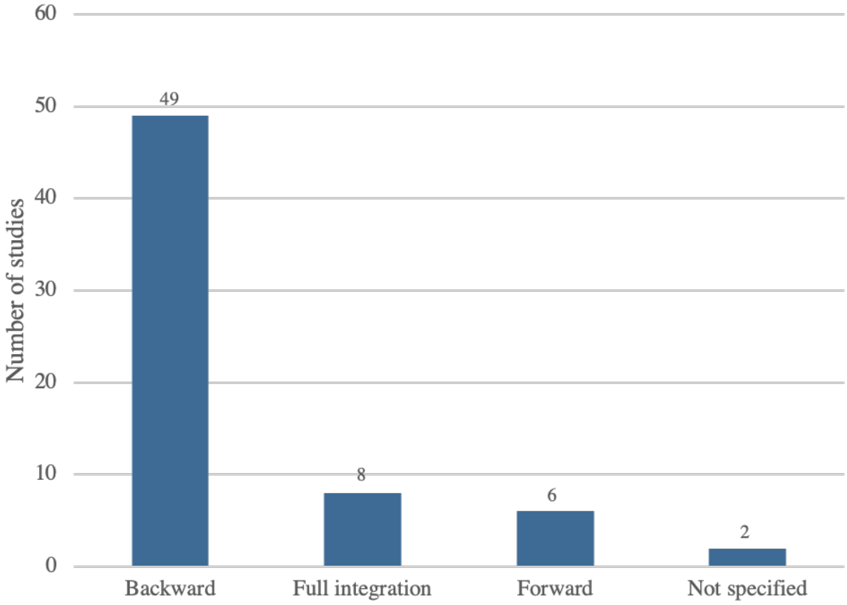


Figure A.3. Which type of vertical integration is the most studied?

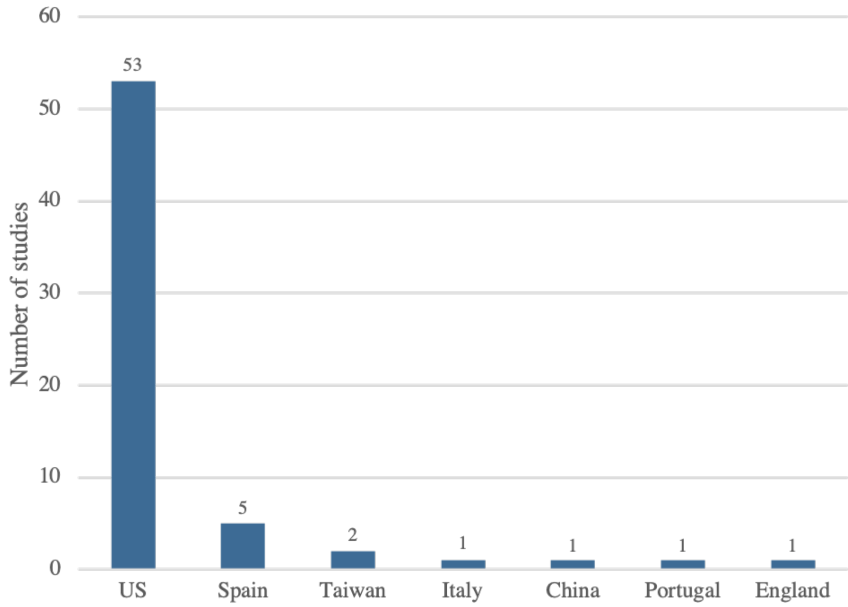


Figure A.4. Which countries do most studies belong to?

Table A.2. What are the journals in which there have been more studies published?

Journal	Number of studies	Total number of citations	Impact factor 2019-20	Scimago H-index
Health Services Research	11	514	2.010	116
Journal of Health Economics	6	595	3.660	117
Health Affairs	4	460	4.950	168
Medical Care	3	64	3.840	170
Inquiry	3	150	0.830	42
Health Services Management Research	2	12	0.970	31
NBER (National Bureau of Economic Research) Working Paper Series	2	NA	NA	NA
International Journal of Integrated Care	2	61	2.010	28
Healthcare Management Science	2	95	2.070	54
Healthcare Management Review	2	74	2.300	53
JAMA (Journal of the American Medical Association) Internal Medicine	2	199	14.570	329
Health Policy	2	108	2.260	85
Other	23	445	-	-

Table A.3. Which of the selected articles are most cited (Top 10)? Source: Google Scholar

Paper	Number of citations	Journal	Impact factor 2019-20	Scimago H-index
Baker et al. (2014)	234	Health Affairs	4.950	168
Cuellar and Gertler (2006)	223	Journal of Health Economics	3.660	117
Goes and Zhan (1995)	147	Health Services Research	2.010	116
Ciliberto and Dranove (2006)	122	Journal of Health Economics	3.660	117
Kralewski et al. (2000)	104	Health Services Research	2.010	116
Gillies et al. (2006)	101	Health Services Research	2.010	116
Carey (2003)	101	Inquiry	0.970	31
Neprash et al. (2015)	98	JAMA Internal Medicine	14.570	329
McWilliams et al. (2013)	95	JAMA Internal Medicine	14.570	329
Casalino et al. (2014)	91	Health Affairs	4.950	168

A.2 Meta-analysis

Tables A.4 to A.7 represent statistically significant results, for the significance level of 5%, of the meta-analysis accomplished in Chapter 2.

Table A.4. Methods vs. Type of vertical integration.

	Backward	Forward	Full	Not specified
Regression model	X[a]			
DEA				
DID				X[b]
SFA				
Other	X[c]			

[a] p-value = 0.042 and Phi = 0.280; [b] p-value = 0.021 and Phi = 0.318; [c] p-value = 0.009 and Phi = -0.357.

Table A.5. Type of vertical integration vs. Costs

	Costs increases	Costs decreases	Costs unchanged
Backward model	X[a]	X[b]	
Forward		X[c]	
Full		X[d]	
Not specified			

[a] p-value = 0.027 and Phi = 0.426; [b] p-value = 0.023 and Phi = -0.438; [c] p-value = 0.033 and Phi = 0.411;
[d] p-value = 0.033 and Phi = 0.411.

Table A.6. Methods vs. Costs

	Costs increases	Costs decreases	Costs unchanged
Regression model			
DEA			
DID			
SFA			
Other	X[a]	X[b]	

[a] p-value = 0.003 and Phi = -0.575; [b] p-value = 0.008; Phi = 0.509.

Table A.7. Methods vs. Efficiency

	Efficiency increases	Efficiency decreases	Efficiency unchanged
Regression model			
DEA	X[a]		
DID			
SFA			X[b]
Other			

[a] p-value = 0.038 and Phi = 0.655; [b] p-value = 0.016; Phi = 0.764.

A.3 Impact of vertical integration on efficiency

Table A.8 refers to the explicit analysis of the efficiency-related studies, accomplished in Chapter 2. Table A.9 presents the analysis of the papers that study the length of stay, also referring to Chapter 2.

Table A.8. Explicit analysis of the efficiency-related studies included in the sample.

Study	Methodology	Type of efficiency	Inputs	Outputs
		Increases		
Leleu et al. (2017)	DEA	TE	CMI; FTE RN + LPN; other trainees; FTE other; beds; post admission days.	Nº of admissions (no surgical); nº of inpatient surgeries; nº of outpatient surgeries; nº of ED visits; nº of other visits; mortality rate; readmission rate.
Caballer-Tarazona and Vivas-Consuelo (2016)	DEA	TE	HR costs and others; nº of beds; nº of ORs.	Adjusted surgical patients; overall score in the Management Agreements; adjusted admissions; adjusted outpatients.
Cho et al. (2014)	DEA	TE	Nº of beds; service mix; FTE employees; non-labor expenses.	CMI adjusted admissions; outpatient visits.
Rosko et al. (2007)	SFA	TE	Prices	Outpatient visits; inpatient admissions.
Chu et al. (2003)	DEA	TE	Personnel, medicine and depreciation costs.	Total revenue.
Chu et al. (2002)	DEA	TE	Physicians; nurses; ancillary labor; beds.	Ambulatory and emergency visits; inpatient days; inpatient visits.
Wang et al. (2001)	SEM	Efficiency of production	NA[a]	NA[a]
		Decreases		
Chunn et al. (2018)	Regression models	Efficiency of production	NA[b]	NA[b]
		Unchanged		
Comendeiro-Maaløe et al. (2019)	SFA	TE	Beds; FTE physicians; FTE nursing staff'.	Discharges weighted by DRG; outpatient activity.
Alonso et al. (2014)	DEA	TE	Nº of beds, nº of FTE physicians; nº of FTE nursing staff.	Nº of discharges; nº of outpatient visits.
Cuellar and Gertler (2006)	SFA	TE	Prices	Nº of hospital admissions; length of stay; CMI; nº of outpatient visits.
Carey (2003)	SFA	TE	Average annual salary; beds.	Adjusted admissions; adjusted patient days; CMI.

[a] The authors use indicators of productivity: adjusted admissions per bed and adjusted admissions per full time employees.

[b] The authors use a clinical productivity measure, expressed in work relative value units.

Table A.9. Length of stay: An indicator of efficiency.

Length of stay		
Better (n = 1)	Worst (n = 2)	Unchanged (n = 3)
David et al. (2011).	Gupta et al. (2019); Konetzka et al. (2018).	Henke et al. (2018); Li et al. (2018); Scott et al. (2017).

A.4 Impact of vertical integration on quality of care

Table A.10 presents the overview of the main quality measures analyzed in the sample of studies, referring to Chapter 2.

Table A.10. Overview of the main quality measures analyzed in the sample of studies.

Mortality		
Better (n = 4)	Worst (n = 1)	Unchanged (n = 5)
Henke et al. (2018); West et al. (2017); Rhoads et al. (2015); Ugolini and Nobilio (2003).	Chukmaitov et al. (2015).	Li et al. (2018); Scott et al. (2017); Chukmaitov et al. (2009); Huckman (2005); Madison (2004).
Readmissions rate		
Better (n = 5)	Worst (n = 1)	Unchanged (n = 7)
Gupta et al. (2019); Lopes et al. (2017); West et al. (2017); Al-Amin (2016); David et al. (2011).	McWilliams et al. (2013).	Ho et al. (2019); Pesko et al. (2018); Konetzka et al. (2018); Casalino et al. (2018); Li et al. (2018); Scott et al. (2017); Madison (2004).
Process of care measures		
Better (n = 6)	Worst (n = 0)	Unchanged (n = 3)
Short and Ho (2019); Herrel et al. (2017); Falces et al. (2011); Leibert (2011); Gillies et al. (2006); Brickman et al. (1998).	—	Rossiter (2018); McWilliams et al. (2013); Curry et al. (2013).