HOLLY BEA MERELIE

MEASURING THE EFFECT OF DEPRIVATION ON PRIMARY HEALTH CARE PERFORMANCE IN THE NHS USING DATA ENVELOPMENT ANALYSIS AND MALMQUIST INDICES

UNIVERSIDADE DO ALGARVE
Faculdade de Economia
2023

HOLLY BEA MERELIE

MEASURING THE EFFECT OF DEPRIVATION ON PRIMARY HEALTH CARE PERFORMANCE IN THE NHS USING DATA ENVELOPMENT ANALYSIS AND MALMQUIST INDICES

Mestrado em Gestão de Unidades de Saúde

Dissertação sob orientação:
Professora Doutora Carla Alexandra E. Filipe Amado
Professor Doutor Sérgio Pereira dos Santos

UNIVERSIDADE DO ALGARVE
Faculdade de Economia
2023

MEASURING THE EFFECT OF DEPRIVATION ON PRIMARY HEALTH CARE PERFORMANCE IN THE NHS USING DATA ENVELOPMENT ANALYSIS AND MALMQUIST INDICES

Declaração de autoria de trabalho

Declaro ser a autora deste trabalho, que é original e inédito. Autores e trabalhos consultados estão devidamente citados no texto e constam da listagem de referências incluída.

Holly Bea Merelie

____

ACKNOWLEDGEMENTS

I express an enormous amount of gratitude…

To my family, who kept on pushing me to finish this dissertation and to reach for my dreams, even when times get tough.

To my boyfriend, that put up with my terrible mood swings when I couldn't find a solution to problems in this thesis and continuously cheered me on.

To Professora Doutora Carla Amado and Professor Doutor Sérgio Santos, my advisors, for all their time, availability, solicitude and dedication throughout this project.

To my friends, that got used to me declining invites for social events to work on this dissertation.

To the friendly people that gave me an encouraging smile when walking passed me sitting on the lawn of the lovely Southampton Common, while I was reading a great number of articles and typing away on my laptop.

Index

Index of Figures

Index of Tables

List of Abbreviations

BAME	Black, Asian and Minority Ethnic
CCG	Clinical Commissioning Group
COPD	Chronic Obstructive Pulmonary Disease
DEA	Data Envelopment Analysis
DMU	Decision Making Unit
EU	European Union
GP	General Practitioner
IMD	Index of Multiple Deprivation
IOD2019	Indices of Deprivation 2019
LE	Life Expectancy
MI	Malmquist Indices
NHS	National Health Service
OECD	Organization for Economic Co-operation and Development
PHC	Primary Health Care
SMR	Standardized Mortality Ratio
UK	United Kingdom
WHO	World Health Organization

Resumo

As desigualdades em saúde são definidas como diferenças evitáveis e injustas que ocorrem entre grupos de pessoas, sendo também importante considerar que este termo faz referência aos cuidados de saúde recebidos e à oportunidade dos indivíduos de adotarem um estilo de vida saudável (The King's Fund, 2022). Desta forma, as desigualdades em saúde envolvem: o estado de saúde, o acesso aos cuidados de saúde, a qualidade e a experiência dos cuidados, riscos comportamentais e determinantes de saúde mais alargados (The King's Fund, 2022). As desigualdades em saúde deverão ser uma preocupação central para todos os países e respetivos governos, como referido na Declaração de Alma-Ata em 1978 (WHO, 1978). Esta declaração enfatiza que os cuidados de saúde primários são um fator chave para colmatar estas desigualdades (WHO, 1978).

Os cuidados de saúde primários deverão ser o primeiro ponto de contacto com o sistema de saúde, pelo que apresentam elevada responsabilidade em providenciar cuidados de saúde efetivos com uma utilização eficiente dos recursos. Os cuidados de saúde primários do Serviço Nacional de Saúde Inglês (NHS) enfrentam diversos desafios, desde uma população em expansão com utentes cada vez mais idosos, com multi-morbilidades (necessidades de saúde mais complexas), mudanças na complexidade das consultas, utentes com reduzida satisfação face aos cuidados de saúde, até à dificuldade de recrutamento e retenção de profissionais de saúde (Neri, Cubi-Molla e Cookson, 2022). Assim, avaliar o desempenho e implementar medidas de melhoria dos cuidados de saúde primários é de elevada importância.

A técnica de *Data Envelopment Analysis* (DEA) tem sido frequentemente utilizada para avaliar e comparar o desempenho de sistemas de saúde e, mais especificamente, dos cuidados de saúde primários. Esta técnica tem-se demonstrado eficaz em produzir informação útil para os decisores políticos e gestores dos serviços de saúde para justificar e fundamentar as tomadas de decisão. Este facto determinou que fosse esta a técnica aplicada neste estudo para avaliar o desempenho dos cuidados de saúde primários do NHS, quanto à equidade na alocação de recursos, à eficiência dos serviços prestados e à efetividade desses serviços.

Recolheram-se dados para 191 *Clinical Commissioning Groups* (CCGs), destes, 3 foram excluídos por falta de dados para o ano completo em estudo e 9 foram excluídos por terem sido identificados como outliers. Três modelos da técnica DEA foram

analisados, o primeiro com o intuito de avaliar a equidade na distribuição de recursos humanos face à lista de utentes registados; o segundo, avaliou a eficiência dos cuidados de saúde primários, tendo em conta os profissionais de saúde presentes e a atividade por eles produzida; e o terceiro modelo avaliou a efetividade dos cuidados em reduzir admissões hospitalares e em maximizar a satisfação dos utentes. Para cada um dos modelos os resultados permitiram identificar os grupos com melhores níveis de desempenho e os grupos com piores níveis. Esta técnica permitiu também identificar unidades de referência (benchmarks) para as unidades ineficientes, pelo que seria vantajoso criar oportunidades de aprendizagem, como entrevistas, discussões de grupo e formações entre estas unidades de modo a conseguirem melhorar os seus resultados de desempenho.

Após a avaliação do desempenho dos CCGs, foram calculados os Indices de Malmquist desenvolvidos por Camanho e Dyson (2006), para contrastar o desempenho de grupos. Neste estudo, este método foi aplicado com o objetivo de avaliar o efeito da privação socioeconómica no desempenho dos cuidados de saúde primários. Para tal, os 179 CCGs foram divididos em dois grupos, o primeiro grupo apresenta maior nível de privação no Índice de Privação Múltipla (IMD) e o segundo apresenta menor nível de privação. O Índice de Privação Múltipla é a medida oficial para medir privação relativa na Inglaterra, pelo que inclui 7 domínios: privação de rendimentos; privação de emprego; privação de educação, competências e formação; privação de saúde e incapacidade; crime; barreiras à habitação e serviços; e privação do ambiente onde se vive (Ministry of Housing, Communities and Local Government, 2019).

Os Índices de Malmquist foram calculados para os três modelos aplicados na técnica de DEA, deste modo obteve-se o efeito da privação na equidade de afetação de recursos, na eficiência e na efetividade dos cuidados de saúde primários. No que respeita à equidade, os resultados demonstram a existência de discriminação positiva para combater a privação. Em específico, verifica-se maior afetação de recursos ao grupo de CCGs com maior nível de privação quando comparado com o grupo com menor nível de privação. Relativamente ao efeito da privação na eficiência dos cuidados de saúde primários do NHS, os resultados evidenciam que as políticas de saúde em vigor em 2019 estavam a ser eficazes na redução de desigualdades em saúde derivadas da privação, demonstrando que o nivel de eficiência e efetividade dos cuidados de saúde primários no grupo com maior privação era superior ou igual ao grupo com menor privação.

Os resultados obtidos neste estudo, estão de acordo com alguns resultados obtidos por outros autores. Assim, esta informação deverá ser utilizada pelos decisores políticos para fundamentar o processo de tomada de decisão. Deste modo, é possível perceber que estas duas técnicas permitem avaliar o passado e planear melhorias para o futuro dos cuidados.

Palavras-chave: Cuidados de Saúde Primários; Desigualdades; Serviço Nacional de Saúde; Índice de Malmquist; DEA.

Abstract

Life expectancy is cut short in areas that experience highest levels of deprivation in England. For this reason, policymakers and managers of primary health care in the National Health Service (NHS) should concentrate efforts on reducing these health inequalities by producing effective care with efficient spending.

The main objective of this study is to measure the effect of deprivation on primary health care performance. To reach this goal, two techniques were applied: Data Envelopment Analysis measured performance levels for the 179 Clinical Commissioning Groups (CCGs) studied, then Malmquist Indices calculated the effect of deprivation on their performance level.

The performance results of this study identify variation throughout the CCGs, suggesting that in some CCGs there is significant room for improvement. The Malmquist Indices demonstrate that there is evidence of positive discrimination in terms of resource allocation in favour of CCGs facing higher levels of deprivation. In specific, CCGs with higher levels of deprivation receive, on average, higher levels of resources than those with lower levels of deprivation. Furthermore, service efficiency and effectiveness were equal or superior in the group with higher deprivation level compared to the group with lower deprivation. Ultimately, this reflects that health and social policies in place in 2019 were successful in addressing health inequalities due to deprivation.

This study demonstrates the potential of these techniques in studying the effects of policies on health inequalities.

Keywords: Primary Health Care; Deprivation; National Health Service; Data Envelopment Analysis; Malmquist Index.

1. Introduction

Health inequalities can be defined as avoidable and unfair differences in health between groups of people, additionally, the term refers to the care people receive and the opportunity these have to lead a healthy life (The King's Fund, 2022). As per The King's Fund (2022), health inequalities therefore involve: health status, access to care, quality and experience of care, behavioural risks to health and wider determinants of health (e.g. housing).

In 1978, the International Conference on primary health care (PHC) held in Alma-Ata, emphasized that the inequality in health status between people in developed and developing countries as well as within countries was unacceptable, reason why it should be a priority to all countries (WHO, 1978). In 2019, the OECD studied inequalities in health and health systems among 33 OECD and European Union conuntries, by looking into socio-economic differences in the exposure to risk factors to health, health status, the utilisation of health services, unmet health care needs and coverage (OECD, 2019). Some conclusions of this study were that poor health behaviour tends to be more prevalent among the disadvantaged; the least educated are more likely to be in poor health; people with low income are less likely to see a doctor; access to preventive services is concentrared among the higher income groups; unmet needs for care are concentrated among lower income groups (OECD, 2019).

Since the 1970s the Ministry of Housing, Communities and Local Government (and its predecessors) have measured deprivation in England (Ministry of Housing, Communities and Local Government, 2019). The Index of Multiple Deprivation (IMD) is the official measure of relative deprivation in England, defining deprivation to include a wide range of an individual’s living conditions, being the most recent report the English Indices of Deprivation 2019 (IoD2019) (Ministry of Housing, Communities and Local Government, 2019). In England, data from 2017-2019 show that there is a systematic relationship between deprivation and life expectancy (LE), where there is a gap in life expectancy of almost 8 years between women who live in the 10% least-deprived areas (higher LE) compared to those in the 10% most-deprived areas (lower LE) (The King's Fund, 2022). This gap is even wider among men – about 9.4 years between the least and most deprived areas (The King's Fund, 2022).

The Alma-Ata Declaration of 1978 points the responsibility to the governments for the health of their people, which can be fulfilled by the provision of adequate health and social measures, identifying, therefore, primary health care as the key to reaching this goal (WHO, 1978). According to the WHO (1978, 3) primary health care "forms an integral part of both the country's health system, of which it is the central function and main focus, and of the overall social and economic development of the community. It is the first level of contact of individuals, the family and community with the national health system bringing health care as close as possible to where people live and work, and constitutes the first element of a continuing health care process". Primary health care within the NHS in England, follows these principles, by providing a wide range of services to people with acute and chronic conditions, throughout their life cycle.

Primary health care in the NHS is facing multiple challenges from a growing population with increasingly older and multi-morbid patients (with more complex health needs), changes in the complexity of consultations, poor patient satisfaction, to the difficulty in recruiting and retaining workforce (Neri, et al., 2022). In addition to these challenges, we also have to consider the limited resources allocated to the health system, thus it is paramount to deliver effective care with efficient spending.

The pursuit of efficiency is of central concern to health policy-makers and managers (Cylus, Papanicolas and Smith, 2016). Inefficient care may lead to: unnecessarily poor outcomes for the patients, in terms of their health improvement and their satisfaction with the health system; denying treatment and health improvement to patients who would have received it if resources had been better used; the need to divert resources from other productive sectors of the economy (public services) (Cylus, et al., 2016). Furthermore, the lack of evidence that a system is performing efficiently can damage confidence in the institutions and reduce the willingness of governments and their citizens to pay for the health system (through their taxes) (Cylus, et al., 2016).

Measuring efficiency in primary health care has been less researched than hospital settings (Neri, et al., 2022). In hospital settings, volume indicators (e.g. patients treated) are commonly used to measure output, whereas in primary health care measuring output is more complex, as it provides a generalist, holistic and long-term service to individuals and families (Neri, et al., 2022).

Although some authors identify primary health care as a complex setting to measure performance, Campbell, Braspenning, Hutchinson and Marshall (2002) provide some insight into an ideal performance measure. This measure should have certain characteristics that can be checked by answering a set of questions such as "acceptability: is the indicator acceptable to both those being assessed and those undertaking the assessment? Feasibility: are valid, reliable, and consistent data available and collectable, albeit contained within medical records, health authority datasets or on videotaped consultations? Reliability: minimal measurement error, organisations, or practitioners compared with similar organisations or practitioners (comparability), reproducible findings when administered by different raters (inter-rater reliability). Sensitivity to change: does the indicator have the capacity to detect changes in quality of care? Predictive validity: does the indicator have the capacity for predicting quality of care outcomes?" (Campbell, et al., 2002, 359). Considering these characteristics and considering the research from other authors, Data Envelopment Analysis (DEA) is the chosen technique to measure performance in this research.

Data Envelopment Analysis developed by Charnes, Cooper and Rhodes (1978) is a non-parametric programming technique that was developed to measure efficiency of not-for-profit entities in public programs (Bankers, Charnes and Cooper, 1984). DEA measures relative efficiency of homogenous Decision-Making Units (DMUs) (Bankers, Charnes and Cooper, 1984). It employs mathematical programming to control and evaluate past accomplishments and to aid in planning future activities (Bankers, Charnes and Cooper, 1984). This technique will identify benchmarks, which are efficient DMUs that can provide valuable teaching to inefficient units. This aspect of DEA is particularly important to improve PHC, as best practices are identified and can be adopted by less efficient units. A recent systematic review carried out by Neri, et al. (2022) showed that DEA is the most frequently utilised approach to measure performance in PHC.

The second technique used in this research is the Malmquist Index (MI) developed by Camanho and Dyson (2006). Camanho and Dyson (2006) applied the Data Envelopment Analysis technique with Malmquist Indices to measure group performance. This method is an adaptation to the Malmquist Index developed by Fare, et al. (1994) where it no longer measures productivity change over time but allows a cross-sectional performance comparison of groups of DMUs operating in different conditions at one moment in time (Camanho and Dyson, 2006).

The aim of this research is to assess the effect of deprivation on primary health care performance by utilising DEA with Malmquist Indices. Specifically, we aim to: 1. Identify Clinical Commissioning Groups that demonstrate resource equity, service efficiency and service effectiveness; 2. Identify benchmarks that may be used to improve the performance of underperforming CCGs; 3. Quantify the effect of deprivation on CCGs' performance in terms of resource equity, service efficiency and service effectiveness. Ultimately, these objectives intend to provide useful information for policy-makers and managers within the NHS.

To achieve the goals set out previously, this dissertation is organised as follows: section 2 provides an initial description of how PHC is structured within the NHS and some health inequalities that are present, followed by a literature review of relevant studies that have applied DEA or MI to PHC; section 3 presents the DEA and MI models, then the results of these models are discussed; finally section 4 concludes this dissertation identifying key messages, limitations and suggestions for further research.

2. Literature Review

2.1.Primary Health Care in the NHS and Health Inequalities

The National Health Service (NHS) in the United Kingdom (UK) was founded in 1948 and is currently composed by many different organisations (NHS, 2013a). It has seven key principles that guide its services: "The NHS provides a comprehensive service, available to all; access to NHS services based on clinical need, not an individual's ability to pay; the NHS aspires to the highest standards of excellence and professionalism; the patient will be at the heart of everything the NHS does; the NHS works across organisational boundaries; the NHS is committed to providing best value for taxpayers' money; the NHS is accountable to the public, communities and patients that it serves" (NHS, 2013b, 3). The NHS has six values that should underlie all its activities: "working together for patients, respect and dignity, commitment to quality of care, compassion, improving lives, everyone counts" (NHS, 2013b, 5). The set of principles and values connect the people it serves and the staff who work for it (NHS, 2013b).

The NHS is funded by taxation with a fixed budget to spend on its services, which raises the challenge of using this budget in a way that produces the best possible outcomes for the people it serves (NHS, 2013b). The planning and purchasing of NHS services is carried out by commissioners, who assess population needs and buy services that should meet these needs in an affordable and high-quality manner (NHS, 2013b). Due to the complexity and scale of the health care system, the planning and commissioning of health care happens at a population level, however, there are some services (such as specialised services for rare diseases) that are more appropriately commissioned at a national level by NHS England (the NHS Commissioning Board) (NHS, 2013b). Primary health care is another example of service that is directly commissioned by NHS England (NHS, 2013b).

Clinical Commissioning Groups were established as part of the Health and Social Care Act in 2012 replacing Primary Care Trusts (PCT). The Clinical Commissioning Groups' (CCG) governing bodies have a General Practitioner, a Nurse, secondary representatives, and a minimum of two 'lay' members who are not NHS professionals (NHS, 2013b). The main duty of these groups is to commission, this is a continual process of planning, agreeing, and monitoring services, it ranges from assessing health-needs, clinically based design of patient pathways, to service specification and contract negotiation or procurement, while continuously assessing quality (NHS, 2023b). The CCGs can commission "planned hospital care, rehabilitative care, urgent and emergency care (including out-of-hours and accident and emergency services), most community health services, mental health and learning disability services" (NHS, 2013b, 6).

The NHS defines that primary health care "focuses on the treatment of minor injuries and illnesses and deals with minor surgery and the on-going management of chronic conditions. It is the first point of contact most people have with the NHS, and is delivered by a wide range of professionals, including family doctors (General Practitioners), nurses, dentists, pharmacists, and opticians" (Department of Health, 2015, 11). Thus, general practice, community pharmacy, dental and optometry services are the services provided by primary health care in the NHS (Department of Health, 2015).

NHS England allocate funds on a 'weighted capitation' basis, budgets are calculated based on certain criteria (NHS, 2019a). The funding attributed to each CCG for primary medical care (the focus of this study) is calculated based on two main aspects: need and Market Forces Factor (cost) (NHS, 2019a). Need is estimated based on factors like population age, index of multiple deprivation (MDI) and health professional workload (NHS, 2019a). To cover possible health inequalities or unmet needs the Standardised Mortality Ratio for those under 75 years of age (SMR<75) is also included in the weighted capitation (NHS, 2019a). Cost is calculated using the Market Forces Factor, which adjusts the unavoidable differences in costs between areas due to their geographical location alone (NHS, 2019a).

As mentioned previously, primary health care should be the first point of access to the health care system, therefore it has increased responsibility for delivering effective care and efficient spending (Neri, et al., 2022). In view of this, a proactive effort to improve efficiency and productivity within primary health care in the NHS is needed to certify that the government's investment meets the needs of a growing population with constantly more complex and expensive health care needs (Neri, et al., 2022). The complexity of primary health care consultations is reflected in: the increasing duration of the consultations; the number of results dealt with by GPs; and the number of patients with complex and potentially hazardous prescription regimes (Neri, et al., 2022). On the other hand, the number of GPs per capita has fallen since 2009 and workforce recrutiment and retention are a significant challenge (Neri, et al., 2022). The pressures faced by primary health care result in declining GP satisfaction and patient satisfaction with the access to services (Neri, et al., 2022).

Health inequalities should be a central concern of all countries, as defined in the Alma Ata declaration in 1978: "The existing gross inequality in the health status of the people particularly between developed and developing countries as well as within countries is politically, socially and economically unacceptable and is, therefore, of common concern to all countries” (WHO, 1978, 2).

The King’s Fund (2022) defines health inequalities as unfair and avoidable differences in health between different groups of people within society, these may arise due to the conditions in which we are born, grow up, live, work and get old. These conditions affect opportunities for good health, they influence our mental health, physical health and overall wellbeing (The King's Fund, 2022). Furthermore, health inequalities may be analysed in four types of factors: socio-economic factors; geography; the individual’s characteristics such as sex, ethnicity, or disability; and socially excluded groups for example, people that are homeless (The King's Fund, 2022).

Wilkinson and Pickett (2006) reviewed 168 analyses in 155 international papers that report findings on the association between income distribution and population health. Their results show that 87 of these analyses were wholly supportive of the association between greater income inequality and poorer population health, 44 were partially supportive (not all of the associations reported were significantly positive associations) and 37 were unsupportive (no statistically significant positive associations) (Wilkinson and Pickett, 2006). These authors conclude that income distribution is related to health, as it serves as a measure of the scale of social class differences in society (Wilkinson and Pickett, 2006). They explain that inequality should not be seen as a new risk factor for health, but that it informs about already widely recognised health effects of socio-economic status and class, adding that a more unequal society suffers widespread health disadvantage due to status competition and class differentiation (Wilkinson and Pickett, 2006).

Although it is very frequent to look at inequality in terms of income and socio-economic factors, other factors such as disability, sex and ethnicity have also been shown to generate health inequalities. Ubido, Huntingdon and Warburton (2002) studied inequalities in access to health care faced by women who are deaf in Cheshire, UK, with their results showing that 49% would be more likely to use health services if help for deaf women was in place. Some of the issues identified in the questionnaires were difficulty in making appointments or obtaining repeat prescriptions – had to go in personally or had to get someone else to phone the practice; problems in the waiting rooms related to not hearing their name called and consequently missing their turn; communication – only 7% of the deaf women usually fully understood what was being said, and even written instructions were not easy to understand; lack of awareness – health service staff demonstrate poor attitude and insensitivity; language – some deaf people do not acquire full comprehension of Sign Language and some medical terms are unfamiliar in Sign Language; issues relating to hearing tests and aids – uncertainty about when to get them changed, uncomfortable aids, unsuitable rooms for hearing tests, and difficulty to get to hearing aid clinics/hospitals; health information – they faced a general problem of lack of information in matters such as sex education, contraception and childbirth (Ubido, et al., 2002).

The NHS was founded to provide universal access to health care, however other important factors influence health, such as the social and economic environment and the lifestyle choices made by individuals and their families (NHS, 2019c). Life expectancy continues to improve for the most affluent 10% of the population, despite this, it is either stagnant or falling for the most deprived 10% (NHS, 2019c). According to NHS (2019c) premature mortality is twice as high in Blackpool, the most deprived area of the country, in comparison to least deprived areas; women in most deprived areas spend 34% of their lives with poor health, compared to 17% in wealthiest areas; multimorbidity is higher in deprived areas; Black Asian Minority Ethnic (BAME) communities are at higher risk of poor health and early death; adults with learning disabilities die, on average, 16 years earlier than the general population; people with mental health illnesses die 15-20 years earlier than those without; people affected by homelessness die, on average, 30 years earlier than the general population.

Due to the previously mentioned findings by NHS (2019c), the NHS Long Term Plan in 2019, concentrated on reducing health inequalities and unwarranted variations in care. To achieve this, the NHS set out strategies such as: NHS England will continue to target higher share of funding towards geographies with high health inequalities – this includes more accurate assessment of health need and ensuring the allocations formula is more responsive to the greatest health inequalities and unmet need; will set out specific and measurable goals for narrowing inequalities that will be assessed in 2023/24 – such as all CCGs ensuring that screening and vaccination programmes are designed to narrow health inequalities; by 2024 75% of women from BAME communities, and other deprived groups, will receive continuity of care throughout pregnancy, labour and postnatal period; by 2020/2021 at least 280 000 people living with severe mental health problems have their physical needs met; invest up to an extra £30 million on meeting the needs of rough sleepers; identify and support carers particularly those from vulnerable communities; and more (NHS, 2019c).

2.2. Data Envelopment Analysis (DEA) applied to Primary Health Care

Farrel (1957) was the first author to develop an empirical method to calculate relative productive efficiency of a set of decision-making units. This author defined efficiency in two aspects: technical efficiency which reflects success in producing the maximum amount of outputs with the set inputs, and price efficiency where an optimal ratio of inputs, according to their prices, is utilised (Farrel, 1957).

This method was further developed by Charnes, Cooper and Rhodes (1978). Data Envelopment Analysis, developed by Charnes, Cooper and Rhodes (1978) is a non-parametric programming technique that was developed to measure efficiency of non-for-profit entities in public programs. Data Envelopment Analysis employs mathematical programming to control and evaluate past accomplishments and to aid in planning future activities (Banker, et al., 1984).

Charnes, Cooper and Rhodes (1978) introduced a ratio definition of efficiency, known as CCR ratio, which generalizes the single-output to single-input classic ratio to multiple outputs and inputs without requiring preassigned weights (Banker, et al., 1984). This method assesses the performance of homogenous decision-making units with common inputs (resources) and outputs (products/services) (Charnes, et al., 1978). The CCR model identifies technical inefficiencies, which are failures to achieve the best possible output levels and/or usage of excessive amounts of inputs (Banker, et al., 1984).

In the Constant Returns to Scale (CRS) model developed by Charnes, Cooper and Rhodes (1978), also known as the CCR model, the efficiency rate of each DMU is the maximum of a ratio of weighted outputs to weighted inputs subject to the condition that similar ratios for every decision-making unit be less than or equal to unity. Mathematically demonstrated as:

$$E_0 = \max h_0 = \frac{\sum_{r=1}^{s} u_r y_{r0}}{\sum_{i=1}^{m} v_i \, x_{i0}}$$

subject to

$$\frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \leq 1 \text{ for j=1, \ldots , n}$$

$$u_r, v_i > 0$$

E_0 represents the efficiency rate of DMU$_0$, y_{rj}, x_{ij} represent the known outputs and inputs, of the jth DMU, respectively, and the $u_r, v_i > 0$ are the variables weights to be determined by the solution of the problem. $\frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \leq 1$ reflects that the virtual output and input ratios of every DMU must be less or equal to unity.

The linear programming for assessing technical efficiency with output orientation under Constant Returns to Scale developed by Charnes, Cooper and Rhodes (1978) can be demonstrated as (Thanassoulis, 2001):

$$E_0 = \text{Max } h_{j0} + \varepsilon \left[\sum_{i=1}^{m} I_i + \sum_{r=1}^{s} O_r\right]$$

Subject to

$$\sum_{j=1}^{N} \alpha_j x_{ij} = x_{ij_0} - I_i \quad \text{i= 1...m}$$

$$\sum_{j=1}^{N} \alpha_j y_{rj} = O_r + h_{j0}\, y_{rj_0} \quad \text{r= 1...s}$$

$$\alpha_j \geq 0, j = 1 \dots N, I_i, O_r \geq 0 \ \forall\ i \ and \ r, h_{j0}\ free$$

ε is a non-Archimedean infinitesimal.

Where N DMUs (j= 1… N) use m inputs to secure s outputs. The model gives priority to maximizing h_{j0}, it identifies a point which offers output levels reflecting the maximum feasible radial expansion of the output levels without raising its input levels (Thanassoulis, 2001). The variables I_i and O_r are slack variables that represent feasible input reductions and/or output augmentations from the production possibility set point that corresponds to the maximum radial expansion of the output levels of DMU j_0 (Thanassoulis, 2001).

The linear programming for assessing technical efficiency with output orientation under Variable Returns to Scale developed by Bankers, Charnes and Cooper (1984) can be demonstrated as (Thanassoulis, 2001):

$$E_0 = \text{Max } z + \varepsilon \left[\sum_{i=1}^{m} I_i + \sum_{r=1}^{s} O_r\right]$$

Subject to

$$\sum_{j=1}^{N} \alpha_j\, x_{ij} = x_{ij_0} - I_i \quad \text{i= 1…m}$$

$$\sum_{j=1}^{N} \alpha_j\, y_{rj} = O_r + z\, y_{rj_0} \quad \text{r= 1...s}$$

$$\sum_{j=1}^{N} \alpha_j = 1$$

$$\alpha_j \geq 0, j = 1 \dots N, I_i, O_r \geq 0 \ \forall\ i \ and \ r, \text{z free.}$$

From the envelopment linear programming models presented above, it is possible to formulate the multiplier version of these models to calculate the input and output weights.

Dyson, Allen, Camanho, Podinovski, Sarrico and Shale (2001) identified multiple pitfalls in the applications of the DEA technique, such as: lack of homogeneity of the units under assessment, the set of inputs/outputs used, measuring those variables and the weights attributed to them in the analysis. Therefore, Dyson, et al. (2001) developed protocols to overcome these issues. Some of these protocols are as follows: non-homogenous units may be compared if clustered into homogenous sets, whereas non-homogenous environments may be analysed if environmental variables are included; when analysing economies of scale, the units should be assessed for the effects of scale and the Variable Returns to Scale model should only be used when scale effects can be demonstrated; the number of variables (inputs and outputs) should follow the rule that the number of units should be at least two times the product of the number of inputs and the number of outputs; weight restrictions should reflect aims or values of the organization, should be fair and justified (Dyson, et al., 2001).

Data Envelopment Analysis has been vastly applied in the health care setting. Below, we present a literature review of some relevant applications of this technique to primary health care and to measure health inequalities.

Neri, et al. (2022) carried out a systematic literature review on different approaches to measure efficiency in primary care, in which they concluded that DEA was the most frequently used technique (31 out of 38 studies). Zakowska and Godycki-Cwirko (2020) also concluded that DEA is the most frequently used tool to analyse efficiency in PHC. Other techniques used were stochastic frontier analysis (SFA), Malmquist Index, methods based on cost functions and regression analysis (Neri, et al., 2022).

Appendix 1 systematises information regarding studies that provide theoretical support to this research paper. Each of these studies has applied DEA to health care and for each one the type and quantity of decision-making units (DMUs), the technological approach, the model orientation and the input and output variables used, are analysed.

These studies have been based in several countries such as the United Kingdom (Amado and Dyson, 2009; Salinas-Jiménez and Smith, 1996; Takundwa, Jowett, McLeod and Peñaloza-Ramos, 2017), Portugal (Amado and Santos, 2009; Camanho and Dyson, 2006; Ferreira, Marques and Nicola, 2013), Turkey (Cinaroglu, 2020), China (Huang, Luo, Wang , Cao , Wang, Bi , Huang and Yi, 2020; Zhang, Wang, Jiang and Wang, 2018), Czech Republic (Dlouhy, 2018), Greece (Oikonomou, Tountas, Mariolis, Souliotis, Athanasakis and Kyriopoulos, 2016), and Spain (Deidda, Lupiáñez-Villanueva, Codagnone and Maghiros, 2014). Some authors have carried out systematic reviews based on international studies that apply the DEA technique to measure primary health care performance (Zakowska and Godycki-Cwirko, 2020; Pelone, Kringos, Romaniello, Archibugi, Salsiri and Ricciardi, 2015; Neri, et al., 2022). Researchers have used regions, health centres, hospitals, district health authorities (DHA) and Clinical Commissioning Groups (CCG) as the Decision-Making Units in their analyses.

The most common objective of these papers has been to use DEA to measure performance, assessing efficiency, effectiveness and equity (Amado and Dyson, 2009; Amado and Santos, 2009; Cinaroglu, 2020; Ferreira, et al., 2013; Jordi, Pley, Jowett, Jaoude and Haghparast-Bidgoli, 2020; Oikonomou, et al., 2016; Salinas-Jiménez and Smith, 1996; Takundwa, et al., 2017). Others, have applied DEA to measure the efficiency of health care resource allocation (Dlouhy, 2018; Huang, et al., 2020; Zhang, et al., 2018). Deidda, et al. (2014) attempted to fill a gap in literature regarding the impact of information and communication technology on primary health care performance. Neri, et al. (2022), Pelone, et al. (2015), and Zakowska and Godycki-Cwirko (2020) intended to review methodological frameworks and the most common variables used in the DEA models.

Amado and Dyson (2009) initially carried out workshops to identify available data, followed by additional data collection through questionnaires applied to the stakeholders. Oikonomou, et al. (2016) also collected data through questionnaires applied to health centres. Conversely, most authors use secondary data from national or international open sources, such as the China Health Statistical Yearbook (Zhang, et al., 2018), Turkish Statistical Institute (Cinaroglu, 2020), WHO Global Health Observatory and Global Health Expenditure Database (Jordi, et al., 2020), NHS Digital (Takundwa, et al., 2017) and others. The systematic review carried out by Pelone, et al. (2015) corroborates this by concluding that most studies based the choice of variables on data availability instead of collecting new data to construct the best possible measure.

DEA is regularly applied in combination with other methods and indices, such as the Lorenz Curve, Gini Coefficient, Theil Index (Zhang, et al., 2018), Robin Hood Index (Dlouhy, 2018), and Simar-Wilson Bootstrapping (Jordi, et al., 2020). Although Camanho and Dyson (2006) did not study health care in their research, they used DEA and the Malmquist Index to measure group performance, which has proven to be a useful method when the objective is to contrast groups with different characteristics, methodology which will be applied in this study.

Zhang, et al. (2018) used the Lorenz Curve, Gini Coefficient, Theil Index and Health Resource Density Index to assess equity in demographic and geographical dimensions, then DEA and MI were used to measure efficiency and productivity. Dlouhy

(2018) emphasizes that the Gini Coefficient is the most popular measure of inequality but defends that the use of the Robin Hood Index (RHI) is more advantageous due to its clear practical interpretation. The Robin Hood Index, applied with the results from the DEA technique, measured the proportion of resources that would need to be reallocated from areas with above-average allocation to areas with below-average provision to reach equal distribution (Dlouhy, 2018).

Jordi, et al. (2020) used the Simar-Wilson bootstrapping method that is a two-stage DEA analysis that adjusts the standard DEA score by the amount of bias created by contextual factors, which allows the authors to investigate the factors associated with the estimated efficiency scores.

Camanho and Dyson (2006) used DEA and Malmquist Indices to measure group performance of bank branches in 4 Portuguese regions. This method enabled them to identify relative strengths and weaknesses of the different regions as well as to distinguish the inefficiencies due to poor management at individual branches from market and managerial policies defined centrally for the branches of a specific region (Camanho and Dyson, 2006). Due to its relevance for the objectives of this work, this methodological approach will be explained in more detail further on in this paper.

Concerning the returns to scale assumption, constant returns to scale (CRS), variable returns to scale (VRS) or both are used in the studies. Constant return to scales is not present if a proportional increase in inputs causes greater (or smaller) than proportional increase in outputs (Ferreira, et al., 2013). If it is believed that a unit is not operating under optimal scale, then the variable returns to scale approach is recommended (Ferreira, et al., 2013). By calculating both approaches it is possible to determine scale efficiency (Ferreira, et al., 2013). Most of the studies reviewed in this context apply a variable returns to scale approach to their studies, justifying that the units under analysis do not operate under an optimal scale (Huang, et al., 2020; Zhang, et al., 2018; Deidda, et al., 2014; Cinaroglu, 2019). Zakowska and Godycki-Cwirko (2020) also came to this conclusion in their systematic review. Albeit, Salinas-Jiménez and Smith (1996), Amado and Santos (2009) and Dlouhy (2018) used a constant return to scale approach defending that for the units under analysis there was no evidence of significant economies of scale. Some authors use both approaches, such as Ferreira, et al. (2013) and Oikonomou, et al. (2016).

Pelone, et al. (2015) explain in their systematic review that the choice for CRS or VRS depends on the context, level and perspective of analysis, and the technology linking inputs and outputs in the transformation process. Furthermore, if the DEA analysis is conducted from the point of view of the policymaker, then a CRS approach should be used and VRS is appropriate when the analysis uses a managerial point of view (Pelone, et al., 2015). Nevertheless, as defended by Ferreira, et al. (2013) the use of both approaches is advantageous as it allows the measurement of scale efficiency, which will provide a more robust assessment of primary health care performance. For this reason, in this research paper, both approaches will be applied.

When it comes to model orientation, it should be based on what managers are able to control – resources or outcomes (Pelone, et al., 2015). Zakowska and Godycki-Cwirko (2020) show in the results of their systematic review that input orientation is the most common orientation used in studies that apply DEA to Primary Health Care, adding that health care outputs are less controllable than the inputs. The systematic review by Pelone, et al. (2015) also concluded that most studies used an input orientation, only 3 out of 39 used both orientations and 9 out of 39 used an output orientation. Although some researchers apply an input orientation to their studies, in health care the focus should concentrate on increasing outputs instead of reducing inputs and costs (Zakowska and Godycki-Cwirko, 2020). Amado and Santos (2009), Jordi, et al. (2020), Huang, et al. (2020), Deidda, et al. (2014) and Oikonomou, et al. (2016) used an output orientation in their models, explaining that if there is evidence of waiting lists, it is important to maximize the services provided with the available resources.

The DEA technique can handle multiple inputs and outputs in the transformation process; however, its' results depend on the relevance of the chosen variables (Pelone, et al., 2015). Zakowska and Godycki-Cwirko (2020) determined in their systematic review that the most common input categories used in DEA applications in PHC were as follows: personnel, PHC centres, consultations or visits, referrals or hospitalizations, and pharmaceutical or prescriptions. The most common output categories were: health care consultations or visits; patients; procedures, treatments, and services; quality; personnel; preventive interventions; and PHC centres (Zakowska and Godycki-Cwirko, 2020). The number of inputs used in a single study varied between 1 to 24 and the outputs ranged from 1 to 21, with a modal value of 3 for both (Zakowska and Godycki-Cwirko, 2020).

Neri, et al. (2022) in their systematic review about approaches to measure efficiency in primary health care classified input variables according to the traditional categorisation of labour, capital and intermediate inputs and the outputs were classified depending whether they were pure measures of primary care utilisation (volume of activities performed or patients treated, without considering standards of care or health outcomes), measures of primary care utilisation with 'quality adjustment' (monitored quality of care and performance targets achieved), health outcomes, or the impact on health-related quality of life (HRQoL) (Neri, et al., 2022).

Pelone, et al. (2015) demonstrated that few studies based their outputs on health outcomes produced, many used primary care activities and others defined outputs in terms of primary care quality indicators. They explain that the disadvantage of basing the efficiency model on primary care activities is that it may reward units that produce more activities, as they are operating at a lower quality level, emphasizing the need for quality variables to be used (Pelone, et al., 2015). For inputs, these authors identify 3 main categories: labour, capital and consumable resources (Pelone, et al., 2015). Amado and Santos (2009) warn that the use of full time equivalent (FTE) of health care professionals as input and number of consultations as outputs leads to rewarding shorter consultations, additionally they explain that frequent consultations may not equate to superior primary health care.

Various authors included quality indicators in their analysis. Salinas-Jiménez and Smith (1996) used quality indicators as outputs to assess performance; Ferreira, et al. (2013) measured quality through the number of complaints but admit that the application of patient satisfaction surveys would have been more beneficial.

Many researchers include exogenous or environmental variables in their performance analysis. However, the specific identification of which variables are used and how they are applied in the studies differ. Deidda, et al. (2014) alert for the effect of not taking proper account of environmental variables, leading to erroneously classifying inefficient DMUs as efficient (and vice-versa). Neri, et al. (2022) grouped these into two categories: population characteristics and organisational features. Population characteristics include demographics, economic or employment status, mortality rates, geographic area, and more, while organisation features focus on the configuration and organisation of primary care delivery, financing level, external market influences, level of clinical performance and others (Neri, et al., 2022). Amado and Dyson (2009) also identified two environmental variables: area-related factors and patient related factors. Deidda, et al. (2014) used mortality index and patients over 65 years of age as their exogenous factors. Salinas-Jiménez and Smith (1996) used standardised illness ratio and unemployment as exogenous variables. Ferreira, et al. (2013) chose population, population density, percentage of patients aged 65 years or older, mortality rate, percentage of patients without designated doctor, distance to nearest hospital, and purchasing power as exogenous variables. Takundwa, et al. (2017) utilised a subset of public health outcomes framework that reflects deprivation, population demographics, burden of disease and lifestyle factors.

After identifying which environmental variables were being used it was crucial to understand how these were being applied by the authors. Amado and Dyson (2009) analysed the impact of environmental variables on performance results without including them in the DEA model. Salinas-Jiménez and Smith (1996) included the environmental variables in the DEA model as inputs. Takundwa, et al. (2017) used Tobit regression to quantify the relationship between efficiency and environmental variables. The efficiency scores from DEA were used as dependent variables and environmental factors were used as explanatory variables (Takundwa, et al., 2017). Ferreira, et al. (2013) used the order-m test to assess the effect of the exogenous variables on efficiency results.

Another aspect that is important to look at when using the DEA technique is the inclusion of weight restrictions. Incorporating weight restrictions involves value judgements and decision-makers should be included in the process of developing the restrictions (Pelone, et al., 2015). According to the systematic review carried out by Pelone, et al. (2015) most studies (31 out of 39) did not use weight restrictions to avoid compromising the objectivity of DEA and partly due to the lack of decision-makers' involvement in the modelling process.

Few authors of the studies that support this research inform about the use of weight restrictions. Amado and Dyson (2009) introduced two weight restrictions in their models that reflect technologically realistic trade-offs between inputs. These weight restrictions guarantee that no practice would be classified as effective simply due to having the lowest number of patients in a specific input (i.e. patients under regular treatment with statins). Amado and Santos (2009) introduced 5 weight restrictions to their efficiency model intending to ensure meaningful results and to account for production trade-offs between inputs and outputs. The first three restrictions represent trade-offs between inputs (number of doctors, nurses and administrative, technical and other support professionals) that could occur without detriment to the level of outputs (activities) (Amado and Santos, 2009). The fourth and fifth restrictions implemented by these authors state that the sum of weights attributed to the number doctors and the number of administrative, technical and other support professionals (restriction 4) or to the number of nurses and the number of administrative, technical and other support professionals (restriction 5), must be equal to the sum of weights attributed to all the outputs related to the doctors activity (restriction 4) or to the nurses activity (restriction 5), respectively (Amado and Santos, 2009). This way, health centres cannot place a high weight on a specific input while ignoring the outputs related to the activity of that input (Amado and Santos, 2009). These restrictions assume that administrative, technical, and other support professionals are auxiliary staff to the activity of the nurses and doctors (Amado and Santos, 2009). Oikonomou, et al. (2016) based their two weight restrictions on the consensus panel.

All studies identify variations in performance levels throughout the DMUs analysed, demonstrating DEA as a valuable technique. Amado and Dyson (2009) showed that regarding equity of service utilization, 23% of patients had not received some of the essential services in the period of the study. They identified three PCTs as technically efficient, only one trust was cost efficient, five trusts were clinically effective and two were patient-focused effective. Amado and Dyson (2009) found that high deprivation caused lower scores of clinical and patient-focused effectiveness, suggesting that these practices should have different outcome targets and should only be compared with each other. The conclusions reached by Amado and Santos (2009) were that higher population density lowers equity results, there are geographical inequalities in the distribution of doctors among the District Health Authorities, the average score of efficiency was 84.4% with DHA Lisbon and Tejo Valley scoring the lowest efficiency score (18%), while Aveiro, Évora and Leiria scoring highest. For this reason, the authors suggest the establishment of learning networks to promote efficiency improvement (Amado and Santos, 2009). These authors also demonstrated that there was asymmetry in the level of patient satisfaction with the care they received, which could reflect high expectations and not strictly poor quality of service (Amado and Santos, 2009).

Zhang, et al. (2018) concluded that 80% of primary health care institutions are inefficient, with a declining productivity from 2012 to 2016 due to technological retrogression. Dlouhy (2018) concluded that inefficient regions had excess health resources, the Robin Hood Index was superior for doctors than Nurses, however both resources need to be reallocated between regions. Jordi, et al. (2020) concluded that high-income countries produce on average Universal Health Coverage outputs more efficiently than other income groups, additionally higher-income, better governance, and more years of schooling within a country are associated with greater technical efficiency in converting health spending into Universal Health Coverage goals.

Salinas-Jiménez and Smith (1996) demonstrated that 51% of Family Health Service Authorities (FHSAs) were efficient and 49% were inefficient, under the input oriented model. However, under the output oriented model, the FHSAs showed better performance, as in this model all the outputs must be susceptible to improvement, while in the input oriented model only expenditure was susceptible to improvement (Salinas-Jiménez and Smith, 1996). More than 40% of regions in Hunan were inefficient, using excessive health resources to produce current level of health services (Huang, et al., 2020). The disproportionate distribution of the number of medical equipment and midwives between different regions was the main source of inequity (Huang, et al., 2020).

Ferreira, et al. (2013) concluded that efficiency scores were similar across orientations, there was a most common benchmark – Health Care Centre 17 (ACES), Nursing was the service with lowest efficiency scores and that purchasing power, percentage of patients aged 65 or older and population size were the exogenous variables that affect efficiency negatively. Deidda, et al. (2014) conclude that when not including the role of exogenous factors we can incorrectly consider inefficient DMUs as efficient and vice-versa. Also, Information and Communication Technology is under-used, increasing its use would lead to higher efficiency levels (Deidda, et al., 2014). Cinaroglu (2019) demonstrated that K-means clustering with DEA is useful to identify efficiencies of public hospitals located in provinces with similar welfare status. Through this technique, 5 province groups with similar welfare state indicators were created, of which the number of technically inefficient public hospitals was greater than the number of technically efficient public hospitals (Cinaroglu, 2019). Oikonomou, et al. (2016) showed that the main form of inefficiency was technical inefficiency and that the Heath Centres in the 6th Health Prefecture of Greece could produce 33% more output if they operated at maximum efficiency. Takundwa, et al. (2017) identified 47 of the 208 Clinical Commissioning Groups (CCG) as efficient, with the inefficient units showing an average technical efficiency score of 90%. These authors also demonstrated that three environmental factors were significant predictors of efficiency: CCGs with smaller population sizes were more efficient, high unemployment and high prevalence of Chronic Obstructive Pulmonary Disease (COPD) decreased efficiency, additionally deprivation was not a significant predictor of efficiency (Takundwa, et al., 2017).

Pelone, et al. (2015) conclude in their systematic review that studies attribute little effort into enhancing the accuracy of study findings and few adopt approaches to improve model specifications such as: selection of inputs and outputs, weight restrictions, super efficiency models to deal with outliers, bootstrapping technique to make the model more robust, which would provide policymakers with more reliable information. Neri, et al. (2022) concluded in their systematic review that reference to health outcomes was limited, consideration of allocative or cost efficiency or determinants of productivity was rare, research on the impact of technological change on productivity, role substitution across staffing and allocative efficiency is either dated or inadequate for the current PHC setting in England. Zakowska and Godycki-Cwirko (2020) agree that DEA can support health economic analyses in PHC, however, its main issues lie in creating a model that reflects the range of products and services in the primary health care sector.

Most authors discuss limitations to the Data Envelopment Analysis technique in their papers. Jordi, et al. (2020) describe DEA as deterministic, relying on the data selected, not being able to perform when data are absent, nor can it produce estimates when the underlying data are inaccurate; it provides a relative measure of efficiency, therefore excluding certain peers could influence the relative performance and other results. Zakowska and Godycki-Cwirko (2020) corroborate this explaining that DEA scores depend on the choice of variables, models and weighting. These authors add that the DEA technique ignores noise in the data and that the efficiency measures are very sensitive to sample size and outliers (Zakowska and Godycki-Cwirko, 2020). Amado and Santos (2009) alert that this technique assumes that it is possible to fully characterise the production of health care services by selecting inputs, outputs and outcomes of production, however some outputs and outcomes of primary health care are not measurable.

The future research suggestions by the authors mainly follow similar topics, emphasizing the need for further research in health care performance measures. Zakowska and Godycki-Cwirko (2020) recognize that appropriate input and output variables and a suitable DEA model for assessing PHC need to be researched further. Dlouhy (2018) suggests that further research should be based on the use of DEA models for inequality measurement. Takundwa, et al. (2017) recommend further research on appropriate measures of output in health care, and the development of guidelines on measuring efficiency in health care which would promote more robust methodological processes and increase comparability across studies. Amado and Santos (2009) advise some future research regarding: how different stakeholder views may be incorporated in the models to reflect multiple perspectives; how to appropriately measure primary health care long-term outcomes; DEA models that can reflect current performance and prior inheritances and investments in the achievement of long-term outcomes; in between others. Amado and Dyson (2009) suggest that variation in results, implicit trade-offs between performance measures and the impact of area deprivation are other issues that need further investigation. Huang, et al. (2020) propose the use of longitudinal data to provide more valuable information for policy makers. Deidda, et al. (2014) recommend looking into the use of Information and Communication technology in improving primary health care efficiency. Cinaroglu (2019) propose the inclusion of welfare indicators in future efficiency analyses. Neri, et al. (2022) suggest that the impact of technological change and skill-mix on primary health care efficiency should be studied and that future research should include more direct patient care staff, instead of focusing only on doctors and nurses.

The studies mentioned in this section serve as a theoretical support to this study. After this literature review, it is possible to conclude that this technique demonstrates various advantages for being used to assess primary health care performance and to measure inequality, such as: it can handle multiple inputs and outputs without requiring unit prices, which is useful in the health care setting where distinct services are provided and various outcomes are desired; it uses the data to construct the best practice empirical frontier, to which each non-optimal production point is compared; it does not require specification of the functional form that links inputs to outputs; performance can be assessed using several alternative orientations to the best practice frontier; as it is conceptualized as a linear program, a lot of useful information can be used for performance improvement (Amado and Dyson, 2009). Nonetheless, it has become more apparent that there are certain methodological choices that are crucial when creating the DEA models, for which the reviewed studies have provided guidelines towards the construction of the models applied in this research paper.

2.3. Malmquist Indices applied to Primary Health Care

The Malmquist Index was introduced by Caves, Christensen and Diewert (1982) and further developed for performance assessment by Färe, Grosskopf, Norris and Zhang (1994). The Malmquist Productivity Index (MPI) developed by Färe, et al. (1994) measures productivity change over time. This index can be decomposed into an efficiency change index (C) and a technological change index (F) (Camanho and Dyson, 2006). The efficiency change index (C) is also known as the catch-up effect which measures changes to the efficiency level of a decision-making unit from the first period (t) to the second (t+1). The technological change index (F) is also known as the frontier shift which measures changes to the efficiency frontier, therefore technological changes, from the first period (t) to the second period (t+1).

The MPI produced by Färe, et al. (1994) defined an input-oriented productivity index[1] as the geometric mean of the two Malmquist Indices developed by Caves, et al. (1982), which refer to the technologies at the time periods t and t+1, and can be demonstrated as:

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

where in the time period t the DMUs are using inputs X^t to produce outputs Y^t. $E^t(X^t,Y^t)$ is the efficiency rate of a DMU operating in period t with regards to the best practice frontier of period t. $E^{t+1}(X^{t+1},Y^{t+1})$ is the efficiency rate of a DMU operating in period t+1 with regards to the best practice frontier of period t+1. $E^t(X^{t+1},Y^{t+1})$ is the efficiency rate of a DMU operating in period t+1 with regards to the best practice frontier of period t and $E^{t+1}\,(X^t,Y^t)$ is the efficiency rate of a DMU operating in period t with regards to the best practice frontier of period t+1. In consistency with the recommendation of Cooper et al. (2007), when calculating these two last measures of efficiency, the super-efficiency model, proposed by Andersen and Peterson (1993), is used allowing the measures to assume values above one.

The Malmquist Index is decomposed into an index reflecting the change in technological efficiency and an index for the change in frontier of the production possibility set (Camanho and Dyson, 2006). These components are obtained by:

---

[1] This index can also be calculated using an output orientation. In this case, it is defined as the maximum radial expansion of the outputs, keeping the observed inputs. For examples of MI with an output orientation the reader is referred to Teixeira, et al. (2016) and Amado, et al. (2019).

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

where the ratio outside the square brackets measures the input technical efficiency change between the two periods (t and t+1) and the geometric mean of the two ratios inside the square brackets reflects the technological change between the two periods (Camanho and Dyson, 2006).

Regarding the technological approach, constant returns to scale should be used as the Malmquist Index provides an inaccurate productivity measure when a variable returns to scale approach is used (Camanho and Dyson, 2006).

Camanho and Dyson (2006) applied the DEA technique with Malmquist Indices to measure group performance. The method implemented by these authors is an adaptation to the Malmquist Index, where it no longer measures productivity change over time, but allows a cross-sectional performance comparison of groups of DMUs operating in different conditions at one moment in time (Camanho and Dyson, 2006). Their aim was to identify best performing bank branches in 4 regions in Portugal (Camanho and Dyson, 2006). This performance index can be decomposed into two indices: one that reflects efficiency spread among DMUs operating in similar conditions and another that reflects the productivity gap between the best-practice frontiers of the different groups (Camanho and Dyson, 2006). Thus, better performance is associated to less dispersion in efficiency levels within the group and dominance of the best practice frontier (Camanho and Dyson, 2006). Camanho and Dyson (2006) explain that the main advantage of this method is that it will use all the DMUs instead of using one "typical DMU" to represent the group. These researchers define an overall measure for comparison of performance between two groups of decision-making units (group A and B), which perform under different programs, using an input orientation. Our study uses an output orientation and, for that reason, we have adapted the formulas to be consistent with this. The formulas used in our study are as follows:

$$I^{AB} = \left[\frac{\left(\prod_{j=1}^{\delta A} E^{A}\left(X_j^{A}.Y_j^{A}\right)\right)^{\frac{1}{\delta A}}}{\left(\prod_{j=1}^{\delta B} E^{A}\left(X_j^{B}.Y_j^{B}\right)\right)^{\frac{1}{\delta B}}} \cdot \frac{\left(\prod_{j=1}^{\delta A} E^{B}\left(X_j^{A}.Y_j^{A}\right)\right)^{\frac{1}{\delta A}}}{\left(\prod_{j=1}^{\delta B} E^{B}\left(X_j^{B}.Y_j^{B}\right)\right)^{\frac{1}{\delta B}}}\right]^{\frac{1}{2}}$$

The two ratios inside the square brackets evaluate the efficiency of the DMUs with regards to a single reference technology, where the first evaluates the average efficiency, measured relative to group A's frontier, of DMUs from group A divided by the average efficiency of DMUs from group B. The second ratio follows the same principle relative to group B's frontier (Camanho and Dyson, 2006).

The overall performance measure presented previously can be decomposed into two sub-components, as follows:

$$I^{AB} = \frac{\left[\prod_{j=1}^{\delta A} E^{A}\left(X_j^A . Y_j^A\right)\right]^{\frac{1}{\delta A}}}{\left[\prod_{j=1}^{\delta B} E^{B}\left(X_j^B . Y_j^B\right)\right]^{\frac{1}{\delta B}}} \cdot \left[\frac{\left(\prod_{j=1}^{\delta A} E^{B}\left(X_j^A . Y_j^A\right)\right)^{\frac{1}{\delta A}}}{\left(\prod_{j=1}^{\delta A} E^{A}\left(X_j^A . Y_j^A\right)\right)^{\frac{1}{\delta A}}} \cdot \frac{\left(\prod_{j=1}^{\delta B} E^{B}\left(X_j^B . Y_j^B\right)\right)^{\frac{1}{\delta B}}}{\left(\prod_{j=1}^{\delta B} E^{A}\left(X_j^B . Y_j^B\right)\right)^{\frac{1}{\delta B}}}\right]^{\frac{1}{2}}$$

where the ratio outside the square brackets compares within-group efficiency spread and the ratio inside the square brackets evaluates the productivity gap between the group frontiers (Camanho and Dyson, 2006).

The index that compares within-group efficiency spread is given by the ratio of the geometric means of the efficiency of the DMUs with respect to their group specific frontier:

$$IE^{AB} = \frac{\left[\prod_{j=1}^{\delta A} E^{A}\left(X_j^A . Y_j^A\right)\right]^{\frac{1}{\delta A}}}{\left[\prod_{j=1}^{\delta B} E^{B}\left(X_j^B . Y_j^B\right)\right]^{\frac{1}{\delta B}}}$$

A value of IE^{AB} less than one reflects the efficiency spread is greater in the DMUs in group A than in those of group B, in other words, there is lower consistency in efficiency levels within group A compared to group B (interpretation adapted from Camanho and Dyson, 2006).

The other component of the new index compares frontier productivity of the two groups. The index for measuring the distance between the best-practice frontiers of groups A and B is given by:

$$IF^{AB} = \left[\frac{\left(\prod_{j=1}^{\delta A} E^{B}\left(X_j^A . Y_j^A\right)\right)^{\frac{1}{\delta A}}}{\left(\prod_{j=1}^{\delta A} E^{A}\left(X_j^A . Y_j^A\right)\right)^{\frac{1}{\delta A}}} \cdot \frac{\left(\prod_{j=1}^{\delta B} E^{B}\left(X_j^B . Y_j^B\right)\right)^{\frac{1}{\delta B}}}{\left(\prod_{j=1}^{\delta B} E^{A}\left(X_j^B . Y_j^B\right)\right)^{\frac{1}{\delta B}}}\right]^{\frac{1}{2}}$$

A value of IF^{AB} less than one reflects lower productivity of the frontier of group A compared to group B (interpretation adapted from Camanho and Dyson, 2006).

These authors developed measures to compare performance of groups of DMUs operating under different conditions. Attributing performance differences to their source is a very important issue in any managerial context (Camanho and Dyson, 2006). This was addressed by decomposing the overall performance measure into efficiency spread among DMUs in each group reflecting internal managerial efficiency; whereas the gap in frontier productivity between groups reflects the context in which the DMUs are required to operate (Camanho and Dyson, 2006).

The method developed by these authors will be applied in this study, using the level of deprivation as factor to group the DMUs.

To the best of our knowledge, there are no published studies that have used this index to contrast the performance of primary health care providers according to the level of deprivation of the region served.

3. Empirical Analysis

For this research, Data Envelopment Analysis was the technique chosen to measure performance of each DMU individually, regarding equity of resource allocation, service efficiency and effectiveness. Secondly, the DMUs were divided into two groups and the Malmquist Indices proposed by Camanho and Dyson (2006) were used to analyse group performance and identify health inequalities according to the Index of Multiple Deprivation (IMD).

3.1. DEA Models and Malmquist Indices

In the research papers that support this study, authors emphasize that choosing which variables to include in the DEA model is a very important and determinant step in this technique. The choice of variables in this study is consistent with those applied in previous studies, subject to available data. The models obey the rule of thumb established by Banker et al. (1984) that states that the number of DMUs should be at least three times the combined number of inputs and outputs. In this study, the EMS Data Envelopment Analysis Software, developed by Scheel (2000), was used to run the DEA models.

Three models were run to assess equity in resource distribution and to measure service efficiency and service effectiveness. These models intend to provide insightful information regarding access and utilization of primary health care and achievement of health outcomes. An output orientation was used in all models. Although researchers should base orientation on what managers are able to control – resources or outcomes (Pelone, et al., 2015), in health care the focus should concentrate on increasing outputs instead of reducing inputs and costs (Zakowska and Godycki-Cwirko, 2020). An output orientation is particularly relevant when there is evidence of waiting lists for services, as it is the case in the UK.

For each model, weight restrictions were applied following the production trade-offs method proposed by Podinovski (2004). Then the super-efficiency procedure for outlier identification by Banker and Chang (2006) was utilised to identify outlying DMUs.

In the first model, the aim was to measure human resource distribution equity based on registered patients list, by age group. In health care, horizontal equity means providing equal treatment to equal needs.

As discussed in Neri, et al. (2022) there are challenges to workforce recruitment and retention within primary health care in the NHS, however the expansion of workforce is recommended by the General Practice Forward View, as this sector is heavily reliant on human resources to deliver effective health care. This model intends to identify inequities that may exist in the distribution of precious human resources according to need, throughout the DMUs.

As shown in table 1, the chosen input variables were number of patients, stratified by age. This stratification was chosen as it follows the stratification applied by NHS Digital in the Health Survey for England 2019. In this model, age is used as a measure for health care need by the population. As shown by Ferreira, et al. (2013) there is a negative influence of patients aged 65 or older on efficiency, as aging increases need for health care services which leads to an increment in costs. Neri, et al. (2022) also recognise that a growing population characterised by old and multi-morbid patients has increased the demand for primary health care services.

As seen in the table 1, as output variables, Full-Time Equivalent (FTE) of each professional group is used, to inform about the human resources available in each CCG. Full-time equivalent is defined by NHS Digital as the proportion of full-time contracted hours the post holder is contracted to work, where 1 FTE would indicate full time (37.5h/week). Therefore, it becomes evident that this measure is more useful for the analysis than simply the headcount of professionals, where the number of service hours is not reflected. According to NHS Digital (2020) Direct Patient Care Staff refers to Health Care Assistants, Care Coordinators, Dispensers, Phlebotomists, Pharmacists, Podiatrists, Physiotherapists, Physician Associates, Apprentices, Paramedics, Nursing Associates, Trainee Nursing Associates, Social Prescribing Link Workers, Improving Access to Psychological Therapies (IAPT) staff, Trainee IAPT Staff, Health Support Worker, and Other Direct Patient Care. Whereas Administration and Non-clinical Staff refers to Managers, Management Partners, Medical Secretaries, Receptionists, Telephonists, Estates and Ancillary, Apprentices, and Other Admin/non-clinical (NHS Digital, 2020). The inclusion of all these professions intends to realistically demonstrate the functioning of primary health care services, following the recommendation of Neri, et al. (2022).

Table 1 - DEA model to measure human resource distribution equity based on registered patients list, by age group.

Inputs {I}	Outputs {O}
Input 1: Number of patients aged [0-15]	**Output 1:** General Practitioner Full time equivalent (FTE)
Input 2: Number of patients aged [16-24]	**Output 2:** Nurses Full time equivalent (FTE)
Input 3: Number of patients aged [25-34]	**Output 3:** All Direct Patient Care Staff Full time equivalent (FTE)
Input 4: Number of patients aged [35-44]	**Output 4:** Administration and Non-clinical Staff Full time equivalent (FTE)
Input 5: Number of patients aged [45-54]	
Input 6: Number of patients aged [55-64]	
Input 7: Number of patients aged [65-74]	
Input 8: Number of patients aged ≥75	

This model was applied with a Constant Returns to Scale (CRS) assumption. In this respect, to guarantee complete equity, the quantity of resources that each CCG should have if they operated under the best scale, to serve the registered patients, will be identified.

The second model aimed to measure service efficiency, analysing the quantity of services delivered versus the resources used to produce these. In health care, efficiency is of central importance as available resources are limited, therefore an efficient health system is one in which resources are optimally converted into population gains (Smith, 2009).

As shown in table 2, the chosen input variables were Full-Time Equivalent (FTE) of each professional group, to inform about the resources available in each CCG. The description of which professions are included in the direct patient care staff group and in the administration and non-clinical staff group is presented in the previous model. As output variables activity indicators were used. In particular, we used count of appointments by mode. According to NHS Digital (2020), the variable "Count of appointments by mode: other" reflects appointments that were held via telephone, video or online and those that were classified as unknown mode.

This model follows the research developed by Amado and Santos (2009), however these authors alert that this choice of variables may lead to rewarding units that provide shorter consultations, without regard to quality. These authors also warn that frequent consultations may not mean better primary health care, and suggest that for a fair performance assessment, service effectiveness and quality should be evaluated (Amado and Santos, 2009). Following the recommendation of Amado and Santos (2009), service effectiveness is also measured in the present study (model 3), including the use of a quality indicator.

Table 2 - DEA model to measure service efficiency.

Inputs {I}	**Outputs {O}**
Input 1: General Practitioner Full-time equivalent (FTE) **Input 2:** Nurses Full-time equivalent (FTE) **Input 3:** All Direct Patient Care Staff Full-time equivalent (FTE) **Input 4:** Administration and Non-clinical Staff Full-time equivalent (FTE)	**Output 1:** Count of appointments by mode: face-to-face **Output 2:** Count of appointments by mode: home visit **Output 3:** Count of appointments by mode: other

This model was analysed from a Constant Returns to Scale and a Variable Returns to Scale approach, as it is beneficial to calculate technical efficiency, pure technical efficiency, and scale efficiency, allowing to detangle between the different types of inefficiency. Using this method, we can identify how many CCGs are operating at an ideal size and which have increasing or decreasing returns to scale. This kind of information is highly valuable for managers and policy makers.

The third model measured service effectiveness in reducing hospital admissions and maximizing patient satisfaction. In this model, we analyse if the services delivered produce desired outcomes, which in this case would mean fewer emergency admissions and higher patient satisfaction. Golany and Tamir (1995) define effectiveness as the distance between observed outputs and the goals set.

As shown in table 3, the chosen input variables were the 3 measures of count of appointments by mode which reflect primary health care activity. To be able to use emergency admissions as an output, we transformed an undesirable output into a desirable one as proposed in Scheel (2001). The second output was the estimated number of patients that rate an overall good experience, from the GP Patient Survey, which provides an insight of the level of patient satisfaction towards the primary health care services provided. The use of patients' level of satisfaction is important to track the service quality (Ferreira, et al., 2013). Pelone, et al. (2015) explain that using quality indicators in the DEA model may overcome the drawbacks of basing efficiency models on primary health care activity.

This model follows the principles demonstrated by Amado and Dyson (2009), who calculated clinical effectiveness using activity indicators as inputs and clinical outcomes as outputs.

Table 3 - DEA model to measure service effectiveness in reducing hospital admissions and maximizing patient satisfaction.

Inputs {I}	Outputs {O}
Input 1: Count of appointments by mode: face-to-face	**Output 1:** Emergency admissions saved when compared to worst observed
Input 2: Count of appointments by mode: home visit	**Output 2:** Estimated number of patients that rate an overall good experience
Input 3: Count of appointments by mode: other	

This last model was analysed from a Variable Returns to Scale approach, as using a CRS approach could lead us into producing targets that are not reachable with the current patient list.

The MPI proposed by Camanho e Dyson (2006) was used to measure group performance when accounting for the level of deprivation. This method aims to assess the performance of two groups, one with the highest and the other with the lowest Index of Multiple Deprivation (IMD) scores. These were divided according to median value. The first group had higher than or equal to the median value of IMD, corresponding to higher deprivation, and the second group had lower than median value of IMD, corresponding to lower deprivation.

The Index of Multiple Deprivation is composed of seven domains to which weights are attributed to calculate the overall average score, as follows: Income Deprivation – 22,5%; Employment Deprivation- 22,5%; Education, Skills and Training Deprivation- 13,5%; Health Deprivation and Disability- 13,5%; Crime- 9,3%; Barriers to Housing and Services- 9,3%; Living Environment Deprivation- 9,3% (Ministry of Housing, Communities and Local Government, 2019). The indices of deprivation are a unique measure of relative deprivation at a small local area level (Lower-layer Super Output Areas – LSOAs), which are areas that divide the countries in even areas, each with similar populations, however IMD scores are available for higher-level areas such as Local Authority Districts and Clinical Commissioning Groups (Ministry of Housing, Communities and Local Government, 2019).

Using these two groups, the three models were run to calculate efficiency spread, frontier shift and overall performance. This method intended to produce information about how performance may be influenced by deprivation, studying the presence of health inequalities.

3.2. Equity, efficiency and service effectiveness data and results

In this study, Clinical Commissioning Groups (CCG) were the decision-making units (DMU) used in the models to assess performance and identify inequalities in primary health care in England. In 2019, there were a total of 191 CCGs of which 12 were excluded from the analysis. NHS Bury CCG, NHS Derby and Derbyshire CCG and NHS Devon CCG were excluded due to not presenting complete activity data for the twelve months of 2019 calendar year. Nine more were excluded as they were identified as outliers. Therefore, the study was based on 179 DMUs (CCGs in England).

To identify outliers, the Banker and Chang (2006) method was utilised, where DMUs with a superefficiency score greater than 120% were excluded (Banker and Chang, 2006). This process was repeated for the three models. After identifying and excluding outliers in all three models (table 4), the models were applied with the same 179 DMUs.

Table 4 – Outliers identified using Super-efficiency analysis.

Model 1	Model 2	Model 3
NHS Corby CCG	NHS Nottingham West CCG	NHS Oxfordshire CCG
NHS Camden CCG	NHS Portsmouth CCG	NHS Dartford, Gravesham and Swanley CCG
		NHS Luton CCG
		NHS Darlington CCG
		NHS Castle Point and Rochford CCG

All the data included in these analyses were sourced from the NHS digital website and refer to the calendar year 2019. The choice for this calendar year was based on data availability, however, it would be relevant to run these models with more recent data as it becomes available.

Table 5 presents descriptive statistics for the variables used in the first DEA model - Measuring Human Resource Distribution Equity based on Registered Patients List, by Age Group. This table demonstrates considerable discrepancies throughout the average, maximum and minimum values of the inputs and outputs of the CCGs.

Table 5 – Descriptive statistics for the variables in the first DEA model.

Variables	Average	Standard deviation	Max	Min
Registered patients Ages 0-15 {I}	56475,96	34316,79	282390,25	17699,25
Registered patients Ages 16-24 {I}	32563,35	22338,24	158273,75	8980,42
Registered patients Ages 25-34 {I}	44955,20	30062,11	198337,33	11921,83
Registered patients Ages 35-44 {I}	41867,41	26038,75	184206,17	12467,08
Registered patients Ages 45-54 {I}	42236,32	23451,70	170807,33	14989,92
Registered patients Ages 55-64 {I}	36384,95	19930,71	137618,42	9628,50
Registered patients Ages 65-74 {I}	28971,99	16870,23	101395,42	5863
Registered patients Ages 75 or more {I}	24506,51	15112,08	94658,83	4267,25
GP FTE {O}	172,73	106,46	809,08	48
Nurses FTE {O}	86,10	54,78	339,80	22,17
All direct patient care FTE {O}	72,20	48,85	314,84	18,52
Admin/Non-clinical FTE {O}	353,24	203,29	1467,66	97,36

NHS Birmingham and Solihull CCG has the maximum number of patients registered with 1325684 patients and NHS Surrey Heath CCG has the minimum with 97528. However, these numbers give us limited information if not analysed in context and proportion. The next step was to investigate the average, maximum and minimum number of registered patients per staff group FTE, which is presented in table 6. As age is being used to measure health need, we calculated the average, maximum and minimum number of registered patients that are 65 years of age or older, in the CCGs. Then, the descriptive statistics were analysed for the number of registered patients of 65 years of age or older per staff group FTE, as represented in table 6.

This analysis of the data intends to provide explanatory information when discussing the equity results obtained through the DEA technique.

Table 6 - Number of patients per staff group FTE.

	Value	Corresponding CCG
Average number of patients per GP FTE	1826,05	
Maximum number of patients per GP FTE	2726,37	NHS Thanet CCG
Minimum number of patients per GP FTE	1180,28	NHS Rushcliffe CCG
Average number of patients per Nurse FTE	3833,84	
Max. number of patients per Nurse FTE	9444,68	NHS Redbridge CCG
Min. number of patients per Nurse FTE	1973,68	NHS Durham Dales, Easington and Sedgefield CCG
Average number of patients per DPC FTE	4729,96	
Max. number of patients per DPC FTE	8347,29	NHS Sutton CCG
Min. number of patients per DPC FTE	1457,78	NHS North Norfolk
Average number of patients per Admin/non-clinical FTE	879,16	
Max. number of patients per Admin/non-clinical FTE	1385,72	NHS Hammersmith and Fulham CCG
Min. number of patients per Admin/non-clinical FTE	576,39	NHS North Norfolk CCG
Average number of patients aged 65y or above per CCG	53478,50	
Maximum number of patients aged 65y or above per CCG	195807,83	NHS Dorset CCG
Minimum number of patients aged 65y or above per CCG	10130,25	NHS Bradford City CCG
Average number of patients aged 65y or above per GP FTE	323,89	
Maximum number of patients aged 65y or above per GP FTE	632,39	NHS Hastings and Rother CCG
Minimum number of patients aged 65y or above per GP FTE	94,33	NHS Tower Hamlets CCG
Average number of patients aged 65y or above per Nurse FTE	646,53	
Maximum number of patients aged 65y or above per Nurse FTE	1096,58	NHS Redbridge CCG
Minimum number of patients aged 65y or above per Nurse FTE	183,73	NHS Bradford City CCG
Average number of patients aged 65y or above per DPC FTE	805,59	
Maximum number of patients aged 65y or above per DPC FTE	1477,88	NHS Fareham Gosport CCG
Minimum number of patients aged 65y or above per DPC FTE	195,16	NHS Bradford City CCG
Average number of patients aged 65y or above per Admin/non-clinical FTE	153,42	
Maximum number of patients aged 65y or above per Admin/non-clinical FTE	221,67	NHS Coastal West Sussex CCG
Minimum number of patients aged 65y or above per Admin/non-clinical FTE	47,18	NHS Bradford City CCG

For the first model, twenty-four weight restrictions were implemented (presented in appendix 2), following the production trade-offs method proposed by Podinovski (2004). The use of weight restrictions aimed to produce more meaningful results. The first seven (WR1-WR7) reflect that none of the ouputs (human resources) can be reduced by replacing one patient of a younger age group for one aged 75 or over. The next six (WR8-WR13) reflect that none of the ouputs can be reduced by replacing one patient of a younger age group for one in the age group 65-74 years of age. The further five weight restrictiones (WR14-WR18) reflect that none of the outputs can be reduced by replacing one patient in the age groups between 16 to 54 years of age, by one from the age group 0-15 years of age.

The last 6 weight restrictions refer to possible trade-offs between outputs (human resources), without altering the inputs. These reflect that one nurse, direct patient care staff or adminstration/non-clinical staff could be reduced if one doctor was added.

The same applies as one nurse can replace a direct patient care worker or a adminstration/non-clinical worker, and a direct patient care worker can replace one adminstration/non-clinical employee. All these trade-offs between outputs imply that the inputs would not change, in other words, the patient list would be unaltered.

Having previously justified the choice of variables included in the model, the results will now be presented and discussed. Due to the large number of decision-making units under analysis, it is not possible to demonstrate the results of all of them here. Table 7 demonstrates the equitable Clinical Commissioning Groups, the inequitable CCG with the highest score and the one with the lowest score. These results are also illustrated in figure 1.

The results of the first model reveal that 13 Clinical Commissioning Groups have equitable distribution of human resources for their populations' need: NHS Bradford City CCG, NHS City and Hackney CCG, NHS Durham Dales, Easington and Sedgefield CCG, NHS Hammersmith and Fulham CCG, NHS High Weald Lewes Havens CCG, NHS Islington CCG, NHS Liverpool CCG, NHS North Cumbria CCG, NHS North Norfolk CCG, NHS Rushcliffe CCG, NHS Tower Hamlets CCG, NHS Vale of York CCG and NHS West Norfolk CCG.

The average relative resource equity rate was 85,51%. Of the 166 inequitable CCGs, 55 had a relative resource equity rate below 80%. The CCG with the lowest relative resource equity rate was NHS Basildon and Brentwood with a score of 65,19%.

Table 7 - Equity Scores obtained with DEA.

DMU	EQUITY (%)	patients Ages 0-15 {I}{V} (%)	patients Ages 16-24 {I}{V} (%)	patients Ages 25-34 {I}{V} (%)	patients Ages 35-44 {I}{V} (%)	patients Ages 45-54 {I}{V} (%)	patients Ages 55-64 {I}{V} (%)	patients Ages 65-74 {I}{V} (%)	patients Ages 75 or more {I}{V} (%)	GP FTE {O}{V} (%)	Nurses FTE {O}{V} (%)	All direct patient care FTE {O}{V} (%)	Admin/ Non-clinical FTE {O}{V} (%)	Benchmarks
NHS Bradford City CCG	100	29	0	20	18	0	7	4	23	16	14	14	56	74
NHS City and Hackney CCG	100	24	13	31	0	0	0	5	28	100	0	0	0	53
NHS Durham Dales, Easington and Sedgefield CCG	100	20	10	14	0	0	0	13	43	45	39	3	13	15
NHS Hammersmith and Fulham CCG	100	68	0	0	0	0	0	14	18	68	13	19	0	1
NHS High Weald Lewes Havens CCG	100	28	0	16	18	0	0	21	18	73	4	5	18	0
NHS Islington CCG	100	64	8	6	0	0	0	4	18	1	0	0	0	4
NHS Liverpool CCG	100	43	0	2	19	0	0	11	25	62	6	4	28	144
NHS North Cumbria CCG	100	48	6	0	7	0	0	8	30	53	40	4	4	108
NHS North Norfolk CCG	100	20	11	13	13	0	0	22	20	19	13	19	49	109
NHS Rushcliffe CCG	100	28	13	6	0	0	20	18	15	78	20	1	0	93
NHS Tower Hamlets CCG	100	0	0	0	0	0	0	0	100	66	25	1	8	30
NHS Vale of York CCG	100	32	0	9	0	0	0	22	38	49	23	27	2	1
NHS West Norfolk CCG	100	12	5	3	8	0	8	10	54	37	26	37	0	1
NHS Bradford Districts CCG	99,72	22	11	14	0	0	12	9	32	43	21	5	30	16 (0.15) 37 (0.26) 86 (0.35) 98 (0.11)
NHS Basildon and Brentwood CCG	65,19	21	10	15	0	0	13	10	30	50	15	5	30	86 (0.29) 98 (0.12) 106 (0.20) 119 (0.36)

The virtual weights presented in this table are the proportion that is attributed to each input (or output) of the total of inputs (or outputs).

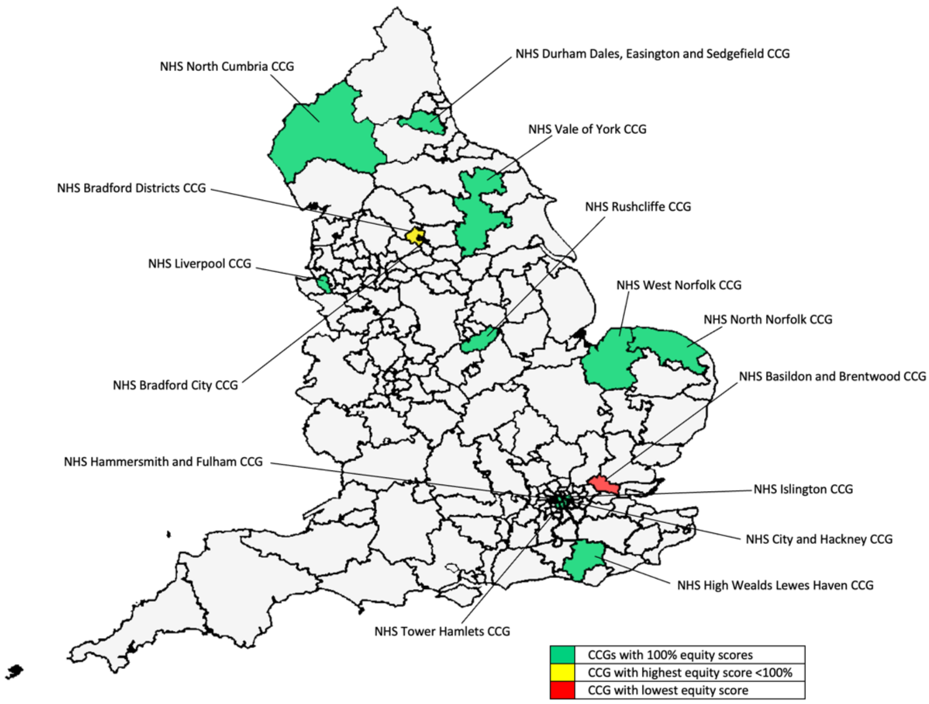

Figure 1 - Map of England representing equity results obtained with DEA for Model 1.

If we combine the information in table 6 with the results obtained through the DEA technique in table 7, some interesting findings have been reached: NHS Thanet CCG has the highest number of registered patients per GP FTE with an equity score of 76,85% in contrast NHS Rushcliffe CCG has the lowest number of registered patients per GP FTE and is 100% equitable; NHS Redbridge CCG has the highest number of registered patients per nurse FTE being the third worst CCG regarding equity score (67,15%) whereas NHS Durham Dales, Easington and Sedgefield CCG has the lowest number of registered patients per nurse FTE and has resource equity. A similar result occurs regarding the maximum and minimum number of registered patients per all direct patient staff FTE. Conversely, the maximum and minimum number of registered patients per administrative and non-clinical staff FTE does not show these results, as both the CCGs (with maximum and minimum registered patients per this staff group) are equitable. These results may suggest that high staff-to-patient ratios for GPs, nurses and direct patient care staff result in reduced resource equity amongst the CCGs.

When it comes to the number of registered patients aged 65 years of age or older per staff group FTE, all CCGs that present the minimum number of patients in this age group per staff group FTE are resource equitable: NHS Tower Hamlets CCG (lowest per GP FTE), NHS Bradford City CCG (lowest per nurse FTE, direct patient care staff FTE and administrative and non-clinical staff FTE). Additionally, the CCGs that present the maximum number of registered patients 65 years of age or older per staff group FTE have resource equity scores under 80%: NHS Hastings and Rother CCG (highest per GP FTE), NHS Redbridge CCG (highest per nurse FTE), NHS Fareham and Gosport CCG (highest per direct patient care staff FTE) and NHS Coastal West Sussex CCG (highest per administrative and non-clinical staff FTE). These results may suggest that greater number of registered patients aged 65 years of age or older per staff group FTE decreases resource equity scores amongst the CCGs.

When analysing the results obtained through DEA, it is also important to investigate the optimal weight structure chosen by each decision-making unit. Table 7 illustrates the virtual weights attributed to the input and output variables. It is important to emphasize that for the CCGs with 100% score in each model, there are multiple optimal weight profiles for the inputs and outputs. The weight structures presented in the tables and discussed in this paper are one of the possible optimal weight structures for these CCGs. The benchmark column represents the number of times that equitable CCGs serve as a benchmark to other CCGs.

If we initially analyse the optimal weight profiles for the input variables of the equitable CCGs NHS Tower Hamlets is the only one that attributes the total of the virtual weight to one input variable – Registered patients aged 75 or above. All other equitable CCGs apply weights on 3 to 7 input variables. There is a common trend amongst the equitable CCGs to attribute zero weight to Registered patients aged 45-54.

Concerning the optimal output weight profiles, table 7 reveals that NHS City and Hackney CCG and NHS Islington CCG apply all the virtual weight on one output – GP FTE. Besides these two CCGs, eight of the equitable CCGs apply highest weight on the output GP FTE, which suggests that this staff group contributes heavily towards increasing resource equity. 5 equitable CCGs exclude the output Admin/non-clinical FTE.

Clinical Commissioning Groups considered to benefit from resource equity are benchmarks to units that suffer from relative inequity. NHS High Weald Lewes Havens CCG is the only equitable CCG that is not a benchmark to any of the inequitable DMUs, this may be due to different management, structures or processes. NHS Liverpool CCG is the decision-making unit that serve as benchmark to the highest number of CCGs (144), followed by North Norfolk (109) (as shown in table 7). This means that inequitable CCGs could improve their relative resource equity by learning from their benchmarks.

The second DEA model intended to measure service efficiency. Table 8 presents descriptive statistics for the variables used in this model. It is possible to identify variations in inputs and outputs throughout the CCGs. Administration and Non-clinical staff represent the highest Full-Time Equivalent, followed by the GPs. From the output variables we can conclude that face-to-face is the most common mode of appointments within primary health care in the NHS, followed by other mode which include telephone, video or online appointments and those that were classified as unknown mode. Home visits are the least frequent mode of appointments.

Table 8 - Descriptive statistics for the variables in the second DEA model.

Variables	Average	Standard deviation	Max	Min
GP FTE {I}	172,73	106,46	809,08	48
Nurses FTE {I}	86,10	54,78	339,80	22,17
All direct patient care FTE {I}	72,20	48,85	314,84	18,52
Admin/Non-clinical FTE {I}	353,24	203,29	1467,66	97,36
Count of appointments by mode: Face-to-face {O}	1243481	759480,9	5754826	202280
Count of appointments by mode: home visit {O}	14798,72	15486,45	98799	117
Count of appointments by mode: other {O}	257201,1	191647,5	1280326	34047

For the second model nine weight restrictions were implemented (presented in appendix 2), following the production trade-offs method proposed by Podinovski (2004). Six of which reflect trade-offs that may occur in the inputs without affecting the outputs, the other three apply the same principal for trade-offs among the outputs.

The first three (WR1-WR3) imply that one doctor could replace one professional of the other three professional groups without detrimentally affecting the quantity of outputs (appointments). The following two weight restrictions indicate that one nurse could replace a direct patient care worker or an administration/non-clinical worker without detrimentally affecting the quantity of outputs (appointments). And finally, one direct patient care employee can replace an administration/non-clinical worker without detrimentally affecting the quantity of outputs (appointments).

The last three weight restrictions indicate that an appointment of one mode can be provided in a different mode without requiring alterations in the inputs (staffing).

As mentioned previously, this model was analysed from a Constant Returns to Scale and a Variable Returns to Scale approach, enabling us to calculate technical efficiency, pure technical efficiency and scale efficiency. Technical efficiency (TE) assesses if the maximum level of outputs is obtained with the set of inputs, whereas pure technical efficiency (PTE) calculates if the units are producing the maximum level of outputs with a set of inputs considering the size of the unit and reflects internal management strategies. Thirdly, scale efficiency (SE) analyses the extent to which the unit is operating at a production scale that maximizes the ratio of outputs
to inputs.

First the model was run with a CRS approach and with an output orientation, which revealed that 17 CCGs were technically efficient (demonstrated in table 9): NHS Bradford City CCG, NHS Thanet CCG, NHS Wirral CCG, NHS North Kirklees CCG, NHS Hastings and Rother CCG, NHS Calderdale CCG, NHS North East Essex CCG, NHS Hounslow CCG, NHS Thurrock CCG, NHS Southport and Formby CCG, NHS Cannock Chase CCG, NHS West Lancashire CCG, NHS Ashford CCG, NHS Swindon CCG, NHS Canterbury and Coastal CCG, NHS Redditch and Bromsgrove CCG and NHS North Hampshire CCG.

The average technical efficiency score was 84,62%. Of the inefficient units, NHS South Warwickshire CCG had the highest score with 99,9% and NHS Richmond CCG had the lowest with 50,37%. Of the 162 inefficient CCGs, 53 had a score under 80%. The efficiency results are also illustrated in figure 2.

Table 9 - Technical Efficiency results obtained with DEA.

DMU	TE (%)	GP FTE {I}{V} (%)	Nurses FTE {I}{V} (%)	All direct patient care FTE {I}{V} (%)	Admin/Non-clinical FTE {I}{V} (%)	Count of appointments by mode: Face-to-face {O}{V} (%)	Count of appointments by mode: home visit {O}{V} (%)	Count of appointments by mode: other {O}{V} (%)	Benchmarks
NHS Bradford City CCG	100	75	13	12	0	100	0	0	1
NHS Thanet CCG	100	100	0	0	0	37	63	0	0
NHS Wirral CCG	100	28	12	8	53	0	100	0	0
NHS North Kirklees CCG	100	47	24	5	24	100	0	0	35
NHS Hastings and Rother CCG	100	83	17	0	0	75	11	14	38
NHS Calderdale CCG	100	51	30	19	0	84	3	13	75
NHS North East Essex CCG	100	36	14	8	42	92	8	0	46
NHS Hounslow CCG	100	38	12	9	41	87	0	13	106
NHS Thurrock CCG	100	52	30	16	3	77	0	23	34
NHS Southport and Formby CCG	100	70	30	0	0	0	100	0	0
NHS Cannock Chase CCG	100	55	27	17	0	88	12	0	68
NHS West Lancashire CCG	100	65	20	14	0	66	34	0	29
NHS Ashford CCG	100	38	22	9	31	70	6	24	39
NHS Swindon CCG	100	42	13	7	38	89	0	11	19
NHS Canterbury and Coastal CCG	100	75	5	4	16	38	56	6	1
NHS Redditch and Bromsgrove CCG	100	29	12	10	49	73	22	5	33
NHS North Hampshire CCG	100	32	14	11	43	58	15	27	77
NHS South Warwickshire CCG	99.90	31	12	12	46	74	15	11	89 (0.17) 128 (0.88) 161 (0.48)
NHS Richmond CCG	50.37	60	17	13	10	33	3	64	54 (0.13) 89 (0.30) 161 (0.48)

Table 9 illustrates the virtual weights attributed to the input and output variables by the CCGs. NHS Thanet CCG is the only technically efficient CCG that attributes the total virtual weight on one input – GP FTE. Additionally, another 10 technically efficient CCGs attribute highest weight on this input variable – this reflects that GP FTE is the input in these groups that most increases their efficiency result, suggesting that this input produces higher output levels. NHS Bradford City CCG, NHS Hastings and Rother CCG, NHS Calderdale CCG, NHS Southport and Formby CCG, NHS Cannock Chase CCG and NHS West Lancashire CCG exclude the output variable Admin/non-clinical FTE – reflecting that this input least promotes their efficiency results. After the input variable Admin/non-clinical FTE, All direct patient care FTE is the most excluded input variable (3 equitable CCGs attribute zero weight to this input variable).

Concerning the optimal output weight profiles, as demonstrated in table 9, NHS Bradford City CCG, NHS Wirral CCG, NHS North Kirklees CCG and NHS Southport and Formby CCG attribute all their output weight structure on one variable – either Count of appointments by mode: face-to-face or Count of appointments by mode: home visit.

Among the equitable CCGs there is a common trend to attribute the highest weight on the input variable Count of appointments by mode: face-to-face, which may indicate that for these groups this appointment mode is the one that requires less inputs. Seven equitable CCGs attribute zero weight to the variable Count of appointments by mode: other, which may suggest that for these groups this is the appointment mode that requires higher levels of inputs, lowering their technical efficiency scores.

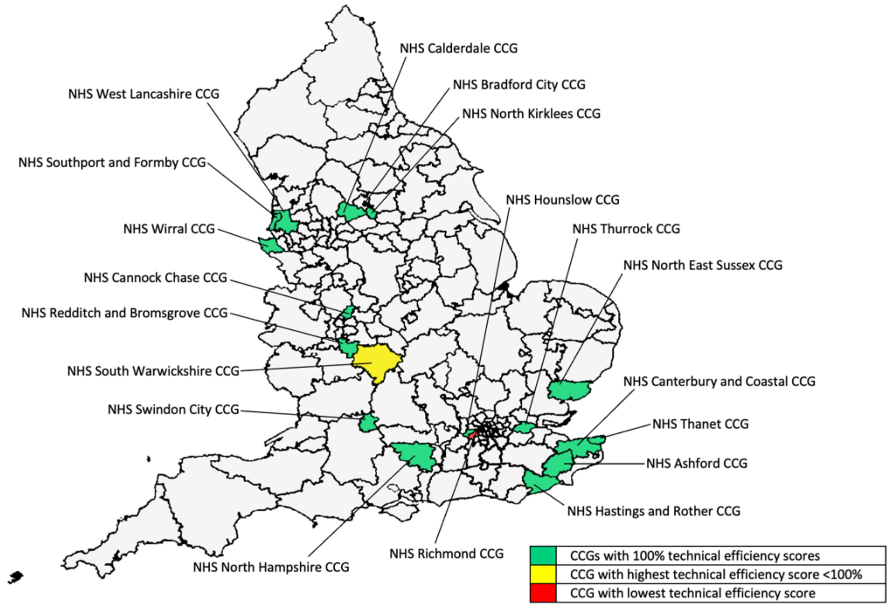

Figure 2- Map of England representing technical efficiency results obtained with DEA for Model 2.

Table 9 also provides information regarding benchmarks, where 3 of the 17 technically efficient CCGs are not benchmarks to any of the inefficient units. NHS Hounslow CCG is the efficient unit that serves as a benchmark to the highest number of inefficient units (106), followed by NHS North Hampshire CCG (77).

After analysing the model from a CRS approach, a Variable Returns to Scale was applied. As demonstrated in table 10 and in figure 3, 33 Clinical Commissioning Groups were pure technically efficient. The average pure technical efficiency score was 87,52%. Of the inefficient units, NHS Wyre Forest CCG was the unit with the highest score (99,5%) and NHS Richmond CCG had the lowest (50,42%). Of the 146 inefficient CCGs, 37 had scores of pure technical efficiency under 80%.

The optimal input weight profiles, as shown in table 10, reveal that NHS Thanet CCG, NHS North Cumbria CCG, NHS Bristol, North Somerset and South Gloucestershire CCG and NHS Cambridshire and Peterborough CCG are pure technically efficient units that apply the total weight on a single input variable – GP FTE. Besides these 4 CCGs, 24 other efficient units apply the highest proportion of virtual weight on this input variable. If we look back at the weight restrictions applied to this model, GP FTE was the only input variable that could substitute all others in the production trade-offs. This choice of applying higher weight on this input suggests that it is the staff group that most contributes towards the efficiency results. Following the input variable GP FTE, the variable Nurse FTE is the one that presents higher weights attributed. Admin/non-clinical staff FTE is the most commonly excluded input variable from the optimal input weight profiles, which may reflect that it is the staff group that least contributes towards increasing the pure technical efficiency results.

Regarding the optimal output weight profiles, we can see from table 10, that none of the efficient units apply total weight, nor the highest weight proportion to Count of appointments by mode: other variable, whereas Count of appointments by mode: face-to-face was the most frequent output variable with the highest virtual weight attributed. In addition, 2 efficient units apply total output weight to this variable (Count of appointments by mode: face-to-face). This reflects that the appointment mode face-to-face is the output that most increases technical efficiency results with these inputs.

Of the 33 CCGs that have pure technical efficiency, 4 CCGs, do not serve as a benchmark to any of the inefficient CCGs. NHS North East Sussex CCG is the efficient unit that serves as a benchmark for the highest number of inefficient units (100), followed by NHS Hounslow CCG (68).

Table 10 - Pure Technical Efficiency results obtained with DEA.

DMU	PTE (%)	GP FTE {I}{V} (%)	Nurses FTE {I}{V} (%)	All direct patient care FTE {I}{V} (%)	Admin/Non-clinical FTE {I}{V} (%)	Count of appointments by mode: Face-to-face {O}{V} (%)	Count of appointments by mode: home visit {O}{V} (%)	Count of appointments by mode: other {O}{V} (%)	Benchmarks
NHS Bradford City CCG	100	74	13	12	0	100	0	0	3
NHS Knowsley CCG	100	54	28	18	0	73	21	6	10
NHS Birmingham and Solihull CCG	100	54	21	8	18	99	1	0	11
NHS Thanet CCG	100	100	0	0	0	45	55	0	0
NHS Leicester City CCG	100	69	25	7	0	80	0	20	6
NHS Wirral CCG	100	59	25	16	0	0	100	0	29
NHS North Kirklees CCG	100	39	19	8	34	100	0	0	25
NHS Swale CCG	100	83	17	0	0	82	7	11	2
NHS Leeds CCG	100	67	19	14	0	91	1	8	4
NHS Hastings and Rother CCG	100	79	21	0	0	71	16	13	45
NHS Calderdale CCG	100	51	30	19	0	83	3	14	35
NHS North East Essex CCG	100	93	4	3	0	96	4	0	100
NHS Ealing CCG	100	47	18	17	17	81	2	17	24
NHS North Cumbria CCG	100	100	0	0	0	43	47	10	1
NHS Hounslow CCG	100	40	13	9	39	87	0	13	68
NHS Thurrock CCG	100	49	28	24	0	77	0	23	18
NHS Bristol, North Somerset and South Gloucestershire CCG	100	100	0	0	0	10	88	2	14
NHS Southport and Formby CCG	100	70	30	0	0	0	100	0	0
NHS Cannock Chase CCG	100	35	17	11	37	95	5	0	44
NHS Crawley CCG	100	46	22	19	13	79	0	21	1
NHS Eastbourne, Hailsham and Seaford CCG	100	71	29	0	0	37	56	7	0
NHS West Lancashire CCG	100	26	14	10	51	0	100	0	10
NHS Ashford CCG	100	42	16	9	34	67	10	23	16
NHS Swindon CCG	100	43	13	7	37	88	0	11	2
NHS Canterbury and Coastal CCG	100	98	1	1	0	37	63	0	0
NHS Redditch and Bromsgrove CCG	100	2	12	10	49	73	24	4	9
NHS Cambridgeshire and Peterborough CCG	100	100	0	0	0	76	0	23	11
NHS Coastal West Sussex CCG	100	56	31	13	0	80	4	16	29
NHS West Leicestershire CCG	100	57	24	18	0	92	0	7	1
NHS North Hampshire CCG	100	40	18	14	27	59	14	27	37
NHS Herts Valleys CCG	100	66	20	14	0	72	14	14	1
NHS South Warwickshire CCG	100	31	12	12	46	68	15	17	8
NHS Surrey Heath CCG	100	33	10	11	46	12	85	4	13
NHS Wyre Forest CCG	99,50	39	12	8	41	53	22	26	50 (0.10) 107 (0.22) 109 (0.52) 161 (0.16)
NHS Richmond CCG	50,42	60	17	13	11	33	3	64	54 (0.02) 89 (0.23) 90 (0.31) 161 (0.45)

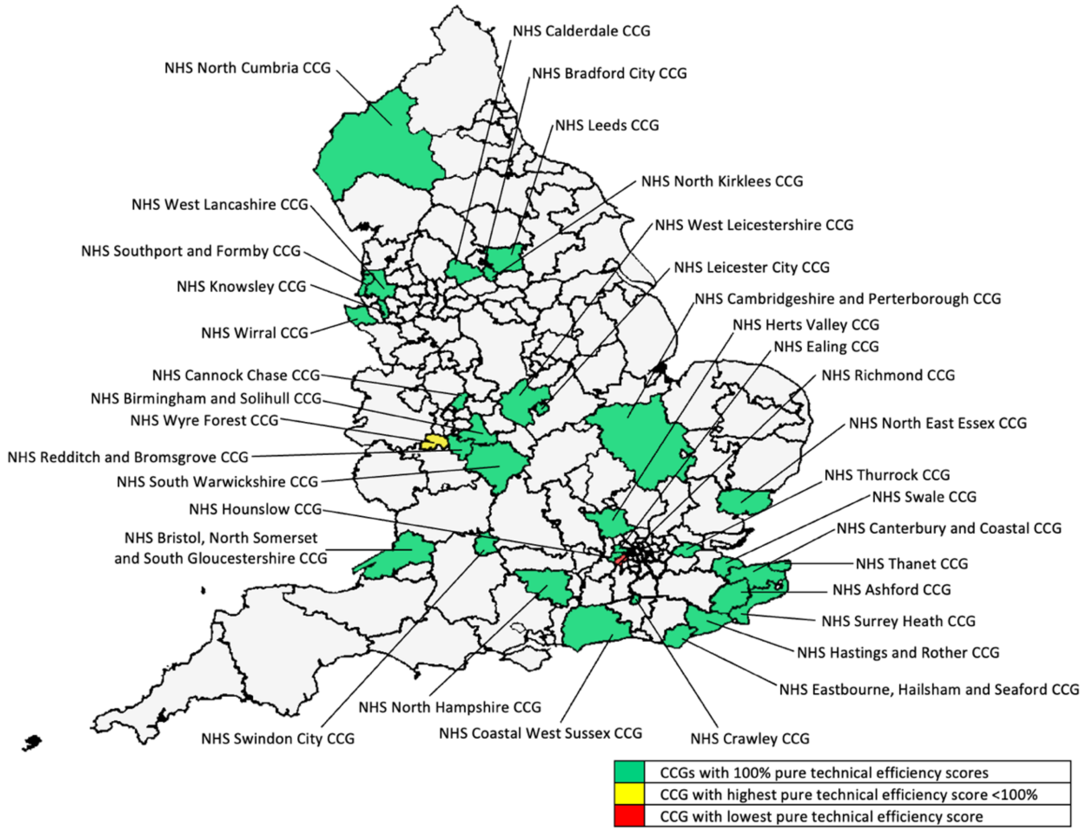

Figure 3 - Map of England representing Pure Technical Efficiency results obtained with DEA for Model 2.

After obtaining technical efficiency and pure technical efficiency through DEA, it is possible to calculate the result of scale efficiency, as shown in table 11 and figure 4. Of the 179 Clinical Commissioning Groups studied only 17 are operating at an optimal scale. NHS Scarborough and Ryedale CCG is the scale inefficient CCG with the highest score of scale efficiency (99,99%), while NHS Kernow CCG has the lowest score of scale efficiency (82,69%).

Table 11 - Scale Efficiency obtained with DEA.

DMU	SE
NHS Ashford CCG	100.00%
NHS Bradford City CCG	100.00%
NHS Calderdale CCG	100.00%
NHS Cannock Chase CCG	100.00%
NHS Canterbury and Coastal CCG	100.00%
NHS Hastings and Rother CCG	100.00%
NHS Hounslow CCG	100.00%
NHS North East Essex CCG	100.00%
NHS North Hampshire CCG	100.00%
NHS North Kirklees CCG	100.00%
NHS Redditch and Bromsgrove CCG	100.00%
NHS Southport and Formby CCG	100.00%
NHS Swindon CCG	100.00%
NHS Thanet CCG	100.00%
NHS Thurrock CCG	100.00%
NHS West Lancashire CCG	100.00%
NHS Wirral CCG	100.00%
NHS Scarborough and Ryedale CCG	99.99%
NHS Kernow CCG	82.69%

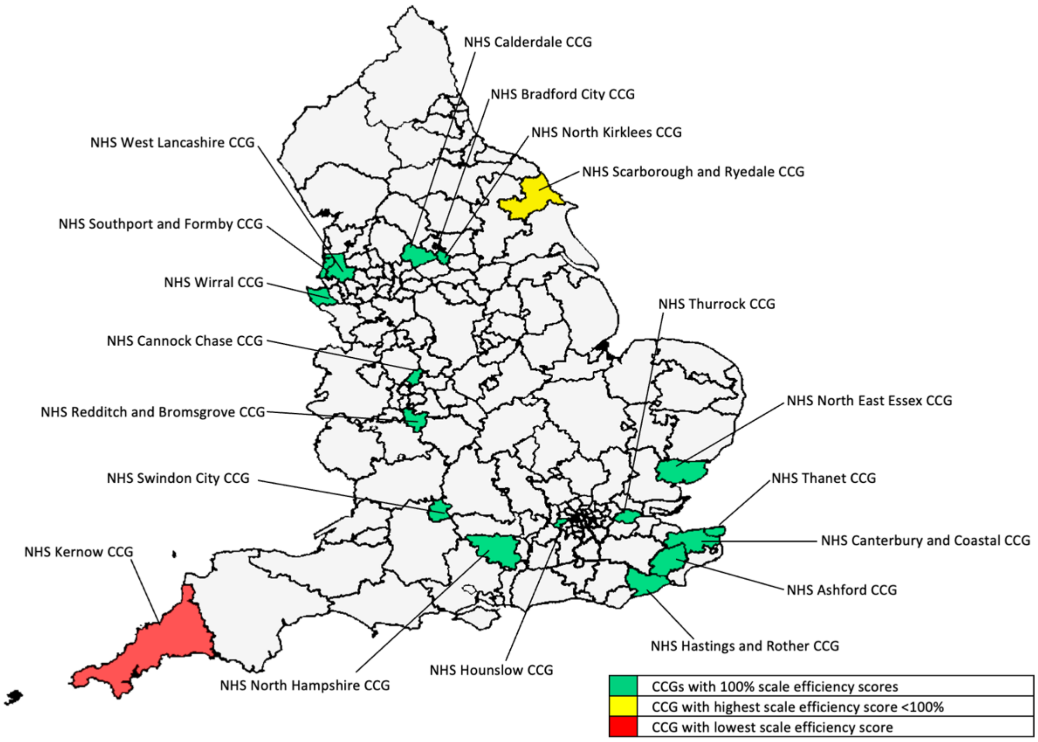

Figure 4 - Map of England representing scale efficiency results obtained with DEA for Model 2.

As demonstrated in table 12 of a total of 162 Clinical Commissioning Groups that are not operating at an optimal scale, 35 have increasing returns to scale, which means they are operating at a smaller than optimal scale. Conversely, 127 are operating at decreasing returns to scale meaning they are operating at a larger scale than optimal. When a unit is operating under increasing returns to scale it means that the output increases by a greater proportion than the increase in input, whereas a decreasing returns to scale occurs when outputs increase by a smaller proportion than the increase in inputs.

Table 12 - Returns to Scale results obtained with DEA.

CCGs operating under decreasing returns to scale		
NHS Barnet CCG	NHS Harrow CCG	NHS Sandwell and West Birmingham CCG
NHS Barnsley CCG	NHS Hartlepool and Stockton-on-Tees CCG	NHS Sheffield CCG
NHS Basildon and Brentwood CCG	NHS Havering CCG	NHS Shropshire CCG
NHS Bedfordshire CCG	NHS Herefordshire CCG	NHS Somerset CCG
NHS Berkshire West CCG	NHS Herts Valleys CCG	NHS South Cheshire CCG
NHS Birmingham and Solihull CCG	NHS Heywood, Middleton and Rochdale CCG	NHS South East Staffordshire and Seisdon Peninsula CCG
NHS Blackburn with Darwen CCG	NHS Hillingdon CCG	NHS South Eastern Hampshire CCG
NHS Bolton CCG	NHS Horsham and Mid Sussex CCG	NHS South Kent Coast CCG
NHS Bradford Districts CCG	NHS Hull CCG	NHS South Lincolnshire CCG
NHS Brent CCG	NHS Ipswich and East Suffolk CCG	NHS South Norfolk CCG
NHS Brighton and Hove CCG	NHS Kernow CCG	NHS South Sefton CCG
NHS Bristol, North Somerset and South Gloucestershire CCG	NHS Knowsley CCG	NHS South Tees CCG
NHS Bromley CCG	NHS Lambeth CCG	NHS South Tyneside CCG
NHS Buckinghamshire CCG	NHS Leeds CCG	NHS South Warwickshire CCG
NHS Cambridgeshire and Peterborough CCG	NHS Leicester City CCG	NHS South Worcestershire CCG
NHS Chorley and South Ribble CCG	NHS Lewisham CCG	NHS Southampton CCG
NHS City and Hackney CCG	NHS Lincolnshire East CCG	NHS Southwark CCG
NHS Coastal West Sussex CCG	NHS Lincolnshire West CCG	NHS St Helens CCG
NHS Coventry and Rugby CCG	NHS Liverpool CCG	NHS Stockport CCG
NHS Croydon CCG	NHS Manchester CCG	NHS Stoke on Trent CCG
NHS Doncaster CCG	NHS Mansfield and Ashfield CCG	NHS Sunderland CCG
NHS Dorset CCG	NHS Medway CCG	NHS Surrey Downs CCG
NHS Dudley CCG	NHS Mid Essex CCG	NHS Tameside and Glossop CCG
NHS Durham Dales, Easington and Sedgefield CCG	NHS Milton Keynes CCG	NHS Tower Hamlets CCG
NHS Ealing CCG	NHS Morecambe Bay CCG	NHS Trafford CCG
NHS East and North Hertfordshire CCG	NHS Nene CCG	NHS Vale of York CCG
NHS East Berkshire CCG	NHS Newcastle Gateshead CCG	NHS Wakefield CCG
NHS East Lancashire CCG	NHS Newham CCG	NHS Walsall CCG
NHS East Leicestershire and Rutland CCG	NHS North Cumbria CCG	NHS Waltham Forest CCG
NHS East Riding of Yorkshire CCG	NHS North Durham CCG	NHS Wandsworth CCG
NHS East Surrey CCG	NHS North East Hampshire and Farnham CCG	NHS Warrington CCG

NHS Eastbourne, Hailsham and Seaford CCG	NHS North Lincolnshire CCG	NHS West Cheshire CCG
NHS Eastern Cheshire CCG	NHS North Norfolk CCG	NHS West Essex CCG
NHS Enfield CCG	NHS North Staffordshire CCG	NHS West Hampshire CCG
NHS Fareham and Gosport CCG	NHS North Tyneside CCG	NHS West Kent CCG
NHS Fylde and Wyre CCG	NHS North West Surrey CCG	NHS West Leicestershire CCG
NHS Gloucestershire CCG	NHS Northumberland CCG	NHS West Norfolk CCG
NHS Great Yarmouth and Waveney CCG	NHS Norwich CCG	NHS West Suffolk CCG
NHS Greater Huddersfield CCG	NHS Nottingham City CCG	NHS Wigan Borough CCG
NHS Greater Preston CCG	NHS Oldham CCG	NHS Wiltshire CCG
NHS Greenwich CCG	NHS Redbridge CCG	NHS Wolverhampton CCG
NHS Guildford and Waverley CCG	NHS Rotherham CCG	
NHS Haringey CCG	NHS Salford CCG	

CCGs operating under increasing returns to scale	
NHS Airedale, Wharfedale and Craven CCG	NHS Newark and Sherwood CCG
NHS Barking and Dagenham CCG	NHS North East Lincolnshire CCG
NHS Bassetlaw CCG	NHS Nottingham North and East CCG
NHS Bath and North East Somerset CCG	NHS Richmond CCG
NHS Bexley CCG	NHS Rushcliffe CCG
NHS Blackpool CCG	NHS Scarborough and Ryedale CCG
NHS Central London (Westminster) CCG	NHS South West Lincolnshire CCG
NHS Crawley CCG	NHS Southend CCG
NHS East Staffordshire CCG	NHS Stafford and Surrounds CCG
NHS Halton CCG	NHS Surrey Heath CCG
NHS Hambleton, Richmondshire and Whitby CCG	NHS Sutton CCG
NHS Hammersmith and Fulham CCG	NHS Swale CCG
NHS Harrogate and Rural District CCG	NHS Telford and Wrekin CCG
NHS High Weald Lewes Havens CCG	NHS Vale Royal CCG
NHS Isle of Wight CCG	NHS Warwickshire North CCG
NHS Islington CCG	NHS West London CCG
NHS Kingston CCG	NHS Wyre Forest CCG
NHS Merton CCG	

Of the units that are operating under decreasing returns to scale, NHS Leeds CCG, NHS Birmingham and Solihull CCG, NHS Bristol, North Somerset and South Gloucestershire CCG, NHS Buckinghamshire CCG, NHS Berkshire West CCG, and NHS East Berkshire CCG, were Clinical Commissioning Groups that were created in 2018 resulting from mergers of other Clinical Commissioning Groups, as demonstrated in table 13.

Table 13 - Clinical Commissioning Groups mergers in 2018.

Mergers in 2018	New CCG	Efficiency in 2019 of the new CCG		Returns to scale
NHS Leeds North CCG NHS Leeds South and East CCG NHS Leeds West CCG	NHS Leeds CCG	TE	86,55%	DRS
		PTE	100%	
		SE	86,55%	
NHS Birmingham South and Central CCG NHS Birmingham CrossCity CCG NHS Solihull CCG	NHS Birmingham and Solihull CCG	TE	86,76%	DRS
		PTE	100%	
		SE	86,76%	
NHS South Gloucestershire CCG NHS Bristol CCG NHS North Somerset CCG	NHS Bristol, North Somerset, and South Gloucestershire CCG	TE	86,28%	DRS
		PTE	100%	
		SE	86,28%	
NHS Aylesbury CCG NHS Chiltern CCG	NHS Buckinghamshire CCG	TE	79,22%	DRS
		PTE	86,69%	
		SE	91,39%	
NHS Newbury and District CCG NHS North and West Reading CCG NHS South Reading CCG NHS Wokingham CCG	NHS Berkshire West CCG	TE	74,99%	DRS
		PTE	83,51%	
		SE	89,79%	
NHS Bracknell and Ascot CCG NHS Slough CCG NHS Windsor, Ascot and Maidenhead CCG	NHS East Berkshire CCG	TE	83,69%	DRS
		PTE	88,65%	
		SE	94,40%	

In table 14 mergers that occurred in the calendar year 2020 are presented. The efficiency results calculated in the second DEA model – TE, PTE and SE for the Clinical Commissioning Groups with data from 2019, provide interesting information regarding the future mergers and the choice of CCGs to merge. There were 18 new CCGs created in 2020 resulting from merging 74 CCGs.

Table 14 - Clinical Commissioning Group mergers in 2020.

Merging CCGs in 2020	TE in 2019 (%)	PTE in 2019 (%)	SE in 2019 (%)	Returns to Scale	New CCG
NHS Bradford City CCG	**100**	**100**	**100**	**CRS**	NHS Bradford District and Craven CCG
NHS Bradford Districts CCG	86,96	91,47	95,07	DRS	
NHS Airedale, Wharfedale and Craven CCG	88,82	89,33	99,42	IRS	
NHS Hambleton, Richmondshire and Whitby CCG	75,01	75,88	98,86	IRS	NHS North Yorkshire CCG
NHS Harrogate and Rural District CCG	80,82	81,47	99,20	IRS	
NHS Scarborough and Ryedale CCG	80,57	80,57	99,99	IRS	
NHS Darlington CCG (outlier)					NHS Tees Valley CCG
NHS Hartlepool and Stockton-on-Tees CCG	84,75	87,69	96,65	DRS	
NHS South Tees CCG	88,90	92,94	95,66	DRS	
NHS Durham Dales, Easington and Sedgefield CCG	90,29	91,89	98,26	DRS	NHS County Durham CCG
NHS North Durham CCG	93,26	94,35	98,84	DRS	
NHS Vale Royal CCG	80,44	90,05	90,80	IRS	NHS Cheshire CCG
NHS South Cheshire CCG	96,17	97,04	99,11	DRS	
NHS Eastern Cheshire CCG	72,12	73,23	98,49	DRS	
NHS West Cheshire CCG	83,17	87,13	95,46	DRS	
NHS Mansfield and Ashfield CCG	87,01	87,18	99,81	DRS	NHS Nottingham and Nottinghamshire CCG
NHS Newark and Sherwood CCG	85,27	86,75	98,30	IRS	
NHS Nottingham City CCG	78,82	82,98	94,99	DRS	
NHS Nottingham North and East CCG	81,55	84,13	96,93	IRS	
NHS Nottingham West CCG (outlier)					
NHS Rushcliffe CCG	83,02	85,25	97,38	IRS	
NHS South Worcestershire CCG	80,86	86,45	93,54	DRS	NHS Herefordshire and Worcestershire CCG
NHS Redditch and Bromsgrove CCG	**100**	**100**	**100**	**CRS**	
NHS Wyre Forest CCG	99,42	99,50	99,92	IRS	
NHS Herefordshire CCG	83,08	84,21	98,66	DRS	
NHS South Lincolnshire CCG	84,53	84,98	99,48	DRS	NHS Lincolnshire CCG
NHS South West Lincolnshire CCG	80,73	81,41	99,17	IRS	
NHS Lincolnshire West CCG	78,66	79,06	99,50	DRS	
NHS Lincolnshire East CCG	78,93	83,04	95,06	DRS	
NHS Corby CCG (outlier)					NHS Northamptonshire CCG
NHS Nene CCG	81,85	91,77	89,20	DRS	
NHS Great Yarmouth and Waveney CCG	76,97	79,40	96,94	DRS	NHS Norfolk and Waveney CCG
NHS North Norfolk CCG	70,47	74,01	95,21	DRS	
NHS Norwich CCG	85,43	86,29	99,01	DRS	
NHS South Norfolk CCG	77,32	79,10	97,75	DRS	
NHS West Norfolk CCG	83,66	84,95	98,48	DRS	
NHS Swindon CCG	**100**	**100**	**100**	**CRS**	NHS Bath and North East Somerset, Swindon and Wiltshire CCG
NHS Wiltshire CCG	82,09	88,99	92,24	DRS	
NHS Bath and North East Somerset CCG	86,51	87,03	99,40	IRS	
NHS Coastal West Sussex CCG	93,50	**100**	93,50	DRS	NHS West Sussex CCG
NHS Crawley CCG	92,30	**100**	92,30	IRS	
NHS Horsham and Mid Sussex CCG	87,19	87,41	99,75	DRS	
NHS Surrey Downs CCG	91,79	95,39	99,77	DRS	NHS Surrey Heartlands CCG
NHS East Surrey CCG	96,45	96,56	99,88	DRS	
NHS Guildford and Waverley CCG	85,98	86,03	99,94	DRS	
NHS North West Surrey CCG	85,40	86,54	98,69	DRS	
NHS Ashford CCG	**100**	**100**	**100**	**CRS**	NHS Kent and Medway CCG
NHS Canterbury and Coastal CCG	**100**	**100**	**100**	**CRS**	
NHS Dartford, Gravesham and Swanley CCG(outlier)					
NHS Medway CCG	73,60	75,43	97,57	DRS	
NHS South Kent Coast CCG	58,96	59,63	98,89	DRS	
NHS Swale CCG	97,46	**100**	97,46	IRS	
NHS Thanet CCG	**100**	**100**	**100**	**CRS**	
NHS West Kent CCG	71,32	79,51	89,70	DRS	
NHS Eastbourne, Hailsham and Seaford CCG	99,12	**100**	99,12	DRS	NHS East Sussex CCG
NHS Hastings and Rother CCG	**100**	**100**	**100**	**CRS**	
NHS High Weald Lewes Havens CCG	69,92	70,08	99,76	IRS	
NHS Barnet CCG	62,64	64,78	96,70	DRS	NHS North Central London CCG
NHS Camden CCG (outlier)					
NHS Enfield CCG	75,59	75,99	99,46	DRS	
NHS Haringey CCG	76,70	76,96	99,66	DRS	
NHS Islington CCG	78,90	79,08	99,77	IRS	
NHS Wandsworth CCG	84,65	89,10	95,00	DRS	NHS South West London CCG
NHS Sutton CCG	91,01	92,26	98,64	IRS	
NHS Richmond CCG	50,37	50,42	99,89	IRS	
NHS Merton CCG	77,30	78,34	98,67	IRS	
NHS Kingston CCG	81,75	83,56	97,83	IRS	
NHS Croydon CCG	77,02	80,46	95,73	DRS	
NHS Southwark CCG	86,81	89,01	97,53	DRS	NHS South East London CCG
NHS Greenwich CCG	59,59	60,32	98,78	DRS	
NHS Lambeth CCG	76,48	81,10	94,30	DRS	
NHS Lewisham CCG	77,98	78,92	98,81	DRS	
NHS Bexley CCG	52,20	52,30	99,81	IRS	
NHS Bromley CCG	84,42	85,54	98,70	DRS	

Analysing the information provided in this table, we can see that there are a combination of different strategies applied to these mergers. Some have joined units with increasing returns to scale (smaller than optimal) with ones that have decreasing returns to scale (larger than optimal) such as NHS Cheshire CCG, NHS Nottingham and Nottinghamshire CCG, NHS Lincolnshire CCG, NHS West Sussex CCG, NHS North Central London CCG, NHS South West London CCG and NHS South East London CCG. Others merge units that are operating under constant returns to scale (optimal size) with units that are not scale efficient such as NHS Bradford District and Craven CCG, NHS Herefordshire and Worcestershire CCG, NHS Bath and North East Somerset, Swindon and Wiltshire CCG, NHS Kent and Medway CCG and NHS East Sussex CCG. Few of the mergers, such as NHS Tees Valley CCG, NHS County Durham CCG, NHS Norfolk and Waveney CCG and NHS Surrey Heartlands CCG take place between units that are all DRS, which may indicate that other types of efficiency were being pursued in detriment of scale efficiency.

Another interesting choice of merger is to combine units that have pure technical efficiency with others operating less efficiently, which occurs in NHS West Sussex CCG and NHS East Sussex CCG. The advantage in this strategy is that optimal management would now be applied to all the prior CCGs, intending to create a larger group with high pure technical efficiency.

A factor that limits the interpretation of this table, is the presence of CCGs that were excluded from our analysis as they were identified as outliers. As mentioned previously, further analyses into these mergers, the rationale and new strategic methods implemented, are necessary.

The <u>third DEA model</u> aimed to measure service effectiveness in reducing hospital admissions and maximizing patient satisfaction. Effectiveness refers to how close the results (observed outputs) came to the objectives. Table 15 demonstrates descriptive statistics for the variables used in the third DEA model. In this model, activity indicators were used as inputs. The number of patients that rate an overall good experience, output variable, is a quality indicator and the number of emergency admissions saved is a health outcome indicator.

Looking at the output variable "Number of patients that rate an overall good experience", there is a large discrepancy between the minimum number of patients with a positive experience in a CCG and the highest number of patients that rate a good experience in a CCG. Ideally, primary health care should provide patients with a high level of satisfaction being the first point of contact with the health system. A high level of satisfaction would indicate that the needs faced by these patients are being met. The output variable "Emergency admissions saved when compared to the worst observed"[2] is approximately 26044 in the CCG that has the highest count of emergency admissions saved, whereas the minimum level of admissions saved is approximately 572 and corresponds to the CCG that presents the worst performance in terms of number of emergency admissions.

In the third model, three weight restrictions were implemented (presented in appendix 2), following the production trade-offs method proposed by Podinovski (2004). These weight restrictions indicate trade-offs that could occur between the three inputs (appointment modes) without detrimentally affecting the outputs.

Table 15 - Descriptive statistics for the variables in the third DEA model.

Variables	Average	Standard deviation	Max	Min
Count of appointments by mode: Face-to-face {I}	1243481,29	759480,85	5754826	202280
Count of appointments by mode: home visit {I}	14798,72	15486,45	98799	117
Count of appointments by mode: other {I}	257201,07	191647,48	1280326	34047
Number of patients that rate an overall good experience {O}	254882,99	148104,34	1062256	82852
Emergency admissions saved when compared to the worst observed {O}	5025,70	3798,85	26044	572

In table 16 and figure 5, the results for this DEA model are illustrated. There are only 5 CCGs that are effective in using their inputs to maximise the desired outputs, in this case to transform the various forms of appointments into patient satisfaction and emergency admissions saved: NHS Barnet CCG, NHS Basildon and Brentwood CCG, NHS Greenwich CCG, NHS Mansfield and Ashfield CCG and NHS Richmond CCG.

[2] The variable Emergency admissions saved when compared to the worst observed was calculated as follows: to the maximum number of emergency admissions per registered patient observed in the sample, the number of emergency admissions per registered patient observed in the CCG analysed is subtracted, then this result is multiplied by the number of registered patients in that CCG.

The average effectiveness score was 59,19%. Of the ineffective units, NHS Bexley CCG scored the highest with 98,97% and NHS Wyre Forest CCG scored the lowest with 35,20%.

Table 16 demonstrates the optimal weight profiles, of the effective CCGs only NHS Mansfield and Ashfield CCG attribute total virtual input weight to one variable – Count of appointments by mode: home visit. The other 4 effective CCGs distribute the proportion of weight by two or three of the input variables in their optimal input weight structure, simultaneously all four attribute highest weight to the variable – Count of appointments by mode: face-to-face.

In terms of the optimal output weight profiles, all of the effective CCGs attribute the total output weight on one of the two output variables.

Table 16 - Service Effectiveness results obtained with DEA.

DMU	Effectiveness (%)	Count of appointments by mode: Face-to-face {I}{V} (%)	Count of appointments by mode: home visit {I}{V} (%)	Count of appointments by mode: other {I}{V} (%)	Number of patients that rate an overall good experience {O}{V} (%)	Emergency admn saved when compared to the worst observed {O}{V} (%)	Benchmarks
NHS Barnet CCG	100	78	6	16	0	1000	40
NHS Basildon and Brentwood CCG	100	56	33	11	0	100	9
NHS Greenwich CCG	100	74	26	0	100	0	84
NHS Mansfield and Ashfield CCG	100	0	100	0	100	0	0
NHS Richmond CCG	100	98	2	0	100	0	152
NHS Bexley CCG	98,9	59	1	40	100	0	117 (1.02)
NHS Wyre Forest CCG	35,2	66	2	32	100	0	117 (1.47)

Regarding benchmarks, NHS Mansfield and Ashfield CCG does not serve as a benchmark to the ineffective units. NHS Richmond CCG serves as a benchmark to the highest number of CCGs (152), followed by NHS Greenwich CCG (84).

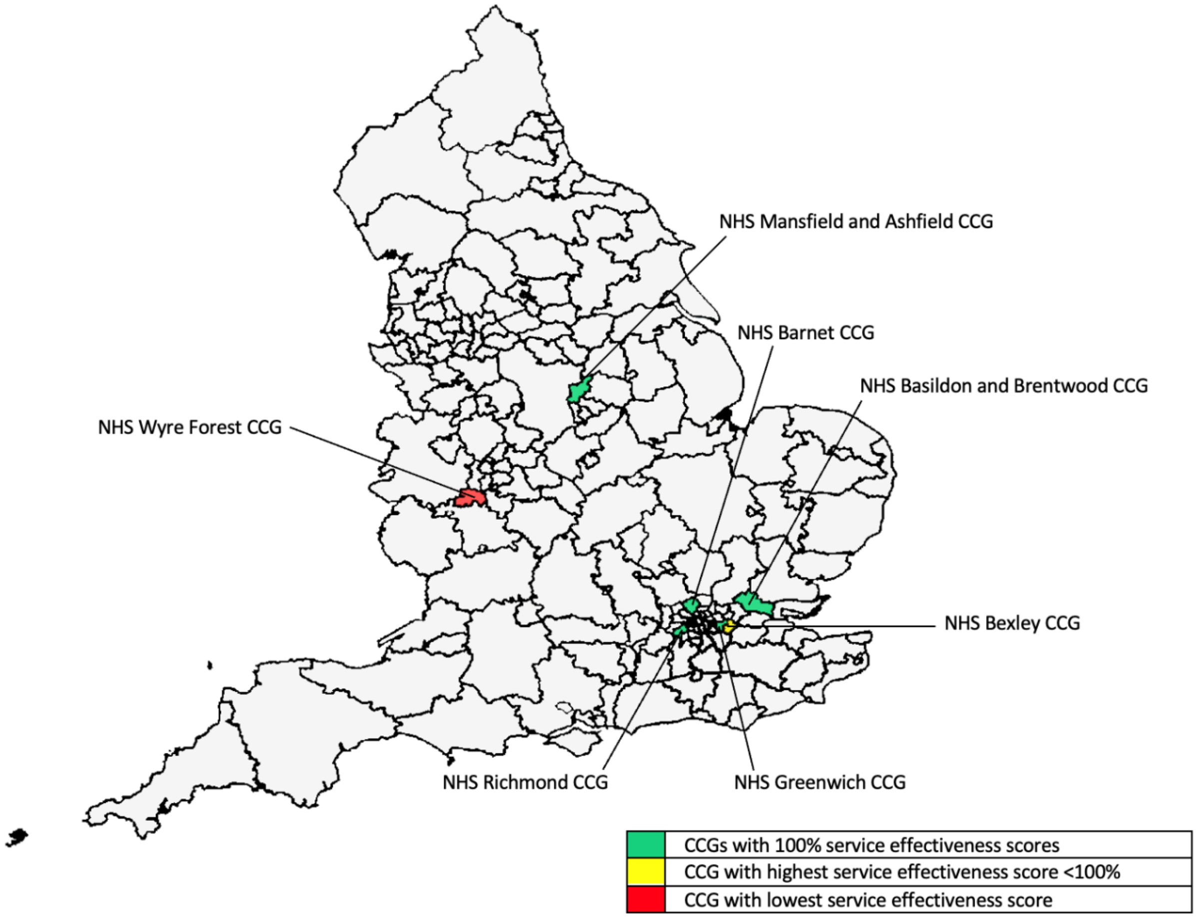

Figure 5 - Map of England representing service effectiveness results obtained with DEA for Model 3.

3.3. Malmquist Indices data and results

In Camanho and Dyson (2006), the authors cluster their DMUs according to regions, here, in this dissertation, the common factor used to group DMUs into two groups was the Index of Multiple Deprivation average scores. After dividing the 179 CCGs into two groups – first group: higher IMD score (89 CCGs) and second group: lower IMD score (90 CCGs), the three models described previously where run using the EMS Software with a constant returns to scale approach as suggested by Camanho and Dyson (2006). In consistency with the analysis presented before, an output orientation was used for the three models. The same weight restrictions described previously for each model were applied. The results obtained allowed us to calculate the values of efficiency spread (IE^{AB}), frontier shift (IF^{AB}), and overall performance (I^{AB}) for each model.

For the three models the aim was to assess if deprivation affects equity, efficiency and effectiveness levels. The results obtained from DEA models assuming CRS, with an output orientation are as follows in table 17.

Table 17 - Results obtained for the three DEA models using the MI by Camanho and Dyson (2006).

Malmquist Indices	***Model 1*** *Human resource distribution equity based on registered patients list, by age group.*	**Model 2** *Service efficiency*	**Model 3** *Service effectiveness*
Efficiency Spread (IE^{AB})	1.02	0.99	1.02
Frontier Shift (IF^{AB})	1.06	1.02	0.98
Overall Performance (I^{AB})	1.08	1.01	1.00

For the first model the result of efficiency spread is 1.02, which implies that equity scores are more consistent (less dispersion) in the most deprived CCGs compared to the least deprived group. In other words, in the most deprived areas, the average resource equity level observed is higher than the average observed in less deprived areas. The value of frontier shift is 1.06, which means that there is dominance of the best practice frontier in the most deprived areas. The frontier of access to resources of the most deprived CCGs is 6% higher than the least deprived CCGs frontier, considering the level of needs. Finally, the overall performance is 1.08, which means that, on average, the available human resources, when accounting for needs, is 8% higher in the most deprived areas than in the least deprived areas. These results emphasize a positive discrimination, where there is an overcompensation of human resources in the most deprived areas in detriment of the least deprived areas. This is consistent to what would be expected.

In table 17, we can see that the efficiency spread for the <u>second model</u> is 0.99, which reflects lower service efficiency level consistency and greater dispersion in the most deprived CCGs. These results demonstrate lower service efficiency levels in the most deprived areas when compared to least deprived areas. The value of frontier shift is 1.02, which means that the best practice frontier of most deprived CCGs present 2% higher productivity than the best practice frontier of the least deprived CCGs. The overall performance is 1.01, which indicates that, on average, service efficiency, according to the number of appointments, in the most deprived areas is 1% higher than the level observed in the least deprived areas.

In the <u>third model</u>, the effectiveness spread is 1.02, which, once again, reflects higher service effectiveness level consistency (lower dispersion) in the most deprived CCGs. The value of the frontier shift is 0.98, which indicates that the most deprived productivity frontier is lower level than that observed for the least deprived areas. In specific, the most deprived CCGs frontier is only at 98% of the level observed in least deprived areas. This reflects that service effectiveness levels in most deprived CCGs is only slightly lower than in the least deprived CCGs. Finally, the overall performance for the third model is 1, which means that, on average, service effectiveness in reducing emergency admissions and maximizing patient satisfaction is equal in most deprived CCGs than in the least deprived CCGs.

The results of the Malmquist Indices reflect that the health and social policies in place in the United Kingdom to reduce health inequalities caused by deprivation are determining similar or slightly better primary health care performance in most deprived areas.

Takundwa, et al. (2017) identified a negative relationship between deprivation and efficiency, however regression analysis showed that this relationship was not statistically significant at a CCG-level. Additionally, Takundwa, et al. (2017) had suggested that the NHS England's adjustment for environmental factors within CCG-level budget allocation had been broadly successful. The findings in this study, provide evidence that NHS England's effort to reduce health inequalities caused by deprivation are being successful.

3.4. The practical and policy implications of the results

Despite the exploratory nature of this study, there are important considerations that can be derived from the results previously presented. The results from this research paper can be divided into two components: performance measurement of primary health care in the NHS using Data Envelopment Analysis and measurement of the effect of deprivation on primary health care performance using Malmquist Indices.

The performance results obtained through Data Envelopment Analysis assess equity, service efficiency and service effectiveness. The equity results identify that only 13 of the 179 CCGs benefit from an equitable human resource distribution based on registered patients list, by age group. Improvement is needed to reach the equity of human resources among CCGs. Although CCG budget allocation is calculated based on health needs of that population (of which age is one criteria), having an adequate budget may not equate to sufficient health professionals. In the UK, there is a significant shortage of health care staff, for which new governmental strategies and increased effort are needed to recruit and retain skilled health professionals. Another option would be to relocate some health professionals to reach equitable distribution.

Service efficiency results based on staff and services produced, obtained through the DEA technique, identified that 17 CCGs were technically efficient, pure technically efficient and scale efficient. However, 53 CCGs had technical efficiency levels under 80%, 37 had scores of pure technical efficiency under 80%, which require improvement. NHS Hounslow CCG, NHS North Hampshire CCG and NHS Calderdale CCG serve as benchmarks to a high number of inefficient units. A practical way to improve efficiency levels of the inefficient units would be to promote learning networks and discussion sessions between them and their respective benchmarks. In these discussion sessions, work methods, staffing and roles, management and policies should be discussed, for inefficient units to learn from and follow the efficient units. It is important to emphasize that primary health care is part of the NHS, which was founded to provide equal care to equal needs, for this to happen, the system should work united to reach this goal. Even though CCGs divide England in areas, they are all part of a public service, for which there should be a joint effort to improve the results of the areas that are not reaching their optimal efficiency. The more efficient the health system is as a total, the better value for money the taxpayer will receive.

In tables 13 and 14, we analysed the mergers that happened in 2018, for which it was possible to calculate efficiency levels for the year 2019, and the mergers that followed in

2020, where we can provide insightful information that could predict the viability of these. Table 14 looks in further detail into the units that were merged and it provides information that policy makers and managers could have used in their decision-making process. This method, once again, shows to be useful and could be used to sustain the decision to merge units and the viability of the future units. As mentioned previously, it would be highly relevant to reproduce these models with data from 2020, including the new CCGs that were created and assess the effect these mergers have had on equity, efficiency and effectiveness levels.

Only 5 CCGs were service effective in reducing emergency admissions and maximizing patient satisfaction, which is worrying as primary health care is the first point of access to the health system. If emergency admissions are not reduced through effective primary health care more patients will require secondary and tertiary care, which ultimately is more costly to the taxpayer and may cause delayed treatment, which consequently worsens health outcomes for the population. Effective primary health care should include health monitoring and disease screening throughout the life cycle, inoculations, prenatal care, chronic illness management, and more. If these basic activities are not being carried out in an effective manner it will lead to emergency admissions to hospital. Simultaneously, promoting a positive patient experience is essential for the patients to access and utilise primary health care services as a first choice and to reach the desired health outcomes. To improve service effectiveness governmental action is required to create better health pathways and encourage primary health care utilisation. At a more local level, ineffective CCGs may benefit from learning from the effective ones, as mentioned previously. The effective CCGs that serve as a benchmark to the highest number of ineffective units are NHS Richmond CCG and NHS Greenwich CCG. Meetings should be held between these units and the ineffective ones to promote learning and improvement.

The results obtained through the Malmquist Indices provide interesting information on the effect of deprivation on primary health care performance. The results for the first model show that there is a positive discrimination, where the most deprived CCGs are receiving more than equitable human resource allocation. On the other hand, the second and third models show that deprivation is being well addressed with this extra resource allocation, as the most deprived areas have higher service efficiency and equal service effectiveness when compared to the least deprived areas. The IMD is used in the formula to calculate CCG budgets, thus these results demonstrate this way of allocating resources is reaching it's objective. However, the NHS Long Term Plan in 2019 informed that the allocation formula was going to be reviewed, aiming to be more responsive to health inequalities and unmet need. Future analysis, using this methodology with data from 2020 and 2021, would be advantageous to assess the effect of the changes to these formulae and if the Long-Term Plan is meeting its' goals.

The two techniques have demonstrated to be valuable tools to assess and improve performance in primary health care. Health care managers and policy makers can identify which CCGs are being most successful in transforming limited resources into precious health outcomes. Not only will they be able to identify which are successful and which are least successful, these techniques also allow them to understand from which successful units it is more beneficial to learn from. Although these techniques provide results from the past, they can contribute to improve the future.

Despite this research paper only covering the calendar year of 2019 and excluding some CCGs, the results achieved provide useful information for policy making, which can contribute to improving access, utilization and outcomes of primary health care in the NHS.

4. Conclusion

Health inequalities have been identified as undesirable and detrimental to the health of a population, nonetheless, the NHS England recognizes that they are a central concern and concentrate efforts to reduce them. The results of this study demonstrate that the policies addressing deprivation that were in place in 2019 were having a positive effect on primary health care performance.

Through the literature review it was possible to conclude that DEA is a commonly used technique to assess performance in health care systems and more specifically in primary health care, however, to the best of our knowledge, there are no published studies that use the Malmquist Indices to contrast the performance of primary health care providers according to the level of deprivation. Therefore, the methodology carried out in this study is innovative.

Despite the exploratory nature of this research, there are some relevant empirical findings from this study. Initially three DEA models were run to assess resource allocation equity, service efficiency and service effectiveness among CCGs. All the models produced relevant results identifying the units that had best performance and those with the lowest. Simultaneously, this technique allows the identification of learning peers (benchmarks), which provides useful information regarding the best practices to learn from to improve the overall performance of PHC providers in the NHS. By identifying the lowest performing CCGs, it creates an opportunity to investigate the causes of this poor performance within these CCGs and compare their practices against better performing CCGs. It is important to remember that health service providers within the NHS should work together to reach the best health outcomes for the national population, therefore a collaborative effort should be applied to improve the performance of all the primary health care system.

In the second model, scale efficiency was calculated for the CCGs, then we investigated the previous (2018) and latter (2020) mergers that occurred to the CCGs within England. This analysis provides a possible method that future mergers could follow.

Of the three aspects of performance (equity, efficiency and effectiveness) measured through the Data Envelopment Analysis in this study, service efficiency was the aspect that presented best performance results among the CCGs studied, albeit all three aspects require significant improvement.

The second technique applied – the Malmquist Indices, allowed us to contrast the performance of the group with highest level of deprivation with the group with lowest level of deprivation. This method provided very interesting results, confirming that the health and social policies addressing deprivation that were in place in 2019 were in fact having a positive effect on performance in all three performance aspects studied in this paper.

One of the main limitations to this research was data availability. As mentioned previously, the data used in this study was sourced from the NHS Digital website. The initial plan was to focus on a wider time interval from 2019-2021 to measure the effect of deprivation on primary health care performance over time, however, due to data unavailability and time constraints, the focus had to be turned to the one full year of data that was available. This dynamic analysis would have been very interesting to study the effect that the COVID-19 pandemic had on performance levels in England, considering deprivation. Despite this, the research that has been completed in this paper, demonstrates an innovative and useful method to measure performance and the effect of deprivation on it. This paper discusses a reproducible method that can be applied with other data sets of other calendar years and even a different common factor to distinguish group performance (besides deprivation).

Creating valid weight restrictions was also a challenging step of this research, as many of the authors mentioned in the literature review do not provide an insight regarding the formulation of these restrictions. Another demanding detail to this study was that most CCG health outcomes presented on the NHS Digital website were undesirable health outcomes, such as Emergency admissions for acute conditions that should not need hospital admissions, which implicated a transformation into a desirable outcome. Despite some challenges, we consider that the goals set out for this research have been reached.

It would be interesting to run these analyses again with more recent data, as since the data used in this study was collected, the long term plan has been in place for four years (since 2019) and in July 2022 CCGs were closed down and Integrated Care Systems (ICS) were created. It would be of high interest to assess the results of these changes in policy and structure of primary health care in the NHS. Furthermore, it would be beneficial to interview policy-makers and managers within primary health care in the NHS to receive their feedback on these results and regarding how to account for them in their decisions and policies. This is a subject for future research.

5. References

Amado, C. A. F. & Dyson, R. G., 2009. Exploring the use of DEA for formative evaluation in primary diabetes care: An application to compare English practices. *Journal of the Operational Research Society,* Volume 60, pp. 1469-1482.

Amado, C. A. E. F. & Santos, S. P., 2009. Challenges for performance assessment and improvement in primary health care: The case of Portuguese health centres. *Health Policy,* Volume 91, pp. 43-56.

Amado, C. A. F., Barreira, A. P., Santos, S. P. & Guimarães, M. H., 2019. Comparing the quality of life of cities that gained and lost population: An assessment with DEA and the Malmquist index. *Papers in Regional Science,* 98(5), pp. 2075-2097.

Andersen, P. & Petersen, N.C., 1993. A Procedure for Ranking Efficient Units in Data Envelopment Analysis. *Management Science*, 39, pp. 1261-1264.

Banker, R. & Chang, H., 2006. The super-efficiency procedure for outlier identification, not for ranking efficient units. *European Journal of Operational Research,* pp. 1311-1320.

Banker, R. D., Charnes, A. & Cooper, W. W., 1984. Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. *Management Science,* 30(9), pp. 1078-1092.

Camanho, A. S. & Dyson, R. G., 2006. Data Envelopment Analysis and Malmquist Indices for Measuring Group Performance. *Journal of Productivity Analysis ,* 26(1), pp. 35-49.

Campbell, S. M., Braspenning, J., Hutchinson, A. & Marshall, M., 2002. Research methods used in developing and applying quality indicators in primary care. *Quality and Safety in Health Care ,* Volume 11, pp. 358-364.

Caves, D. W., Christensen, L. R. & Diewert, W. E., 1982. The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity. *Econometrica,* 50(6), pp. 1393-1414.

Charnes, A., Cooper, W. & Rhodes, E., 1978. Measuring the efficiency of decision making units. *European Journal of Operational Research,* 2(6), pp. 429-444.

Cinaroglu, S., 2020. Integrated K-means clustering with data envelopment analysis of public hospital efficiency. *Health Care Management Science,* Volume 23, pp. 325–338.

Cooper, W., Seiford, L. & Tone, K., 2007. *DATA ENVELOPMENT ANALYSIS: A Comprehensive Text with Models, Applications, References, And DEA-Solver Software*, 2nd edn. Springer Science & Business Media, New York.

Cylus, J., Papanicolas, I. & Smith, P. C., 2016. Health system efficiency How to make measurement matter for policy and management. *Health Policy Series,* Issue 46.

Deidda, M., Lupiáñez-Villanueva, F., Codagnone, C. & Maghiros, I., 2014. Using Data Envelopment Analysis to Analyse the Efficiency of Primary Care Units. *Journal of Medical Systems ,* 38(122).

Department of Health and Social Care, 2023. *Gov.uk.* [Online] Available at: https://www.gov.uk/government/publications/the-nhs-constitution-for-england/the-nhs-constitution-for-england#introduction-to-the-nhs-constitution [Acedido em 4 August 2023].

Dlouhy, M., 2018. Measuring Geographic Inequalities: Dealing with Multiple Health Resources by Data Envelopment Analysis. *Frontiers in Public Health,* 6(53).

Dyson, R. G.,Allen, R., Camanho, A. S., Podinovski, V. V., Sarrico, C .S. & Shale, E. A., 2001. Pitfalls and protocols in DEA. *European Journal of Operational Research,* 132(2), pp. 245-259.

Färe, R., Grosskopf, S., Norris, M. & Zhang, Z., 1994. Productivity Growth, Technical Progress, and Efficiency Change in Industrialized Countries. *The American Economic Review,* 84(1), pp. 66-83.

Farrell, M. J., 1957. The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society,* 120(3), pp. 253-290.

Ferreira, C., Marques, R. M. C. & Nicola, P., 2013. On evaluating health centers groups in Lisbon and Tagus Valley: efficiency, equity and quality. *BMC Health Services Research volume,* 13(529).

Golany, B. & Tamir, E., 1995. Evaluating Efficiency-Effectiveness-Equality Trade-Offs: A Data Envelopment Analysis Approach. *Management Science,* 41(7), p. 1172–1184.

Huang, M., Luo, D., Wang, Z., Cao, Y., Wang, H., Bi, F., Huang, Y. & Yi, L., 2020. Equity and efficiency of maternal and child health resources allocation in Hunan Province, China. *BMC Health Services Research,* 20(300).

Jordi, E., Pley, C., Jowett, M., Jaoude G. & Haghparast-Bidgoli, H., 2020. Assessing the efficiency of countries in making progress towards universal health coverage: a data envelopment analysis of 172 countries. *BMJ Global Health,* 5:e002992.

Ministry of Housing, Communities & Local Government, 2019. *The English Indices of Deprivation 2019 - Statistical Release,* London.

Neri, M., Cubi-Molla, P. & Cookson, G., 2022. Approaches to Measure Efficiency in Primary Care: A Systematic Literature Review. *Applied Health Economics and Health Policy,* Volume 20, pp. 19-33.

NHS Digital, 2020. *Appointments in General Practice.*

NHS Digital, 2020. *General Practice Workforce England Direct Patient Care Tables December 2019,* s.l.: Health and Social Care Information Centre.

NHS Digital, 2020. *General Practice Workforce England Admin/Non-Clinical Tables December 2019,* s.l.: Health and Social Care Information Centre.

NHS England, 2014. *Understanding the New NHS - A guide for everyone working and training within the NHS,* London: BMJ.

NHS, 2013b. *The NHS Constitution- the NHS belongs to us all*.

NHS, 2013a. *Guide to the Healthcare System in England*.

Department of Health, 2015. *Health Building Note 00-08 Addendum 1 – A guide to the healthcare system in England for local planning authorities,* Leeds.

NHS, 2019b. *The NHS long term plan*.

NHS, 2019a. *Technical Guide to Allocation Formulae and Pace of Change For 2019/20 to 2023/24 revenue allocations*.

NHS, 2019c. *The NHS Long Term Plan.*

NHS, 2021. *Gov.uk.* [Online]
Available at: https://www.gov.uk/government/publications/the-nhs-constitution-for-england/the-nhs-constitution-for-england?fbclid=IwAR1_ceKBY63_7-IcJLk7829tN3DprfKvm3HLQRX_BXUjPKymZddZ59mkV8I
[Accessed 03 August 2022].

NHS England, 2021. *NHS England.* [Online]
Available at: https://www.england.nhs.uk/get-involved/get-involved/how/primarycare/
[Accessed 4 August 2022].

NHS, 2023. *NHS England.* [Online]
Available at: https://www.england.nhs.uk/commissioning/who-commissions-nhs-services/ccg-ics/
[Accessed 18 June 2023].

OECD, 2019. *Health for Everyone?: Social Inequalities in Health and Health Systems,* Paris: OECD Publishing.

Oikonomou, N. et al., 2016. Measuring the efficiency of the Greek rural primary health care using a restricted DEA model; the case of southern and western Greece. *Health Care Management Science,* Volume 19, pp. 313–325.

Pelone, F., Kringos, D. S., Romaniello, A., Archibugi, M., Salsiri, C. & Ricciardi, W., 2015. Primary care efficiency measurement using data envelopment analysis: a systematic review. *Journal of Medical Systems,* 39(156).

Podinovski, V. V., 2004. Production Trade-Offs and Weight Restrictions in Data Envelopment Analysis. *The Journal of the Operational Research Society,* 55(12), pp. 1311-1322.

Salinas-Jiménez, J. & Smith, P., 1996. Data envelopment analysis applied to quality in primary health care. *Annals of Operations Research,* Volume 67, pp. 141-161.

Scheel, H., 2000. *EMS: Efficiency measurement system user's manual*, version 1.3. University of Dortmund.

Scheel, H., 2001. Undesirable outputs in efficiency valuations. *European Journal of Operational Research,* 132(2), pp. 400-410.

Smith, P., 2009. *Measuring value for money in healthcare: concepts and tools,* York.

Takundwa, R., Jowett, S., McLeod, H. & Peñaloza-Ramos, M. C., 2017. The Effects of Environmental Factors on the Efficiency of Clinical Commissioning Groups in England: A Data Envelopment Analysis. *Journal of Medical Systems,* 41(9).

Thanassoulis, E., 2001. *Introduction to the Theory and Application of Data Envelopment Analysis - A foundation text with integrated software.* Springer Science+Business Media New York.

Teixeira, C., Santos, S. & Amado, C., 2016. *Diferentes Modelos Organizativos de Cuidados de Saúde Primários Apresentam Diferenças no Desempenho?*.

The King's Fund, 2022. *The King's Fund.* [Online] Available at: https://www.kingsfund.org.uk/publications/what-are-health-inequalities [Accessed 03 August 2022].

Ubido, J., Huntington, J. & Warburton, D., 2002. Inequalities in access to healthcare faced by women who are deaf. *Health and Social Care in the Community,* 10(4), pp. 247-53.

WHO, 1978. *Alma-Ata 1978 Primary Health care.* Geneva.

Wilkinson, R. G. & Pickett, K. E., 2006. Income inequality and population health: a review and explanation of the evidence. *Social Science & Medicine,* 62(7), pp. 1768-1784.

Zakowska, I. & Godycki-Cwirko, M., 2020. Data envelopment analysis applications in primary health care: a systematic review. *Family Practice,* 37(2), pp. 147–153.

Zhang, Y., Wang, Q., Jiang, T. & Wang, J., 2018. Equity and efficiency of primary health care resource allocation in mainland China. *International Journal for Equity in Health,* 7(140).

APPENDICES

APPENDIX 1. STUDIES THAT ASSESS HEALTH SYSTEM PERFORMANCE USING DEA

Study	DMUs, Country	DEA Orientation and Returns to scale	Inputs	Outputs	Environmental Variables
Amado and Dyson (2009)	14 GP Practices, UK	Equity of service utilization: Input oriented, CRS;	• WTE GP with diabetics; • WTE Nurse with diabetics; • Number of diabetes-related medication items prescribed to diabetics.	• Number of diabetics who have had a complete diabetes annual review; • Number of diabetics on targeted anti-hypertensive medication (ACEI), if needed; • Number of diabetics on targeted statins, if needed.	Patient-related factors: • The effect of the level of patient compliance, confidence and understanding on the performance levels; the impact of the proportion of elderly patients, newly diagnosed, diabetes type-1 patients and patients with diabetes-related complications on performance.
		Clinical effectiveness: Output oriented, CRS.	• Number of diabetics who have had a complete diabetes annual review; • Number of diabetics on targeted anti-hypertensive medication (ACEI), if needed; • Number of diabetics on targeted statins, if needed.	• Number of diabetics with BP under control; • Number of diabetics with blood glucose under control; • Number of diabetics with blood cholesterol under control.	Structure and processes: • GP-led; • Nurse-led or balanced model; • Training in diabetes; • Training status of the practices; • Type of contract.
		Patient-focused effectiveness: output oriented, CRS.	• Number of patients who answered the diabetes service questionnaire.	• Average number of diabetics who report that they understand enough about diabetes and its management; • Average of diabetics who answered ‘always’ or ‘most times’ regarding the quality of the diabetes services delivered at the practice; • Average number of diabetics who report being ‘confident’ or ‘very confident’ in taking good control over their diabetes.	
Amado and Santos (2009)	18 District Health Authorities, 351 Health Centres	Output oriented, CRS	• Number of doctors; • Number of nurses; • Number of administrative, technical and other support professionals.	• Number of family planning consultations; • Number of maternity health consultations; • Number of junior health consultations; • Number of adult health consultations; • Number of elderly health consultations; • Number of home visits by doctor; • Number of other consultations by the doctor; • Number of other consultations by the nurse; • Number of home visits by the nurse; • Number of curatives and other treatments; • Number of injections; • Number of vaccinations.	

Cinaroglu (2020)	Turkey, 81 provinces, 688 public hospitals.	Input oriented, VRS	• Number of beds; • Number of physicians; • Number of nurses and midwives.	• Number of admissions; • Number of patients; • Number of operations.	
Deidda et al. (2014)	130 Primary Care Centers in 7 primary care districts in the Basque Region in Spain.	Output oriented, VRS	• Number of GPs; • Number of nurses; • Total number of prescriptions.	• Number of consultations by GPs; • Number of consultations by nurses; • Number of notes per patient within the ICT system.	• Mortality index; • Patients over 65y.
Dlouhy (2018)	13 Czech regions, Czech Republic.	Input oriented, CRS	• Health resources (number of doctors and nurses).	• Regional population.	
Ferreira, et al. (2013)	22 ACES (Health center groups) in Lisbon and Tagus Valley region, Portugal	Input oriented, output oriented, non-oriented, VRS and CRS	• Doctors' working hours; • Nurses' working hours; • Administrative staff working hours; • Total costs.	• Number of adult health conditions; • Number of specialty consultations; • Number of urgency consultations; • Number of home visits by doctors; • Total number of consultations; • Number of group education sessions; • Number of consultations by nurses; • Number of injections, vaccinations, curatives and other treatments; • Number of home visits by nurses; • Total number of nursing services; • Total number of public health activities.	• Population; • Population density; • Percentage of patients aged 65yo or older; • Mortality rate; • Percentage of patients without designated doctor; • Distance to nearest hospital; • Purchasing power.
Huang, et al. (2020)	14 regions in Hunan Province, China.	output oriented, VRS	• Number of health personnel; • Number of beds; • Number of medical equipment with a value of ≥CNY 10 000; • Operating areas.	• Number of live births; • Number of outpatients and emergency visits; • System management rate for children under 3 years old.	
Jordi, et al. (2020)	International - 172 countries	Output oriented, VRS	• UHC service coverage index (comprises 16 tracer indicators); • Proxy for financial risk protection (proportion of current health expenditure).	• Current health expenditure (CHE).	Second-stage DEA: factors outside the health system: income, governance, education and health system capital – social, political, economic and environmental explanatory factors.
Neri, et al. (2022)	Systematic review- 38 International studies, in high-income settings.		Labour, capital and intermediate inputs. Measured in physical units (volume) or via prices (total cost).	• Pure measures of primary care utilisation (did not include information about quality standards of care); • Measures of primary care utilisation adjusted by quality (controlled for the quality if care delivered, or the performance targets achieved); • Health outcomes (final or intermediate);	<u>Characteristics of population:</u> demographics, economic or employment status, case-mix, mortality rates, geographic area of the patient population. <u>Organisational features:</u> configuration and organisation of primary care delivery, the financing level and system in place,

				• Impact on health-related quality of life (HRQoL): expressed in quality-adjusted life years (QALYs) or disability-adjusted life years (DALYs).	external market influences (eg. Number of health facilities in the region), or level of clinical performance.
Oikonomou, et al. (2016)	42 Health centers and their regional surgeries of the 6th Health Prefecture (HP), southern and western Greece.	Output oriented, 1st VRS then CRS	• Medical personnel; • Nursing personnel; • Technological equipment.	• Acute, chronic and preventive consultations.	
Pelone, et al. (2015)	Systematic review - 39 studies	Majority of the studies used an input orientation (27), 3 used input and output orientation and 9 used output orientation.	3 main input categories: • Labour; • Capital; • Consumable resources.	Only 2 studies use health outcomes; 20 used activity indicators; 7 used quality indicators; 5 studies use activity and quality indicators.	
Salinas-Jiménez and Smith (1996)	85 Family Health Service Authorities (FHSA), UK	1st input oriented 2nd output oriented, CRS	• Gross expenditure on general medical services (in pounds) per head of resident population.	Quality indicators: • General medical practitioners per 10 000 patients on list; • The percentage of practices employing a practice nurse; • The percentage of general medical practitioners who had a patient list of less than 2 500 patients; • The percentage of general medical practitioners not practicing single-handed; • The percentage of general medical practitioners who had achieved higher rate of payments for childhood immunization; • The percentage of females aged 35 to 64, registered with the FHSA and who had an adequate cervical smear in the previous five and half years; • The percentage of practice premises which satisfied the minimum standards set out in para 51.10 of the Statement of Fees and Allowances, excluding practices exempt under para 51.11.	• Standardised illness ratio; • Unemployment.
Takundwa, et al. (2017)	208 Clinical Commissioning Groups, England		• Per capita CCG funding allocation; • Number of GPs per 100 000 population.	• Directly standardized average health status score for individuals aged 18 and over; • Directly age and sex standardized respiratory disease survival rate per 100 000 population; • Directly age and sex standardized cancer survival rate per 100 000 population; • QOF cardiovascular disease score; • QOF cancer score; • QOF COPD score;	• Indices of multiple deprivation ranking; • GP registered population; • GP registered population <18y; • GP registered population aged 65y and over; • GP registered population aged 18y and over with a long standing health condition;

				• Percentage of patients who would recommend the GP practice to others.	• GP registered population aged 18y and over who are unemployed; • Estimated smoking prevalence; • Prevalence of obesity; • COPD prevalence; • Cancer prevalence; • CHD prevalence.
Zakowska and Godycki-Cwirko (2020)	Systematic review - 54 studies.	VRS is most common; Some studies apply both.	Most common input categories: • Personnel; • PHC centers; • Consultations or visits; referrals or hospitalization days; • Pharmaceuticals and prescriptions; • Procedures, treatments and services; • Patients. Min. number of inputs: 1 Max. number of inputs: 24 Mean and Mode: 3	Most common output categories: • Consultations or visits; • Patients; • Procedures, treatments, and services; • quality; • Personnel; • Preventive interventions (including vaccinations); • PHC center; • Referrals and hospitalizations. Min. number of outputs: 1 Max. number of outputs: 21 Mean: 4 Mode: 3	
Zhang, et al. (2018)	31 provinces that were divided into 3 groups – eastern, central and western regions, China	Input oriented, VRS	• Number of institutions and beds (capital); • Health workers – physicians, nurses, other clinical staff, administrative and non-clinical staff (labour).	• Average number of visits; • Annual hospitalization rate.	

APPENDIX 2. WEIGHT RESTRICTIONS APPLIED TO THE THREE DEA MODELS

DEA Model 1- weight restrictions	DEA Model 2 – weight restrictions
WR1: $\boldsymbol{v_8 - v_7 \geq 0}$	WR1: $v_1 - v_2 \geq 0$
WR2: $\boldsymbol{v_8 - v_6 \geq 0}$	WR2: $v_1 - v_3 \geq 0$
WR3: $\boldsymbol{v_8 - v_5 \geq 0}$	WR3: $v_1 - v_4 \geq 0$
WR4: $\boldsymbol{v_8 - v_4 \geq 0}$	WR4: $v_2 - v_3 \geq 0$
WR5: $\boldsymbol{v_8 - v_3 \geq 0}$	WR5: $v_2 - v_4 \geq 0$
WR6: $\boldsymbol{v_8 - v_2 \geq 0}$	WR6: $v_3 - v_4 \geq 0$
WR7: $\boldsymbol{v_8 - v_1 \geq 0}$	WR7: $u_2 - u_1 \geq 0$
WR8: $\boldsymbol{v_7 - v_6 \geq 0}$	WR8: $u_2 - u_3 \geq 0$
WR9: $\boldsymbol{v_7 - v_5 \geq 0}$	WR9: $u_1 - u_3 \geq 0$
WR10: $\boldsymbol{v_7 - v_4 \geq 0}$	
WR11: $\boldsymbol{v_7 - v_3 \geq 0}$	
WR12: $\boldsymbol{v_7 - v_2 \geq 0}$	**DEA Model 3 – weight restrictions**
WR13: $\boldsymbol{v_7 - v_1 \geq 0}$	WR1: $u_2 - u_1 \geq 0$
WR14: $\boldsymbol{v_1 - v_2 \geq 0}$	WR2: $u_2 - u_3 \geq 0$
WR15: $\boldsymbol{v_1 - v_3 \geq 0}$	WR3: $u_1 - u_3 \geq 0$
WR16: $\boldsymbol{v_1 - v_4 \geq 0}$	
WR17: $\boldsymbol{v_1 - v_5 \geq 0}$	
WR18: $\boldsymbol{v_1 - v_6 \geq 0}$	
WR19:$\boldsymbol{u_1 - u_2 \geq 0}$	
WR20: $\boldsymbol{u_1 - u_3 \geq 0}$	
WR21: $\boldsymbol{u_1 - u_4 \geq 0}$	
WR22: $\boldsymbol{u_2 - u_3 \geq 0}$	
WR23: $\boldsymbol{u_2 - u_4 \geq 0}$	
WR24: $\boldsymbol{u_3 - u_4 \geq 0}$	