The London School of Economics and Political Science

The impact of variation in critical care organisation on patient mortality: evidence form the United Kingdom

Ritesh Maharaj

A thesis submitted to the Department of Health Policy of the London School of Economics and Political Science for the degree of Doctor of Philosophy

London, September 2022

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I declare that my thesis consists of 51123 words (including appendices but excluding references).

Statement of co-authored work

I confirm that Chapter 4 was jointly co-authored with Professor McGuire and Professor Street, and I contributed 85% of the work. A version of this chapter has been published in a peer reviewed journal. The citation for this publication is: Maharaj R, McGuire A, Street A. Association of Annual Intensive Care Unit Sepsis Caseload with Hospital Mortality From Sepsis in the United Kingdom, 2010-2016. JAMA Netw Open. 2021 Jun 1;4(6):e2115305. doi: 10.1001/jamanetworkopen.2021.15305. PMID: 34185067; PMCID: PMC8243236.

Abstract

The changing landscape of aging population, increasing incidence of critical illness and more constrained national budgets mean physicians, policy makes, and hospital administrators must consider more efficient ways to organise critical care services. In general, policymakers have embraced the idea of centralising services and increased specialisation to improve efficiency in health care. This thesis explores these policies in the context of critical care services in the UK. Evidence of the productivity of critical care services and in particular volume-outcome relationship in critical care and the underlying mechanism by which this relationship operates is scarce.

I consider several aspects of these issues. In the first study I investigate the volume-outcome relationship for sepsis using data from the Intensive Care National Audit and Research Centre which covers all ICUs in the England, Wales, and Northern Ireland. In this cohort study, sepsis case volume in an ICU was significantly associated with hospital mortality from sepsis, and a volume lower threshold of 215 patients per year was associated with an improvement in mortality. The second study explores the underlying mechanism of the volume-outcome relationship. Two possible mechanisms proposed are dynamic learning-by-doing and static scale economies. If the volume-outcome relationship operates through the learning-by-doing mechanism, then patient outcomes would improve by the volume of patients treated over time, making system-wide centralisation unnecessary. This study supports the idea that the underlying mechanism by which volume leads to improved outcomes is through learning-by-doing. ICUs tend to improve by caring for a large volume patients distributed over time. Patients may, therefore, be better served by ICUs organised to achieve minimum volume

standards without centralisation. The third study examines the related role of ICU specialisation in improving mortality. This study found that ICU specialisation do not have significantly lower hospital mortality for critically ill patients in the UK after adjusting for patient characteristics and caseload volume.

Across the three studies I argue that a minimum volume threshold may be effective in improving patient outcomes. Centralisation may not fully leverage the benefits of the learning-by-doing mechanism. Lastly, accounting for volume, there is no compelling evidence of any added value from ICUs specialisation.

Acknowledgements

I am immensely grateful to my supervisors Professor Andrew Street and Professor Alistair McGuire for their invaluable advice and support throughout the PhD program. Over the past 4 years I have enjoyed the benefit of their generous, and constructive feedback that has helped me navigate the many challenges that come with pursing a PhD.

I am also grateful for the faculty in the department of Health Policy at the LSE for their contributions to the PhD seminar sessions during the HP500 seminar series. I would like to thank Andrew Street, Mylene Lagarde, Huseyin Naci, Justin Parkhurst, and Joan Costa-Font for the helpful feedback that improved my PhD.

I would like to thank Prof Irene Papanicolas and Prof Steve Morris for their helpful comments during my PhD viva. Addressing their comments greatly strengthened the work presented in this thesis.

I would like to thank my fellow PhD students in the Department of Health Policy for the lively discussions and community that we created during the challenging times we faced. It has truly been a privilege to share an office with you. I would like to thank Camille, Iva, Nilesh, Arianna, and Nick for the countless moments of humour, warmth, kindness, and friendship.

Thank you to Andrew and Xiaozhou for an enjoyable conference at the Health Economics Study Group meeting. I would like to thank my family for their patience and constant support during my PhD studies. I am grateful to my friends who have helped keep me focused. I would also like to thank my colleagues at Department of Critical Care Medicine, King's College Hospital, for being so accommodating of my various requests and for being so supportive.

Table of Contents

List of tables

List of figures

List of abbreviations

GDP- Gross domestic product

ICU- intensive care unit

PIR- patient-to-intensivist ratio

UK- United Kingdom

US -United states of America

ICS- Intensive Care Society

SCCM- Society of Critical Care Medicine

OECD- Organisation for Economic Co-operation and Development

CMPD- Case-Mix Program Database

APACHE II- Acute Physiology and Chronic Health Evaluation

ICNARC- Intensive Care National Audit and Research Centre

LTACHS- Long Term Acute Care Hospitals

SNF- Skilled Nursing Facilities

RSMR- Risk standardised mortality ratio

ICM- ICNARC Coding Method

RoBINS- Risk of Bias in Non-Randomised Studies

SOFA- Sequential Organ Failure

CABG- Coronary artery bypass graft

PCI-Percutaneous coronary intervention

Chapter One: Introduction

Abstract

This chapter reviews the conceptual and empirical literature on the volume-outcome relationship exploring the economies of scale and learning-by-doing mechanism to give firm foundation to my application in critical care delivery within the UK. Additionally, I describe the related literature on specialisation and quality of care.

Luft first described the inverse relationship between volume and outcome for a range of surgical procedures and mortality in 1979, where mortality was reduced as annual caseload increased. Since then, there has been a large body of literature, focused mainly on surgical procedures, that have reported a similar finding. Nevertheless, several uncertainties remain. The volume-outcome relationship in non-surgical cohorts is less well studied. The role of severity of illness, the complexity of care required, and a defined threshold that can be used to set policies such as minimum volume standards have also been inadequately described.

There is also a paucity of studies examining the underly mechanism of this relationship, disentangling the scale effects and the learning-by-doing effects. The term learning-by-doing was popularised in economics by Arrow in 1962[1]. In the 1970s and 1980s the strategic implications of learning-by-doing were applied to industrial trade policy [2]. In healthcare, studies have been focused on cardiovascular procedures, namely coronary artery bypass graft (CABG) and percutaneous coronary interventions (PCI). Most have not been able to identify a learning-by-doing effect. The learning-by-doing mechanism is less well explored outside of cardiovascular procedures.

More than 40 years ago, Skinner proposed the concept of the focused factory that argued firms would develop competitive advantage by limiting the scope of their work[3]. In industry the focused factory idea has been applied to specific product (or service) lines, processes and competitive priorities[4]. Specialisation on specific groups of patients has been proposed to improve quality and efficiency. Those in favour argue that specialisation brings less uncertainty, and results in the development of higher levels of expertise. Arguments against specialisation are that it increases fragmentation of care, and that greater breadth gives benefits through economies of scope by sharing common resources across groups of patients. Whilst specialisation might appear attractive from the physicians' perspective to focus on one area of clinical practice, it remains unclear if specialised service offers any patient benefits.

1.1 Introduction

Across high income countries, health care expenditure accounts for a significant and ever-increasing proportion of national gross domestic product (GDP). Many countries have enacted reforms to improve health system efficiency and performance, among which are policies regarding the optimal size and configuration of healthcare providers. This thesis will first consider the optimal size by examining the volume-outcome relationship, including exploring the underlying mechanism namely learning by doing compared with static scale economies. The thesis then considers optimal configuration by exploring the effects of specialisation of the quality of care provided. The relationship between the institutional volume caseload and performance has been well described but the source of this relationship remains uncertain. Economic literature has considered economies of scale, learning by-doing, level of specialisation as factors explaining this relationship. Data collinearities make inference difficult. This thesis will explore multiple factors to inform where improvements in quality can be made.

1.2 The volume-outcome relationship

The volume-outcome relationship in healthcare has been widely studied since the 1970's, mostly in the context of complex surgical procedures[5]. Although the volume-outcome relationship has been long recognised, it's role in policies such as minimum volume standards and regionalisation of care has only recently been advocated. Early volume-based policies have been consumer focused and involved making patients aware of hospital procedure volumes when they seek care. However, these had little impact on patient behaviour[6].

Recent polices have been more regulatory in nature, such as incorporating certificates of need for new surgical centres and minimum volume standards for a small range of surgical procedures [7, 8]. Minimum volume standards have emerged as a prominent policy to leverage the volume-outcome relationship into improvements in quality of care. One of the major barriers to the successful implementation of this policy has been that minimum volumes have been set normatively instead of empirically estimated.

The literature evaluating the volume outcome relationship is largely focused on surgical procedures using either departmental and individual surgeon volumes[9]. Institutional volumes reflect infrastructure, staffing, technology, and other institutional characteristics. Individual surgeon volumes capture individual surgeon traits like technical skills and decision making[10]. Whilst both levels may influence the outcome, we would presume the strength of the relationship depends on the contribution the surgical difficulty makes to the outcome or the availability of hospital level resources or non-surgical expertise[11]. Empirical evaluations of the interplay between surgeon volume and hospital volume have largely concluded that even for high-risk surgeries, hospital volume is a more significant contributor to mortality than surgeon volume, although there is no consensus[12]. Low-volume surgeons operating in a high-volume centre appear to have lower mortality and length of stay compared with operating in low volume centres[13].

Much of the early economic literature has focused on the econometric challenges of identifying the causal effects of endogenous selection through selective referral where volume better hospitals attract more patients, making volume endogenous. Selective referral suggests that the observed volume-outcome relationship reflects the referral system that

directs more patients to hospitals that are known to deliver higher quality[5]. Selective referral could also occur without explicit knowledge about quality. Patients or their referring physicians may choose hospitals because of having a good reputation or avoid others because of bad reputations. Both scenarios result in higher quality generating higher volumes.

In general, two approaches have been used to account for selective referral, namely instrumental variables, or longitudinal models with hospital fixed effects. Instrumental variables try to break the endogenous link and results from studies using instrumental variables depend on the validity and strength of the instrument. A major challenge has been identifying a valid instrument. Commonly used instruments are the number of beds or a geographical factor[6, 15]. Hospital beds are related to hospital size and have been shown to directly influence quality making it an invalid instrument. Longitudinal studies with hospital fixed-effects require sufficient within-hospital variation in volumes to identify an effect. These studies have typically used a small sample of hospitals with little variation in within-hospital volumes and are likely underpowered[14]. Fixed effects try to pick up the effects of unobservable variables but may lead to measurement error through reliance on within-hospital variation if this is small. Also, any omitted variable bias due to selection is unlikely to be fixed over time.

Overall, the precise magnitude of selective referral remains unclear and the role it may have on the explaining the volume outcome relationship is likely to be negligible[15]. Flood et al reject the selective referral hypothesis arguing that the effects of differences in mortality are usually small and unlikely to drive referral patterns or patient choice[16]. Many researchers have subsequently assumed volume to be exogenous[15].

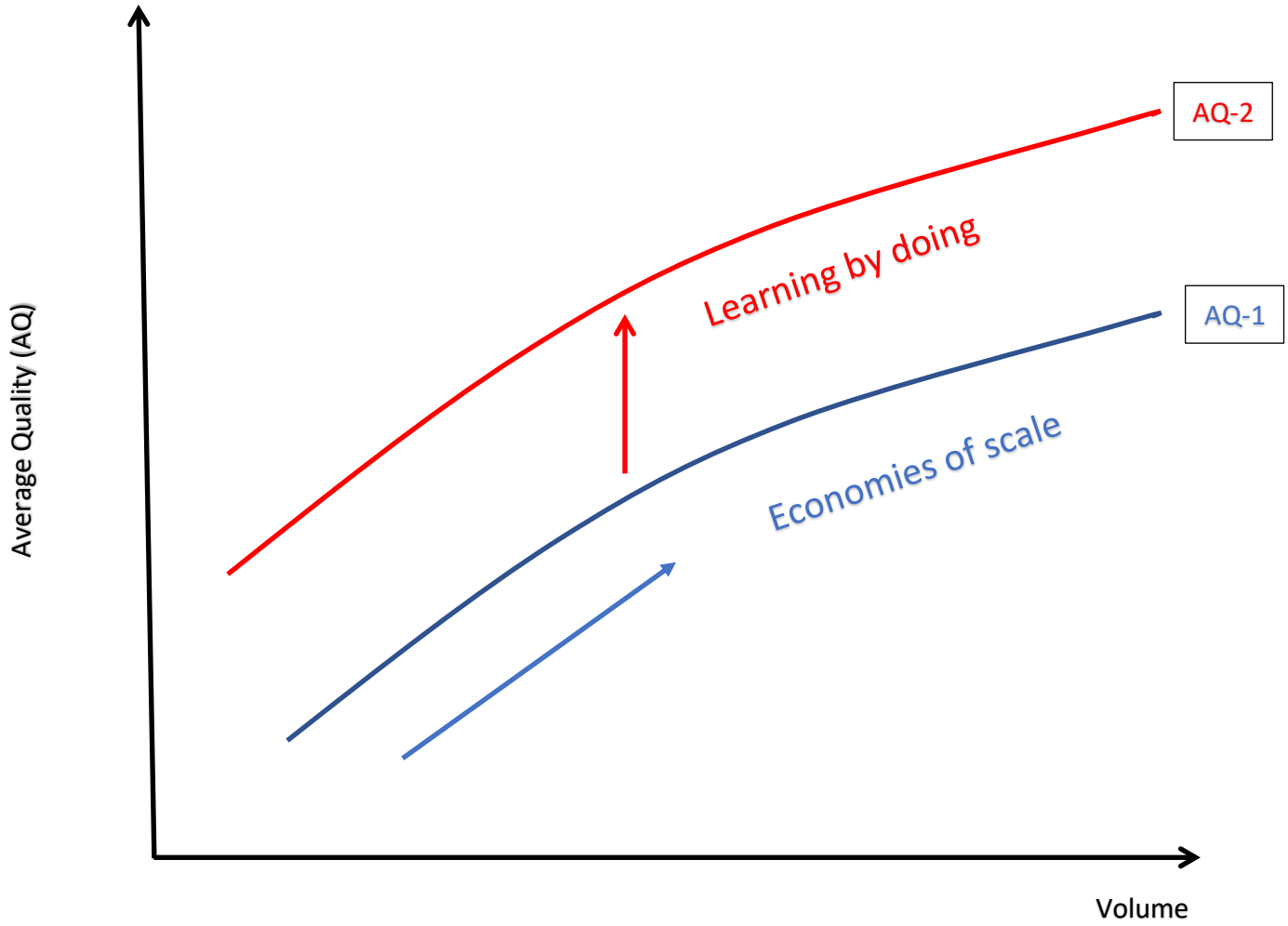
Economies of scale refers to the effects of volume captured by annual patient caseloads. The rationale is that higher volumes institutions are more likely to be better equipped and have

more standardized treatment pathways and processes of care which could potentially lead to better patient outcomes.

1.3 Learning-by-doing

Over the years, the volume-outcome relationship has been used to support quality improvement policies such as the minimum volume standards for surgical procedures and centralisation of high-cost services[17]. Whilst the general approach of "more is better" is superficially attractive, the success of volume-based policies requires an understanding of the underlying mechanism. Policies rely on the premise that volume improves outcome either through economies of scale or learning by-doing. There is an important distinction between these mechanisms (Figure 1). Economies of scale refers to the reduction in average costs or an improvement in quality as related to the quantity (volume) of goods provided and refers to movements along the average quality (cost) curve. In contrast, learning-by-doing refers to the improvements in productivity or quality that occur because of experience and result in higher quality or lower costs at any level of volume. Learning-by-doing results in a shift of the volume-quality curve. Learning-by-doing refers to improvements in quality or reductions in cost that are related to the experience of the firm and expertise in technological improvements and not simply attributable to economies of scale.

Figure 1. Economies of scale versus learning-by-doing.



The blue arrow refers to movement along the Average Quality (AQ 1) curve reflecting effect of economies of scale. The red arrow refers to changes in productivity by experience or learning-by-doing, resulting in a new Average Quality curve AQ-2.

The distinction between static economies of scale and dynamic learning-by-doing is important because if the volume-outcome operated entirely through movement along the economies of scale curve (AQ-1), such as by investments in infrastructure and research and development, then equating static marginal quality to marginal volume would be socially optimal. Consider the example of a transitory shock that raises short-term demand such as a pandemic, assuming demand does not exceed a supply threshold. In such a scenario the economies of scale mechanism predicts no long-run gains to quality when volume returns to baseline. In contrast, the learning-by doing mechanism would predict a permanent improvement quality (AQ-2) from that point onwards.

A simple model of learning-by doing can be described as:

$$mortality_{ij} = E_j + V_j + X'_{ij} + \Phi'_{ij} + e_{ij}$$

Mortality for patient i in ICU j depends on the E experience of the ICU, the V volume of the ICU, X'patient characteristics and Φ' time invariant ICU characteristics and e random error. While this specification appears simple, several issues arise. First, outcomes may be heterogenous across different types of patients. Second, current volume will be determined through the production function where past decisions dictate the capacity to treat in any hospital so that current volume is a function of past volume, making lag effects important. Third, experience is itself a dynamic issue proxied by lagged volumes.

Studies account for heterogeneity of treatments by focusing on specific diseases. Most studies focused on coronary artery disease requiring either coronary artery bypass graft (CABG) or percutaneous coronary interventions (PCIs)[15, 18]. Studies have used lagged

volumes or cumulative volumes to identify the learning-by-doing effect in these procedures and using hospital fixed effects to control for time invariant unexplained heterogeneity[15, 19]. Hospital fixed-effects would mean that the regressions estimate the effect of increases or decreases of lagged or cumulative volumes rather than the differences in lagged or cumulative volume across hospitals on mortality. This approach requires a sufficient variation in volumes to observe a learning-by-doing effect. The absence of an effect in prior studies may be due to small sample size and small changes in volume over time.

Health economists are particularly interested in learning-by-doing because most emerging medical technologies are complex, and the team-based skills required to deliver them would benefit from experience. Early studies found a volume-outcome relationship and attributed this to the learning-by-doing effect. More recent studies have failed to identify a learning-by-doing effect and have cast doubt on this mechanism. One possible explanation for the lack of demonstrable learning-by-doing effect in these studies could be the differences in the approach taken in industry compared to healthcare.

Arrow first applied the term learning-by-doing to firm learning and much of its early use of was in the industrial sector, referring to the growth in productivity by experience[1]. Even when volumes remain the same, improvements in quality can therefore still occur. These improvements are not attributable to economies of scale but rather to firm experience.

There are several reasons why learning-by-doing may be different in healthcare compared with industry. First, in the general industrial organisation literature, the learning-by-doing mechanism was focused on costs or quantity of goods produced and not on product quality.. There is no substantial work from industry evaluating the effect learning-by-doing has on the quality of goods. In contrast the health economic literature has focused on the quality of

health care, particularly mortality. We have assumed that the mechanisms that lead to lower costs in industry are the same as those that lead to lower mortality in healthcare, but this may not be the case. Second, identifying the origin of a product or product line is easier in manufacturing, experience is less easy to identify in healthcare. In more general industrial economics studies, learning-by-doing is identified by cumulative volume. Cumulative experience in industry is straightforward to measure since the cumulative volume in industry is often readily captured. A challenge in health care is being able to identify novel procedure that can be used to track cumulative volume from the beginning of its use. When cumulative volumes since are not available, a common alternate approach in identifying a learning-by-doing effect in established diseases is to use lagged volumes[18].

Experience is increasingly important because medical care is increasingly complex and more frequently delivered by teams. Determining the magnitude of the learning-by-doing effect compared with the economies of scale effect is important for assessing the likely success of small versus large services. If services improve by learning, then maintaining a large number of providers may be beneficial through competition and would argue against the benefits of centralisation.

The health economics literature on learning-by-doing has focused on cardiac surgery, limiting the conclusions that could be drawn from such a limited body of evidence. There may be differing importance of individual doctors in surgical procedures and non-surgical procedures. In non-surgical patient cohorts, outcomes may be more related to the clinical teams rather than the technical skills of individual doctors. The skills of a clinical team refer to the leadership, decision making, communication and co-ordination behaviours used by the multidisciplinary team members[20]. Patient safety research has demonstrated the importance of teamwork in patient outcomes[20, 21]. Poor communication during ward

rounds or handoffs is frequently cited as a major contributor to medical error and teams that report high levels of collaboration between team members also demonstrate lower patient mortality[22]. Communication failures commonly occur at transitions of care, at shift changes or changes of care areas in acute settings[23]. These high-risk interactions in which critical information about patient's plan of care can be miscommunicated, leading to harm, delays or inappropriate therapies are frequent points of failure[24]. The interactions between members of the team can contribute to specific errors. For example, poor communication between physicians, nurses and pharmacists can lead to errors in drug, dose, and route of administration[25].

Complex tasks usually involve a division into smaller components of treatment that requires teamwork and communication. Failure at an organisational level is reflected in co-ordination neglect[23]. The strong tendency to focus on division of labour and less emphasis on the co-ordination and integration of care, has been associated with patient harm explored by seminal work differentiating taskwork from teamwork [26, 27]. Even if one individual provides the highest standard of care, this alone will not protect the patient from harms across the treatment episode because of this interdependence of team performance.

Critical care services are delivered through teams consist of doctors, nurses, physiotherapists, pharmacists, dieticians, and a range of rehabilitation therapists. This team construct has often been compared to the aviation industry[28]. Team performance in the ICU, like the aviation industry, requires communicating priorities and appropriate task delegation.

ICU teams are unusual in that they demonstrate low temporal stability, which can impede team dynamics. Shift patterns and within and between hospital rotations for training schemes

means that team members may not have worked together before and need to reference a common departmental framework in an acute setting. Successful ICU teams can collaborate, share information and co-ordinate towards a common goal or task. This requires striking a balance between inclusiveness and authority in creating shared goals and fostering a sense of collective responsibility. ICU team members are expected to bring shared knowledge to formal and informal handoffs, operating under the emotional and physical constraints of high acuity service. Death and dying occur regularly in the ICU and the environment is characterized by constant moral distress and grief[29]. In this environment teams change often, and it is likely that institutional knowledge depreciates as teams change.

A challenge to identifying a learning-by-doing effect is that cumulative or lagged volumes are often highly correlated with the contemporaneous volume. High volume centres in one year tend to be high volume in previous years. One way to resolve this would be to use a large number of centres over a long period of time to identify sufficient temporal variation in volumes to weaken the collinearity between lagged and contemporaneous volume[18]. Data restrictions normally prohibit such an approach. The current literature largely studied learning-by-doing over years and failed to demonstrate an effect[15]. It is possible that learning occurs over a shorter time epoch than annually, given the frequency with which medical teams change. Short-term learning effect may occur over quarters or months.

In summary, learning-by-doing has been frequently demonstrated in manufacturing but it has been more challenging to identify it in healthcare[30]. Distinguishing scale effects from learning-by-doing has been difficult. Most of the empiric literature has used relatively small datasets and might be therefore underpowered to detect this important distinction. Medical

teams undergo frequent turnover, and it might be that learning occurs over shorter periods than has been previously studied.

1.4 Specialisation

Increasing specialisation is proposed to improve quality and efficiency in the hospital sector. Specialisation outside of healthcare has a long history dating back to Smith (1776) and Taylor (1911)[31, 32]. Specialisation can be conceived as manufacturing system than is limited to a small number of products or technologies and is analogous to the focused factory in the manufacturing sector[4]. The focused factory idea applied to healthcare has suggested that hospitals could improve the quality of care and efficiency by focusing on a specific disease or medical procedure. The main idea of focus is that organisations are required to identify segments in their markets and develops specialist systems for each segment. Segmentation aims to identify a homogenous group of patients with similar and predictable demands[4]

In the hospital sector, a focus on a specific service line usually requires the aggregation of patients with similar diagnosis and needs, as well as to physically co-locate the resources needed[33].

The simplest formulation of specialisation to a segment of the market in critical care would be:

$$mortality_{ic'j} = S_j^c + V_j + X'_{ic'j} + \Phi'_{ic'j} + e_{ic'j}$$

Mortality for patient i from diagnostic group c' in ICU j is related to the magnitude of specialisation S in diagnostic category c as well as X'patient characteristics and Φ' time invariant ICU characteristics and e random error. Patients in the ICU could be of the same diagnostic group c as the specialisation of the ICU ($c' = c$) (e.g., a cardiac patient in cardiac

specialised ICU) or patients could belong to a diagnostic group other than the speciality of the ICU($c' \neq c$) e.g., a cardiac patient in a neurosurgery specialised ICU).

This simplified description of specialisation belies the methodological challenges posed by trying to identify specialisation. There is no agreement about how best to measure specialisation. Deriving information about the extent to which a service is specialised is constructed from the data [34]. The most common approach is to group patients according to their diagnostic group and then constructing some measure based on these groups[34]. Previous measures have been based on proportions. The Information Theory Index (ITI) compares the average proportion of patients across diagnostic categories within a hospital to the average across all hospitals in a region or country[35]. The Herfindhal-Hirschman Index(HHI) and a Gini-derived index developed by Diadone and D'Amico provide similar measures of specialisation [36]. All these measures identify services that are most distant from the average hospital and do not necessarily describe medical specialisation. Hospitals that have low proportions of one diagnostic group will score similarly to hospitals with high proportions because they are equidistant from the average hospital. That is to say that these measures characterise the degree of specialisation as the deviation of the case-mix from the average case-mix across a predefined sample population. Identifying a medical specialisation measure such as a cardiac, trauma or sepsis specialised service may be more informative than an average case-mix measure of specialisation.

The benefits of specialisation could be observed in patients that are within the specialty of the ICU ($c' = c$) as well as on cases that are outside the speciality of the ICU($c' \neq c$) . First, a specialised service model characterised by a narrow range of high-quality products is

assumed to benefit patients within focus i.e., when $(c' = c)$ then patient i^c treated in ICU of S^c specialisation is assumed to benefit. Patients benefit from expertise and resources to treat the disease in focus, reduced variability, and fewer competing priorities. Second, the benefits to specialisation may also have favourable spill over effects on patients outside the focus of specialisation [37]. Specialisation may give rise to knowledge spill overs so that treatment benefits are seen across different types of cases. This occurs when expertise gained by caring for one group of patients may provide transferable skills to caring for other groups of patients. Chandra et al have described a productivity spill over when physicians learn new techniques when exposed to other physicians who are expert in those techniques, which leads to improvements in productivity through knowledge[38].

Despite the persistence of the idea of specialisation, the empirical evidence has been mixed. The reorganisation of services focused on specific service lines has found no improvement in costs, length of stay, increased disruption with lower job satisfaction and professional development [39-41]. Specialist hospitals focused on cardiac and ambulatory surgery have shown divergent outcomes. Studies report higher patient satisfaction and improved patient outcomes as well as higher rates of readmission and higher cost inefficiencies[42-45]. A study of cost efficiencies across 3 states in the US has shown speciality orthopaedic, and general surgical hospitals to have higher levels of cost inefficiency compared to the general hospitals they compete with[45]. Cardiac hospitals by comparison were found to be as efficient as their general hospital competitors[45]. One possible explanation for this might be the scale of operations in that specialist cardiac hospitals operate at a similar scale to general hospitals, but orthopaedic and general surgical hospitals tend to be much smaller than general

hospitals. The benefit of specialisation is therefore likely related to the case-mix, the types of services needed and the degree of customisation.

Part of the challenge of measuring specialisation is the lack of agreement about how to best characterise specialisation. The literature characterises specialisation along lines of types of treatments, patient characteristic's, medical speciality and organisational characteristics[4]. To address this diversity, Herzlinger suggested that specialisation into focused factories should be based on common objectives such as the treatment of specific groups of patients[46]. This contrasts with the traditional approach of organisational units based on specialties. A practical example of this would be stroke centres and heart attack centres instead of neurology and cardiac centres respectively.

Second, evaluating the efficiency of such a diverse definition has been problematic. Some studies have used patient satisfaction and readmission rates and others have used mortality to draw conclusions about operational performance[42]. These difference in outcomes could lead to differences in conclusions as to the value of specialisation

Critical care services are an integral but high-cost component of a functioning health system. Whilst much of the focus of individual critical services has been on the evolution of technology for organ support, the broader question of how best to organise these services to improve patient outcomes and reduce costs has not been addressed. These questions have become increasingly important in the face of increasing demands and tighter budgets. There is substantial variation in the way ICUs are organised within hospitals. Some ICUs admit patients with a wide range of diagnoses. These ICUs are termed general or non-specialist ICUs. Proponents of general ICUs argue that critically ill patients with different diagnoses usually

require similar treatments, obviating the need for more narrowly focused care. For example, lung protective mechanical ventilation for acute respiratory failure in community acquired pneumonia would be similar for respiratory failure in blunt chest trauma. Diversification leads to economies of scope because there is some clinical convergence of the treatment pathways. As such there are knowledge spill overs from one diagnostic group to another. Alternatively, critical care services may be segmented into narrowly focused diagnoses such as cardiac, neurosurgical, medical or elective surgical ICUs[47]. The expected benefits of pooling patients with similar diagnoses are greater predictability, reduced diagnostic variability, the development of specialised expertise, and the development of more standardised treatment pathways. The concern with specialised ICUs is that they may have limited ability to treat patients outside the speciality diagnoses and create shortages elsewhere in the system.

Understanding the benefits of specialisation is therefore important to more efficiently utilising the limited critical care capacity. The relationship between ICU specialisation and mortality has been poorly described[47]. Thus far there has been one large study of ICU specialisation across range of specialities that failed to identify any benefit[47]. Smaller studies specifically of neurosurgical speciality ICUs and cardiac ICUs have shown inconsistent results[48].

The general arguments in favour of increased specialisation suggests that there may be benefits to a narrow spectrum of focus because of more predictability and less conflicts in operational goals. The critically ill patient is increasingly multi-comorbid and prone to multisystem complications requiring a common basket of organ supports. These include circulatory support, renal replacement therapy and mechanical ventilation. A model of care

delivery that is adaptable to the wide range of expertise required to manage the critically ill may be required.

The current evidence for critical care outcomes in speciality versus general ICUs is poor. First, consistent with general literature on specialisation, there is no gold standard for the definition of specialist ICU. Consequently, the identification of specialist ICUs is inconsistent across studies. Most specialist ICUs are self-designated, often by administrators and there are no regulatory requirements to have the title of specialist ICU. The designation may not reflect the case-mix within the ICU. For example, some ICUs considered themselves specialist by implementing a standardised protocol, whilst others considered themselves as general ICUs with a standardised treatment protocol for specific diseases[49]. Second, most studies focused on surgical specialities such as neurosurgery or cardiac surgery, where there may be risk selection, overestimating the benefits of specialisation[50]. Third, most studies did not control for caseload volume, academic affiliation, or patient characteristics, as well as the clustering of patients within ICU. These organisational and patient characteristics may be more relevant in determining mortality than specialisation per se.

In summary, our understanding of the relationship between specialisation and mortality in hospital services is limited by a lack of gold standard in measuring specialisation. Specialist services may reduce mortality because of higher levels of expertise and better organised care. Specialisation may also have spill over effects to patients outside the speciality because of the relatedness of patients within a service. Much of the existing literature on specialisation has failed to account for volume and patient characteristics.

These concepts of economies of scale, learning-by-doing and specialisation will be applied more specifically to critical care. The thesis begins by first by describing the interaction between critical care services and the wider hospital and then develops by describing the data

used in the empirical work. The thesis proceeds with a systematic review of the volume-outcome relationship in the ICU and then evaluates this relationship for sepsis in the UK. The underlying mechanism of the volume outcome relationship is explored in the subsequent chapter followed by a chapter on the benefits of ICU specialisation. The thesis concludes with a discussion of the policy and research implications of this work.

Chapter Two: The organisation of the ICU within the hospital

Abstract

The ICU can be viewed as a "hospital within a hospital" and in this sense models of hospital behaviour can be informative. The history of critical care services can be traced to the Polio epidemic in Denmark in the 1950. The major innovation was the organisation of multidisciplinary teams and developing expertise of organ support that was applicable across many diseases. The flow of patients into and out of the ICU offer some insight into how central the ICU is to deliver both elective and emergency care. There is huge variation between health systems as to the definition of an ICU bed. Some countries define ICU by the intensity of nursing care or monitoring required, others by the number of organs supported and by the severity of illness of the patient.

Despite these differences, and broader variations within the health systems, international comparisons can provide useful insights about the delivery of care and inform the generalisability of interventions or policies. Even amongst countries with similar spending on health, there are huge difference in the quantity of critical care services provided. This provides a useful context to explore the wider effects of over and undersupply of critical care. The COVID-19 pandemic highlighted how the national picture of ICU beds per capita was associated with differences in mortality[51]. The difference in ICU beds also result in different practice patterns. For example, end-of-life care is more frequently delivered in the ICU compared to non-ICU settings in systems where more ICU beds are available[52]. Low risk patients with a predicted mortality of <2% are unlikely, on average, to benefit from critical care. These patients are more likely to be treated in the ICU when beds are plentiful. When

ICU bed are scare then patients are more likely to experience premature discharge or even restricted access to critical care[53].

Within countries there is also variation in ICU bed availability. For example, in England, London as more than twice as many ICU beds per capita than the South (11.1 versus 5.4 ICU beds /100000 population) and ICU beds are increasing in London over time.

2.1 Models of hospital behaviour

Understanding how the various service lines in the hospital interact requires some appreciation for the various models of hospital behaviour. The traditional model of the private firm maximising profit in the context of allocating scarce resources has limitations in its application to hospital economics[54]. Models encompass a wide range of hospital behaviours and empirical evaluation is limited by a lack of clearly defined outputs and pricing mechanism, and hospitals do not respond to purely to profit, but instead also respond to social need[55].

Hospitals vary in their organisational structure so there is no single model that is general enough to describe the entire hospital setup. General theories of hospital behaviour aggregate across different hospital types, however, because of the diversity of hospital types, the more generalised theories are have less predictive value [55]. Another issue is the non-market structure of the hospital sector. The physician is the agent for the patient. In the UK specifically, the patient does not bear the direct costs of treatment. The UK government acts as both funder and health care regulator.

Some theories of the firm may have relevance for predicting hospital behaviour. Firms hire agents to produce outputs from which revenue is gained. The agents are concerned with the production process and the decision-making process related to allocative efficiency. The firm's behaviour is deterministic and responds to market forces in terms of inputs, outputs and prices, assuming perfect competition. There is some separation of management from ownership, some discretionary behaviour, and a degree of monopoly power. Profit maximisation is not a binding constraint. An objective function other than profit maximisation becomes less testable. Behavioural theories of firm behaviour concern themselves with bargaining processes and internal decision making and are less concerned about market

behaviour. These competing theories highlight specific issues when considering the behaviour of hospitals. The two methodological approaches that have been adopted, namely the organism model, where the hospital is considered a single entity, or the exchange approach focused on the behaviour of individuals within the institution[55]. Using the theory of the firm does not resolve the issue of readily identifiable prices or well-defined outputs.

The model proposed by Pauly and Redisch assumes physicians control resource allocation and aim to maximise profits[56]. This model does not consider the agency role and does not consider the market structure in any detail. The profit maximising model assumes that physicians are not employed by the hospital and the physician does not produce services directly sold by the hospital.

Whilst patients may have quality or quantity preferences, they have little chance to express these preferences because it is physicians and not patients who prescribe hospital care. It is the quality-quantity preferences of the physician that prevails.

Newhouse proposed that the administrator is the decision maker [57]. This model makes the simplifying assumption of being a single product firm and gives no consideration to the interdependency amongst producers in the hospital sector. There is general agreement about the objective function of the hospital with some quantity maximising function with a quality constraint. There remains a problem of identifying the decision maker. Harris argued that hospital output must be produced on demand, cannot be stored, and is poorly substituted between patients. Harris identifies both clinicians and administrators as being important but also as having their own objectives and constraints[58]. The model appreciates the ethical and medical constraints that the institution faces that are separate and to the economic motives. Hospital behaviour is geared towards short-run internal allocation problems that cannot be overcome by recourse to a pricing mechanism given the ethical constraints, the

clinicians objective function and the nature of health care goods. The Harris model is significant because it recognises that the internal structure of the hospital rather than the market structure is the dominant force in the theory of hospital economics[58].

In a multi-agent model described by Goldfarb *et al.*, multiple groups interact within the organisation[59, 60]. The various coalitions have conflicting interests, and the firm is a satisfying organisation rather than a maximising one. This translates to a shifting interest depending on which group holds the locus of control. Hospital motivation is the product of multiple actors, but it is still possible for a single actor to dominate. In the hospital setting the major actors are administrators and physicians. In this environment, physicians may wish to expand services in terms of quantity and quality, and managers may wish to work towards revenue maximisation or at least meeting the break-even constraint. If the hospital revenue exceeds the costs, then both hospital physicians and managers meet their objectives and there is no conflict. When hospital costs for providing a service exceed revenue gained then managers and physicians may have divergent goals. When there is no unified objective, the physicians often emerge as the dominant group. Hospitals may pursue an expansion of service even if it does not meet the break-even constraint. Managers may attempt to alter the means of production by exercising control over decisions for which they have some jurisdiction such as restricting capacity or staffing. Physicians do not have direct control of investment decisions but retain the rights of control over hospital resources. Physicians would resist attempts by administrators to constrain the service in the long run. A manager's ability to enforce budgetary restraint is often weak when there is conflict between service demands and cost containment. This can be seen in both elective and emergency care.

The emergency care pathway often includes identifiable individuals in immediate peril. This tension between budgetary restraint and the rule of rescue creates significant political and ethical challenges. Consolidating patients with similar needs may bring savings or improved quality through the more efficient use of physical resources and professionals in those areas. In this sense critical care services can be seen as a hospital within a hospital, where the implicit form of the objective function is:

$$U = u(V, CM, Q, P),$$

where U=utility, $V =$ volume of admissions, $CM =$ case mix, $Q =$ quality of care and $P =$ profit/financial surplus. Because quality is difficult to measure directly, we use mortality as a proxy measure. In meeting its objective function, the ICU faces several constraints, including the availability of technology and resources to provide treatments. The available patient constraint reflects the epidemiological and demographic characteristics of the community. In general there are two types of patients available for admission: necessary and discretionary[60]. When all necessary patients are admitted, then capacity is consumed by discretionary patients. An example of this tension is the trade-off between planned high risk elective surgery and emergency care. High-risk elective surgery such as vascular or major abdominal surgery require predicable ICU capacity. When there are high demand from emergency cases such as evidenced by the recent COVID-19 pandemic, then discretionary admissions, for example, from elective surgery may decrease. To maintain elective surgical capacity, critical care services may have to expand beyond the break-even constraint.

2.2 A brief history of critical care

Critical care services represent an organisational innovation focused on the patients most likely to deteriorate in the hospital. Critical care services are an important component of

health systems, accounting for up to 15-30 % of hospital expenditure and 1% of GDP and up to 3% of all health spending[61]. The ability of hospitals to provide effective and emergency services is contingent on ICU capacity. This makes critical care services an attractive focus for reforms focused on improving the efficiency. The aims of this thesis are to assess the role of two potential reforms, namely centralisation and specialisation, in improving the outcomes for critically ill patients.

The intensive care unit (ICU) is a specialised ward within the hospital where critical care is exclusively delivered. The modern ICU has its origins in the Polio epidemic in Copenhagen in 1952[62]. Copenhagen was the epicentre of the worst Polio epidemics recorded. In 1952, the main way to treat respiratory failure was the 'iron lung', of which there was only one available. An iron lung was a type of ventilator where the patient was completely sealed in a metal case and negative pressure was used to create the respiratory movements. These machines were not widely available and much like the COVID-19 pandemic, demand for mechanical ventilation quickly overwhelmed resources. At the height of the Polio epidemic, there were 50 patients a day hospitalised and about 10% needing mechanical ventilation[63, 64]. Copenhagen only had about 20 anaesthetists and hundreds would be required to provide 24-hour positive pressure ventilation by hand. About 1500 medical and dental students were arranged in 6-hour shifts to manually squeeze air into a patient's lungs via tracheostomy tubes over several months. Mortality from polio involving the brain stem went from 87% to 31%[62]. In addition to advancing our understanding of Polio, other insights emerged from the Polio epidemic. First, that positive pressure ventilation over prolonged periods was feasible. Second, that putting these patients in one place where doctors and nurses had expertise in managing organ failure made it possible to deliver these treatments reliably.

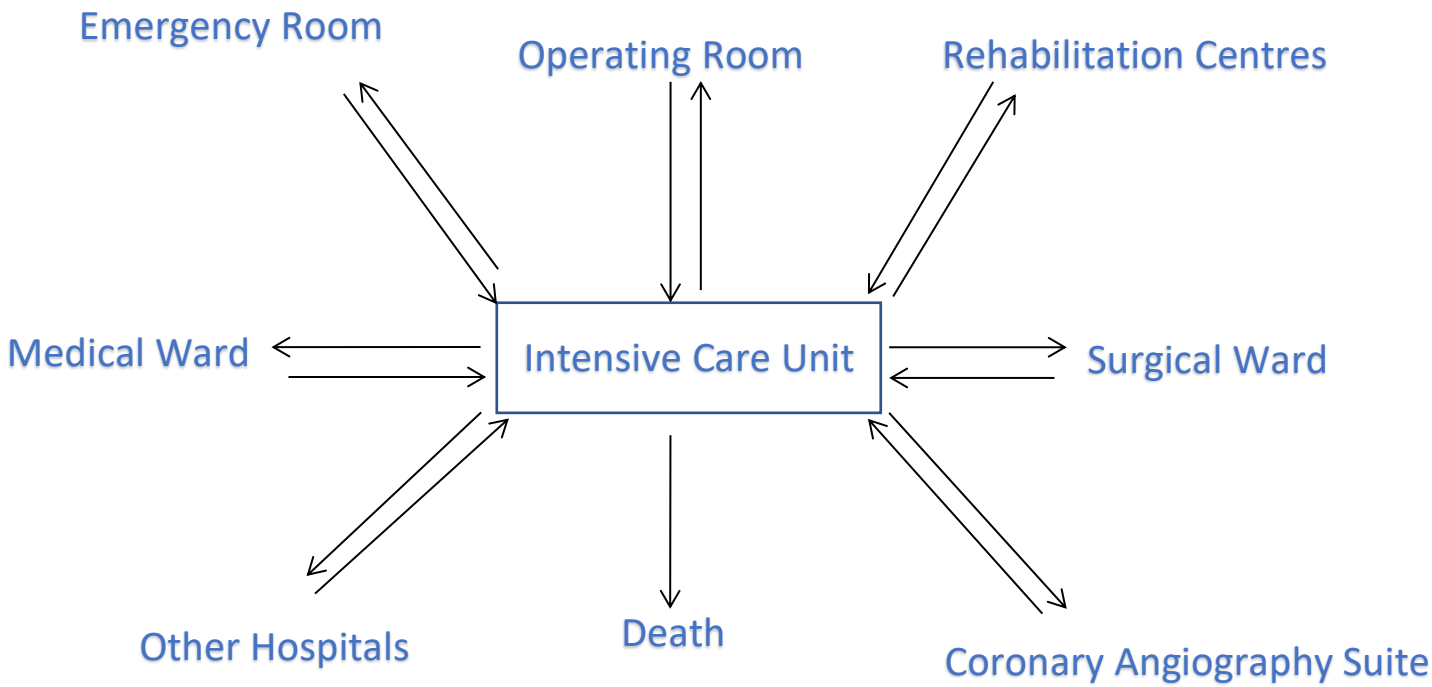
Third, the use of multidisciplinary teams was introduced [65]. In the 1960's the Engstrom became the first commercially available ventilator that was subsequently acquired by General Electric. In the 1960's, ICUs developed in the US, Europe, and Australia.

At the same time, other organ support technologies were being developed. The Dutch physiologist Willem Kolff introduced dialysis in 1943 and defibrillators became available in the 1950s[65, 66]. Technology for automated monitoring of vital signs became available in the late 1960s. The introduction of benchmarking revealed stark difference in outcomes and efficiency. Survival rates varied three-fold for the same severity of illness and the resources needed to produce similar survival for patients with similar severity of illness varied six-fold [65, 67]. These differences in performance were not explained by any single therapeutic intervention or case-mix and were likely due to differences in the organisation of critical care services[65].

2.3 ICU flow

The movement of patients in and out of the ICU is referred to as patient flow and is integrated with and dependant on other aspects of the hospital. Figure 2 describes the many routes in which patients can enter and leave the ICU. A patient may enter the ICU from the Emergency Department with community acquired pneumonia and leave the ICU to a medical ward or to a rehabilitation centre. A patient could require post-operative ICU care from the operating room, then develop an acute coronary event requiring coronary revascularization in the angiography suite then leave the ICU to go to another hospital.

Figure 2. A simplified example of the multiple entry and exit points into a single ICU within a hospital with a medical and surgical ward. Each line represents a potential patient journey.



Whether the patient is admitted to the ICU depends on the availability of ICU beds, whether there are patients that could be discharged from the ICI and whether there is available staff. This homeostasis can easily be disrupted by a surge in demand. Each of the units of demand may have conflicting interests. Since its inception, the organisation of the ICU within the hospital has been debated. Indeed, the ICU can be considered an organisational intervention providing a range of services. Within the hospital the ICU provides the capacity to care for acutely unwell patients. Within the broader health system, ICUs provide levels of care and clinical services that that are unavailable in other parts of the hospital.

There are some aspects of critical care provision that are universal and others that are speciality, geographically, health system or country specific. A recent report by the task force of the World Federation of Societies of Intensive and Critical Care Medicine sought to create an all-encompassing definition of what constitutes an intensive care unit[68].

2.4 What is an Intensive Care Unit (ICU)?

A global definition of an ICU must be sufficiently broad to include the variability that exists in the capacity to care for patients with the highest acuity of illness in health systems across the world. In this sense an ICU can be viewed in relative terms to the realities of a health system and the available resources.

ICUs employ one of two staffing models, namely open or closed units[69]. In the open unit model, the patient is admitted to the ICU by their treating physician and remain primarily under the care of the treating physician. The primary treating physician may have a longitudinal relationship with the patient. Typically, the attending physician does not have any specific training in the management of critical illness and may seek a consultation from an intensivist. In a closed ICU model, the primary responsibility for the patient is transferred to the intensivist. The intensivist in a closed unit would decide which patients are admitted, discharged, and make most treatment decisions within the ICU. Intensivist staffing within the US may not be generalisable to studies done outside the US. Whilst there is move towards closed ICUs in the US, a significant number of ICUs remain open in organisation and about 5% of ICUs have no intensivist cover at all[70]. A meta-analysis of open versus closed ICUs included 27 observational studies and concluded that closed ICUs were associated with lower hospital mortality and lower resource utilisation [71]. This literature has obvious limitations in that studies are exclusively US based and are several decades old. The arguments in favour of closed, intensivist lead ICUs are expertise with commonly used interventions, procedural training, improved consistency of evidence-based therapies, facilitation of multidisciplinary

working and improved care coordination. I shall henceforth consider all ICUs closed as is the case in the UK, Europe, Canada, Australia.

A comparison of the varying definition of ICUs used by several international certifying bodies is described in Table 1. The definitions broadly describe the types of staffing, expertise, types of patients and types of treatments provided. There are important differences between countries in terms of how ICU beds are counted. In the US ICU beds are defined by staffing, in Belgium by the characteristics of patients and in countries like Australia, New Zealand and the UK high dependency beds are considered ICU beds[72]. [73, 74]. Comparisons between countries can be problematic because of the fundamental differences in the definition of an ICU bed, even amongst similarly high-income countries.

Table 1.Definitions of an ICU (adapted from Marshall et al).

Certifying Body	Year	Definition
The Intensive Care Society (UK)	2013	An Intensive Care Unit (ICU) is a specially staffed and equipped, separate and self-contained area of a hospital dedicated to the management and monitoring of patients with life- threatening conditions. It provides special expertise and the facilities for the support of vital functions and uses the skills of medical, nursing, and other personnel experienced in the management of these problems. It encompasses all areas that provide Level 2 (high dependency) and/or Level 3 (intensive care) care as defined by the Intensive Care Society Document *Levels of Critical Care for Adult Patients* (2021)[75].
Society of Critical Care Medicine (US)[76]	1999	ICU serves as a place for monitoring and care of patients with potentially severe physiological instability requiring artificial life support. The level of care in the ICU is greater than that available in the ward and intermediate care unit
College of Intensive Care Medicine of Australia and New Zealand Minimum Standards for intensive Care Units Colleges of Intensive Care Medicine	2011	An ICU is a specially staffed and equipped, separate, and self-contained area of a hospital dedicated to the management of patients with life-threatening illnesses, injuries, and complications, and monitoring of potentially life-threatening conditions. It provides special expertise and facilities for support of vital functions

There are four domains that separate an ICU from routine hospital care.

1. Physical space

   A discrete geographic locale within the hospital where the sickest patients are cared for is a central component of the definition of an ICU.

2. Support and monitoring technology

   The ability to perform continuous monitoring of physiological parameters is a central differentiating feature of ICU care compared to ward care. This data is continuously displayed and used by clinicians to in caring for the patient. There are a range of organ supports available. Respiratory support includes oxygen, and the many modes of mechanical ventilation. Haemodynamic support is pharmacologic and mechanical. Renal support includes the various modes of renal replacement therapy. Nutritional support is in the form of enteral and parenteral nutrition. The specifics of organ support available would depend on the available resources and the population served. Sub-speciality ICUs such as neurosurgical ICU would have different organs supports and monitoring compared with a transplant or cardiac ICU.

3. Human resources

   The ICU team has specialist qualifications, and the level of care is more intensive than any other part of the hospital. In addition to physicians and nurses the team includes physiotherapists to support mobilisation and rehabilitation, respiratory therapists to manage mechanical ventilation, dieticians to manage nutrition of complex patients, a

pharmacist to manage drug interactions and optimise dosing in critical illness, a social worker to provide family and patient support, a microbiologists to assist in the management of infections and other staff to support patients and families during the ICU stay. [68]. Patients should have medical teams immediately available. The ratio of patient-to-intensivist ratio (PIR) should allow for sufficient attention to each patient. The ideal configuration for staff ratios is not well established and there may be interactions between different levels of staffing. A higher patient to nurse ratio may be safe if the patient to pharmacist or patient to physiotherapist ratios are lower. The optimal ratio may also be influenced by other factors. These include case-mix and turnover, other duties of the physician, physician support from other medical professionals and trainees, and technology support[77].

Many national bodies and health care payers specify the required patient to staff ratios to service the ICU. The current recommendation for PIR by the UK ICS is 12-14 patients to one intensivist. The best evidence for the effects of PIR in the UK context identified 7.5 as the optimal number at which hospital mortality was lowest[78]. A major limitation of this study was that details of other members of care teams were not included in the study[78]. The SCCM in the US does not make specific recommendations for PIR, suggesting individual ICUS should decide what number allows for the delivery of the service commensurate with the institution's expectations.

The SCCM task force on ICU staffing does recommend that PIR less favourable than 14:1 negatively impact perceptions of quality, stress, and patient care[77].

The UK ICS also makes recommendations about patient to nurse ratios, patient to physiotherapist ratios and patient to resident doctor (doctors in training) ratios. The

UK ICS recommends one patient to one nurse for level 3 patients and two patients per one nurse for level 2 patients. These recommendations are similar to those in the position paper from the European Federation of Critical Care Nursing.

2.5 Critical care services provided

An ICU can be described by the types of services offered beyond the immediate support for failing organs. The ICU can offer specialist services to specific cohorts of patients like burns. cardiac or neurosurgical patients. The ICU could be a regional referral centre for smaller local hospitals. The ICU could also offer limited levels of care aimed at specific groups of patients such as monitoring for post-surgical patients in high dependency units that are sometimes considered part of the integrated critical care service. Critical care services may vary in terms of the specific staffing configuration and types of patients admitted.

ICU bed provision appears to have some relationship to the case-mix within the ICU. Data from six European countries contributing to a multicentre clinical trial (namely the UK, France, Germany, Spain, Netherlands, and Belgium) suggests that the percentage of patients in the ICU with sepsis increases when fewer ICU beds are available[79]. This would be expected given that demand for ICU for sepsis patients is exogenous and sepsis patients would displace more endogenous types of patients such as those needing elective surgery.

The UK Intensive Care Society has standardised the level of intensity of treatment as level 0(ward level), 1 (enhanced care), 2(critical care) to 3(critical care) (Table *2*). These levels of treatment form the basis on which hospitals allocate resources and seek remuneration in the UK.

Table 2. Levels of care as designated by the Levels of Adult Critical Care Second Edition[75].

Level of care	Description	Example
0-Ward level care	Requires hospitalisation Needs can be met through normal ward care	A patient stepped down from a higher level of care at low risk of deterioration.
1- Enhanced care	Patient discharge from higher level of care Patients in need of additional clinical interventions or clinical input Patients receiving critical care outreach service support	Patients requiring active treatment by the critical care outreach team. Requires continues oxygen therapy Boluses of intravenous fluids Diabetes- receiving continuous insulin Parenteral nutrition
2- Critical care	Patients stepping down from Level 3 to Level 2 care Patients needing single organ support (exceptions Basic cardiac and respiratory simultaneously is considered Level 2 and Advanced respiratory alone is considered Level 3) Requiring 2 or more forms of basic organ support or monitoring Patients with severe physiological derangements, who cannot be cared for elsewhere.	Mask/ hood CPAP Patients receiving renal replacement therapy Invasive neurological monitoring Continuous infusions to manage seizures

3-Critical care	Patients receiving advanced respiratory support Patients receiving at least two forms of organ supports at an advanced level (non-respiratory). Patients with delirium or agitation already receiving level 2 care.	Invasive mechanical ventilation Patient receiving renal replacement therapy and vasopressor support

ICU care is expensive and increases in ICU spending is a major contributor to rapidly escalating national health care expenditure[80, 81]. The more efficient organisation of critical care services will help improve quality and lower costs and is an important policy priority. The rise in ICU costs may be partly due to differences in practice patterns at individual hospitals. Differences in local resources, financial incentives and a lack of clearly defined indications and guidelines has also led to significant variations in the way ICUs are organised. Much of ICU care falls within the scope of the grey area in medicine with the following attributes:

1. There are few clinical guidelines
2. The potential for marginal harm is relatively small. The harms of reducing access to other patients or unnecessary invasive procedures are not always identifiable in advance.
3. The benefits to the patient are idiosyncratic.

The organisation of the health system is a key component of optimal critical care services and evaluating its efficiency requires an understanding of the broader health system[82-86]. Interpreting studies that describe outcomes from critical illness across countries is challenging because of the differences in how it is utilised. Differences in observed outcome need to consider the differences in types of patients admitted to the ICU, discharge practices, end of life care practices and alternate care settings such as hospice and skilled nursing facilities.

One commonly cited comparator is perioperative mortality, particularly following high-risk surgery. First, one large difference that might explain differences in outcome is the types of patients that are offered surgery in the first place. A study comparing the population of Alberta with Massachusetts found that the per capita rate for major vascular reconstructive surgery was 3.4 times higher in the US, with significantly more ICU utilisation and worse mortality[87]. Second, post-operative ICU admission varies substantially, even amongst countries with similar ICU capacity. In Europe about 6.3% of high-risk patients in the UK are admitted to the ICU[88]. In Sweden this figure is 3.2% and in Spain 12.5% and in Latvia 21% of high-risk patients are admitted to the ICU[88]. About 73% of the high-risk patients that died in the European Cohort study were never admitted to the ICU at any point[88].

The role of critical care services in providing end-of-life care is hugely variable even amongst high income countries. A study of patients dying with cancer in 7 high income countries, namely the US, Canada, Belgium, England, Germany, Norway and the Netherlands, found ICU admission was at least twice (40.3%) in the US compared to the other 6 countries, (<18%) [89]. The US has a high utilisation of ICU services to deliver end of life care and one in 5 Americans die in the ICU[90]. By comparison only 5% of deaths in the UK occur in the ICU[91].

These wide variations in discretionary patients raise the spectre of the epidemiological phenomenon referred to as the "Will Rogers effect" [92]. This refers to the moving of a patient from one location to another and lowering the mortality of both groups. If large numbers of perioperative patients with a low risk of death on average or palliative patients with a high chance of death are admitted to the ICU then any comparison of overall ICU mortality between critical care services is likely to reflect the case-mix more than the quality of care provided. Comparisons across disease specific cohorts are therefore more informative.

Sepsis is an important public health problem amounting to almost 50 million incident cases in 2017 and accounting for 20% of all deaths worldwide[93]. Sepsis refers to infection associated with organ dysfunction and is usually treated within an ICU setting. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3) is described in Appendix one[94]. Sepsis is treatable and requires timely intervention. Although mortality from sepsis has decreased in the last three decades, it remains the most common cause of death in hospitalised patients[95]. There are wide variations in sepsis mortality between countries. The highest mortality is in sub-Saharan Africa and the lowest mortality is in North America and Europe[93]. Amongst high-income countries mortality varies substantially. A systematic review of sepsis mortality across North America, Europe and Australia found 30-day mortality to be 19.6%, 23.6% and 18.7% respectively[96]. Some of the observed differences maybe attributable to adherence to best practice guidelines, such as the surviving sepsis bundle, though the benefits of adherence to the sepsis bundles have been inconsistent [97]. In one study a very marginal increase in adherence to the sepsis protocol in Spain was associated with a 5% reduction in mortality[98]. In contrast, a study of sepsis patients in 114

US hospitals found a more than 60% improvement in adherence to the sepsis protocol but no change in the trend for mortality[99]. These findings are consistent with a landmark study by the RAND group, which found variations in mortality in hospitalised patients with myocardial infarction and heart failure were not explained by variations in processes of care[100]. Some hospitals that provided a high level of therapies identified to be effective but were observed to have poor outcomes and other hospitals achieved good outcomes without providing these effective therapies [100]. Even within the same country mortality varies widely. Within the US, data from about 3000 hospitals found a 23% to 56% risk adjusted mortality for sepsis [101]. In Japan, the hospital mortality from sepsis ranged from 12% to 48%[102]. These studies point to the idea that improvements occur through complex pathways and not necessarily through measurable single processes. Indeed, despite the absence of a "magic bullet" to treat sepsis, steady improvements in mortality have continued to occur[103]. The organisation of critical care services, namely through centralisation or specialisation, and the resources available may explain these observations. More needs to be understood about these organisation strategies so that we can better implement them.

2.6 Regional Differences in critical care availability

Across the world, there is heterogeneity in patient populations, the underlying structure of the health system, health spending and cultural values that result in wide differences in the supply and utilisation of critical care systems. Despite these differences, comparisons between countries can provide useful insights about the delivery of care or illuminate mechanisms that may improve outcomes if implemented elsewhere. Two valuable lessons have been learned from international comparisons of critical care services. First, responding to the growing threat of disasters and pandemics requires an understanding of available resources both nationally and internationally. The COVID-19 pandemic has highlighted the value of knowing the capacity of the critical services to respond to surges in demand and the value of timely access to critical care services. A study across 14 European countries found substantial national and regional variation in spatial access to ICU beds[104]. Countries with lower ICU accessibility had higher COVID-19 case fatality ratios. Germany had a large, widely distributed pool of ICU beds and was well positioned to respond to a surge in demand. In contrast countries like Denmark, Italy and Sweden had fewer ICU beds per capita that were spatially concentrated resulting in localised shortages of ICU beds[104]. Second, the generalisability of research undertaken in one system to other health systems depends on understanding the structure of the health system in which the study is done, understanding differences in staffing, the types of patients treated, and practice patterns. The decision to admit patients to the ICU reflects patient characteristics, and wider organisational aspects of the hospital system such as the array of services available in and outside the ICU. Studies

describing the variation in ICU utilisation suggest that factors other than case-mix contribute to the broad variation in ICU utilisation.

A common currency used to compare the structure of critical care services is the number of ICU beds standardised to the population. ICU bed availability varies considerably worldwide and although there is no consensus about the right number to serve a population, we can make judgements about the adequacy of bed supply by practice patterns[105]. We could infer that a high proportion of very sick patients denied ICU admission because of bed availability, premature ICU discharge and improvements in mortality as ICU bed supply increases would suggest an under-supply of ICUs[106-108]. Alternatively, having up to 40% of patients admitted to the ICU with a predicted risk of death <2% or 20% less than with a risk of death <1% , would also suggest an inefficiency in ICU utilisation as this group of patient are unlikely to derive any benefit from ICU care[109, 110]. Comparisons of ICU beds per capita across countries provide useful information about the critical care resources available and hence the context in which reforms are considered.

ICU bed availability was moderately correlated with the percent of GDP spending on health(r=0.632)(Figure 3) [105]. The US is the biggest spender on health care (16.8% of GDP). The Organisation for Economic Co-operation and Development (OECD) average for GDP spending on health is 8.8%[111]. The UK spends 10.2% of GDP on health, similar to most other European countries.

Spending on critical care services accounts for about 1% of GDP in the US and about 0.1% in the UK[81]. This huge difference in the supply and spending of ICU services invariably results in international variations in utilisation. In the US Veterans Affairs Health System, low risk

hospitalised patient (with a predicted mortality of <2%) make up about half of all ICU admissions[110]. In the UK, low-risk patients account for a median of 13%, (IQR 9-16%) of ICU admissions, ranging from 0% to 36.6%. The underlying supply of critical care services therefore plays an important role in the context of considering further reforms.

Even amongst countries with similar GDP spending on health and similar overall standards of living across Europe, there is a substantially large difference in ICU bed provision. There are two ways to consider the population level provision of critical care services. One is to consider the number of beds per capita of population and the other is to consider ICU beds in relation to acute hospital beds.

In 2020, the average number of ICU beds across the Organisation for Economic Co-operation and Development (OECD) countries was 12 beds per 100,000 population, ranging from 2.3 per 100,000 population in India, 2.7 and 4.7 per 100,000 population in Indonesia and China respectively to 25.8, 28.9 and 33.9 per 100,000 population in the US, Austria and Germany respectively[112](Figure 4). In 2020 the UK is below the OECD average at 10.5 ICU beds per 100,000 population.

Figure 3. Relationship between ICU availability and percentage Gross Domestic Product (GDP) spent on health care [72].

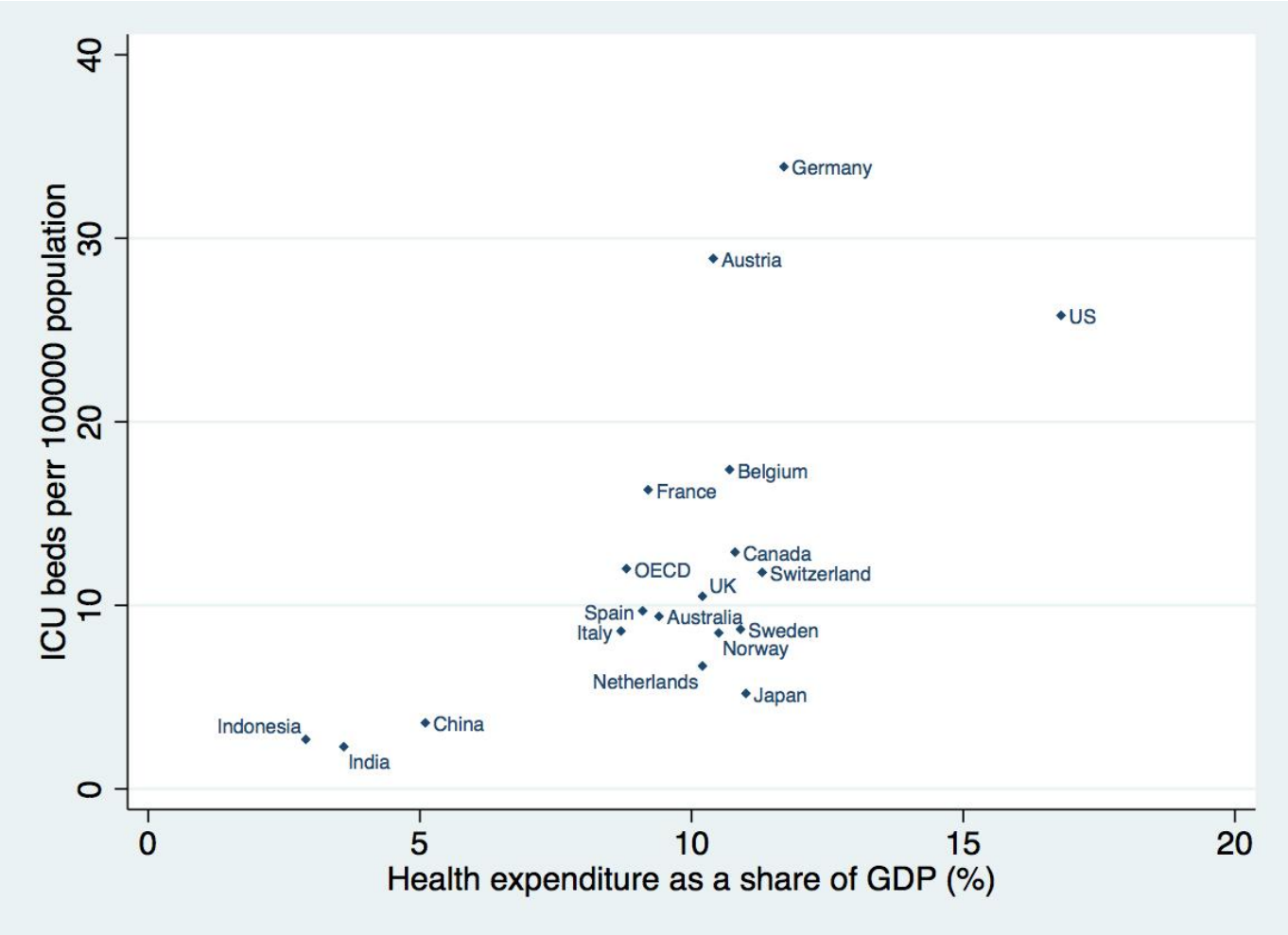
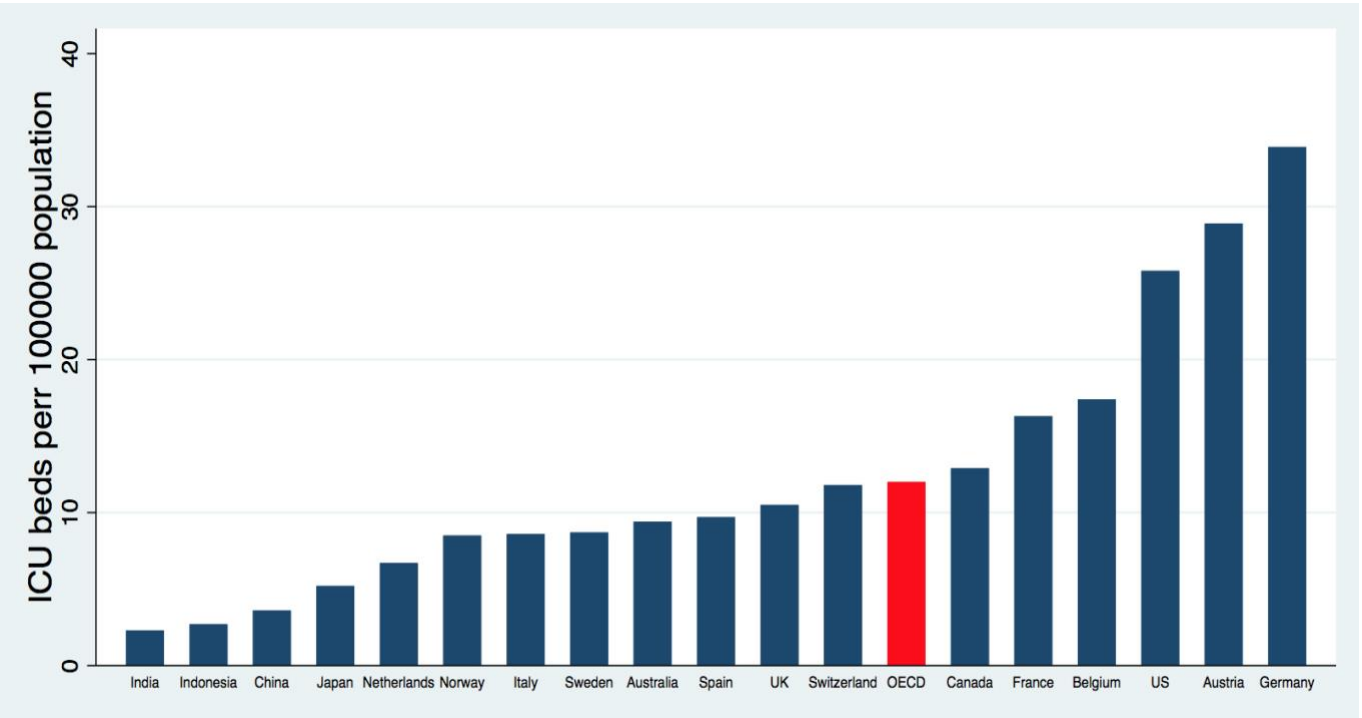


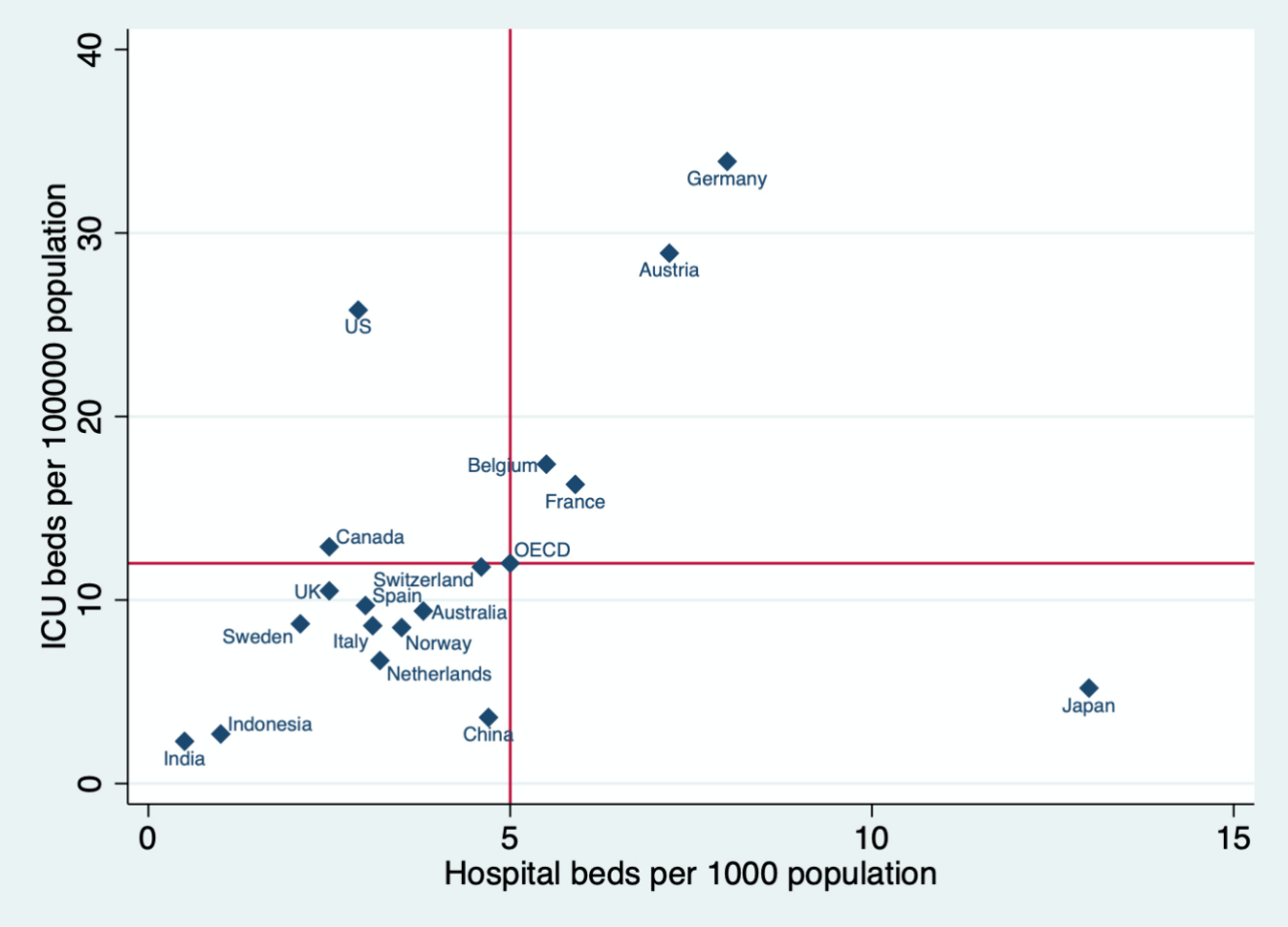
Figure 4. Variation in ICU bed availability across OECD countries [72].



ICUs are organised within hospitals and a shortage of available acute hospital beds will impact how an ICU functions. Shortages of hospital beds delays stepping patients down from the ICU, increasing the inefficient use of ICU and impeding patient flow. The differences in resources outside the ICU such as the availability of Long-Term Acute Care Hospitals (LTACHS) and Skilled Nursing Facilities (SNF), may explain variability in length of stay and discharge practices

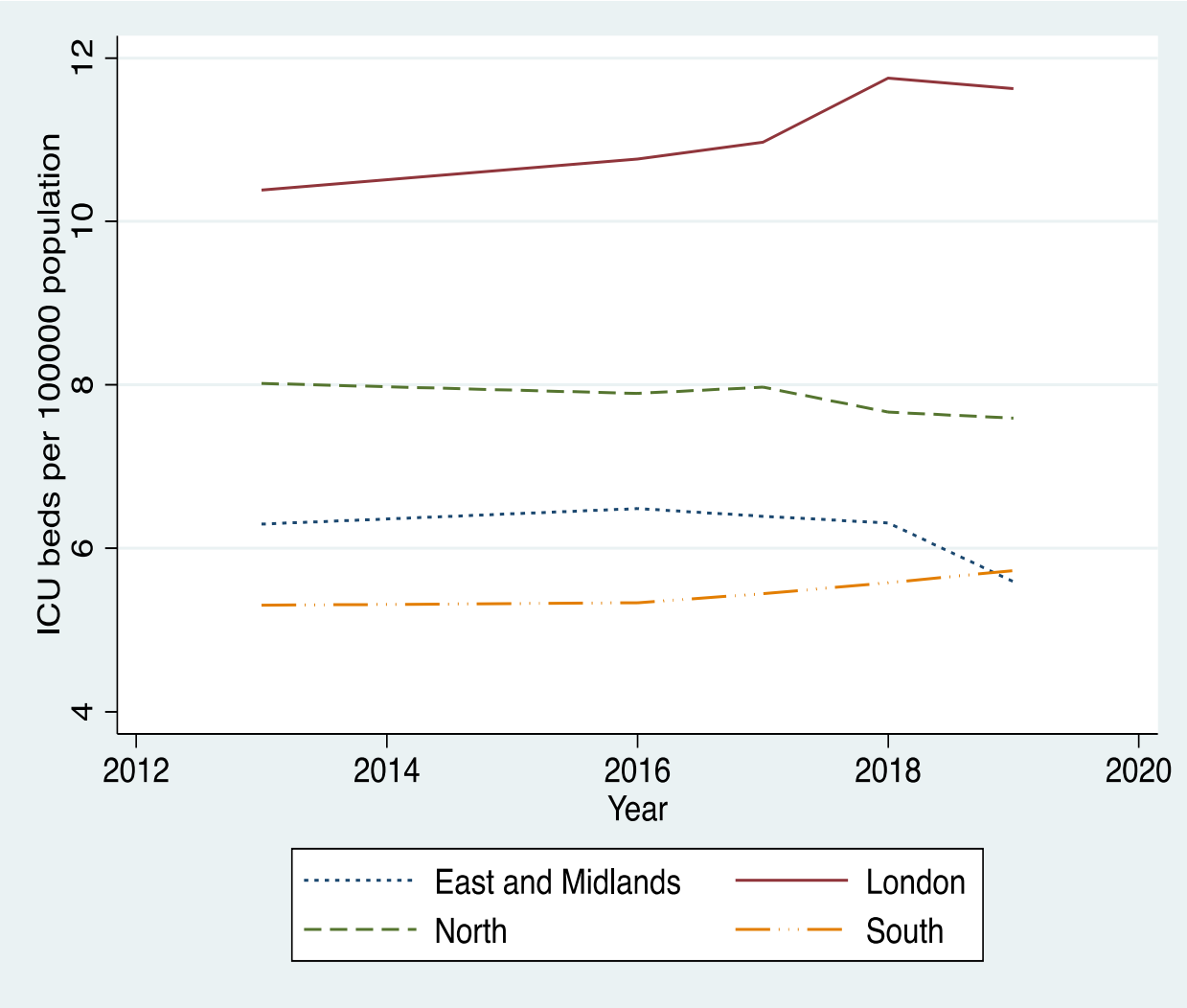
between countries[113]. The provision of ICU beds is weakly correlated with acute hospital bed provision ($r = 0.646$) but varies significantly between countries, Figure 4 and Figure 5 [105]. The OECD average for acute hospital bed provision is 5 beds per 1000 population. The US has relatively low acute hospital bed provision (2.9 per 1000 population) and high ICU bed provision resulting in ICU beds accounting for about 8% of acute hospital beds. The OECD average for the percentage of ICU beds to acute hospital beds is about 2.4%. Japan, on the other hand, has high hospital bed provision (13 per 1000 population) and low ICU bed availability (5.2 ICU beds per 1000000 population). Austria and Germany have both high ICU bed provision and high acute bed provision. The UK has low acute hospital bed provision (2.5 per 1000 population) and ICU bed provision below the OECD average. This relative under supply of both ICU beds and acute hospital beds underscores the importance of optimising the organisation of these scarce resources and the need for co-ordination across individual organisations.

Figure 5. The relationship between ICU and acute hospital bed availability.



Differences in critical care provision are not limited to the country-level. Even within countries there is substantial regional variation in ICU provision. In the US, the West region has the lowest ICU bed provision (24.2 ICU beds per 100,000 population) compared with the South East (31.7 ICU beds per 100,000 population)[114]. Figure 6 describes ICU bed provision across 4 regions in England between 2013 and 2019. London had the highest ICU bed provision,(mean 11.1 ICU beds per 100,000 population, 95% CI 10.6-11.6) [115, 116]. ICU bed provision was lowest in the South (mean 5.4 ICU beds per 100,000 population, 95% CI 5.3-5.6). The North and East and Midlands had 7.8 ICU beds per 100,000 population 95% CI 7.7-8.0) and 6.4 ICU beds per100,000 population 95% CI 5.9-6.5). In addition to London having higher ICU per provision per capita, it appears that ICU capacity in London is growing, and other regions are static or decreasing (Figure 1.5). This would suggest that supply of ICU beds is not driven by the underlying needs of the population. The recent COVID-19 pandemic highlights the regional problems that may occur when there is a sudden influx of patients putting the system under strain, even when the national picture appears stable[117]. Higher levels of capacity strain have been associated with higher mortality from COVID-19[117, 118]. Surges in demand from COVID-19 cause capacity strain, that has potentially eroded the gains from emerging treatments[86]. The capacity strain that occurred during the COVID-19 pandemic also increased mortality for non-COVID-19 patients[118].

Figure 6. Regional variation in ICU bed provision across four geographic areas- London, North, South and the East and Midlands between 2013 and 2019.



[Data from NHS England and the Office for National Statistics [115, 116]

Chapter three: Data

Abstract

Aims: This chapter describes the Case-Mix Program Database (CMPD), the ICNARC Coding Method (ICM), the $ICNARC_{H\text{-}2018}$ risk adjustment model.

Method: The method of data collection and validation adopted by the ICNARC for the Case-Mix Program Database. The criteria for describing a high-quality clinical database as developed by the Directory of Clinical Databases (DoCDat) were used to rate the CMPD. The $ICNARC_{H\text{-}2018}$ risk adjustment model is also described. The data from the CMPD used in this thesis is from between 1 January 2010 and 31 December 2016 and are briefly summarised.

Results: The mean quality level of the CMPD in terms of the DoCDat criteria is 3.7 (with 1=worst and 4 =best). The total number of ICU episodes was 1,008,023. There were 41868 readmissions, 31358 inter-ICU transfers and 1513 patients that were both readmitted and transferred between ICUs during the same episode. The total number of unique patients was 933,284, from 231 ICUs. The $ICNARC_{H\text{-}2018}$ risk adjustment model predicted that acute hospital mortality would be 19.4% and the observed mortality was 21.2%.

Conclusion: The CMPD is a high-quality clinical database that is representative and reliable. The $ICNARC_{H\text{-}2018}$ risk adjustment model compares favourably with other commonly used models in terms of overall fit and discrimination for predicting acute hospital mortality.

3.1 The Case-Mix Program dataset

Clinical databases are immensely valuable in evaluative research, and in informing planning and organisation of services [119]. Clinical databases are designed with medical decision making in mind, and include more detailed data on patient history, comorbidities, procedures, and outcomes than administrative databases. Inaccurate and poor-quality data will inevitably lead to misleading conclusions. To promote quality and appropriateness of use, a quality assessment checklist tool was developed, similar to the CONSORT checklist used in randomised control trials[119]. The resulting instrument consists of 10 items, four relating to coverage, and six relating to reliability and validity of the data. The instrument is described in Table 3. Level I represents the least rigor and Level IV represents the most rigor.

The Intensive Care National Audit and Research Centre (ICNARC) is an independent charity established in 1994.[120]. ICNARC co-ordinates national audit from adult intensive care units in England, Wales, and Northern Ireland as part of the Case-Mix Program (CMP). Data is validated locally and then centrally and then pooled into the Case-Mix Program Database (CMPD). Adult intensive care units are defined as ICUs, mixed ICUs, and High Dependency Units (HDUs), combined general and coronary care units and medical and surgical units admitting patients greater than 16 years old. The CMPD covers 100% of all adult ICUs in England, Wales, and Northern Ireland although participation is voluntary. The CMP is listed in the Department of Health's Quality Accounts as a national audit by the National Advisory Group on Clinical Audit and Enquiries (NAGCAE) for Acute Care. The CMP is open to both public and private (independent) sector ICUs.

CMP data is collected prospectively and abstracted into standard forms by trained data collectors according to precise definitions. Data collectors are trained prior to data collection.

In addition to data collectors, training includes critical care physicians and nurses. Retraining of staff also happens on a regular basis. Other training opportunities include regular one-day Data Workshops aimed at data collection/validation, Teleconferences to discuss common problems, clarify definitions and share best practice, and the CMP Process Guide providing practical guidance for effective data collection.

Data abstraction is performed by chart review. Data is collected from consecutive admissions and submitted for inclusion into the CMPD at least every six months. Local validation occurs in accordance with ICNARC CMPD specification. The data then undergoes central validation for completeness, and inconsistencies with data are returned to the local ICU for correction. This validation process is repeated until all queries are resolved and the data is then added to the CMPD. The CMPD team undertake over 600 validation checks. Each ICU receives a Data Analysis Report which provides trends over time of how the ICU is performing relative to other ICUs.

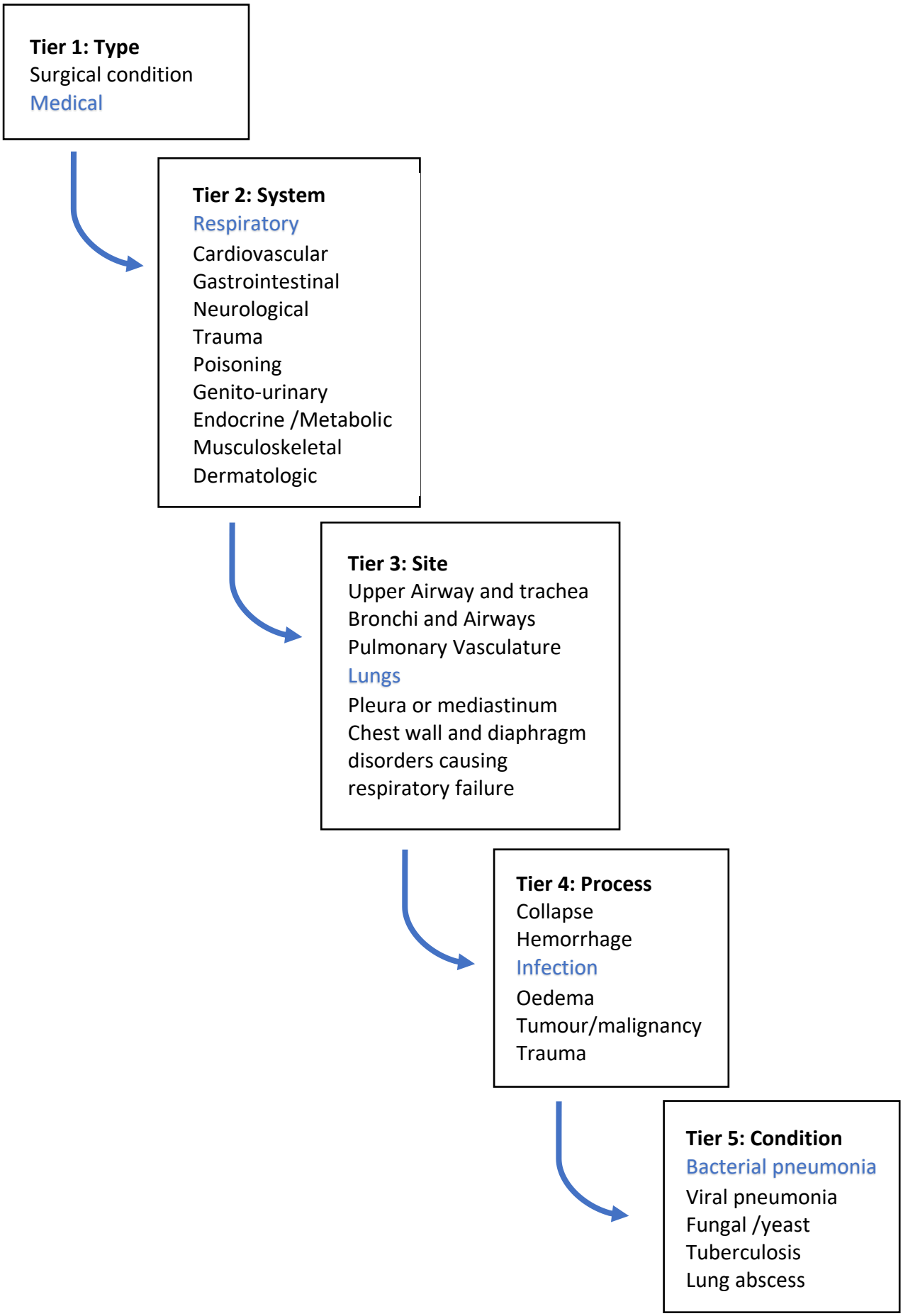
The information on reason for admission to the ICU is coded using the ICNARC Coding Method (ICM) [121]. The ICM is a five-tiered hierarchical method for coding based on 1. Medical or surgical condition 2. The body system 3. Anatomical site 4. The pathological condition and 5. The condition requiring admission.

Figure 7 describes the ICM with the illustrative example of bacterial pneumonia.

Table 3. Directory of Clinical Databases (DoCDat) checklist[119]

	Level I	**Level II**	**Level III**	**Level IV**
A. Extent to which the eligible population is representative of the country	No evidence or unlikely to be representative	Some evidence eligible population is representative	Good evidence eligible population is representative	Total population included
B. Completeness of recruitment of eligible population	<80% or unknown	Many -80-89%	Most -90-97%	All or most->97%
C. Variables included in the database	• Identifier • Admin info • Condition or intervention	• Identifier • Admin info • Condition or intervention • Short-term or long-term outcome	• Identifier • Admin info • Condition or intervention • Short-term or long-term outcome • Major Known confounders	• Identifier • Admin info • Condition or intervention • Short-term or long-term outcome • Major Known confounders
D. Completeness of variables	Few <50%	Some 50-79%	Most 80-97%	Almost all >97%
E. Collection of raw data	Few<70%	Some 70-89%	Most 90-97%	Almost all >97%
F. Explicit definition of variables	None	Some <50%	Most 50-97%	Almost all >97%
G. Explicit rules for deciding how variables are recorded	None	Some <50%	Most 50-97%	Almost all >97%
H. Reliability of coding of conditions and interventions	Not tested	Poor	Fair	Good
I. Independence of observations of primary outcome	Outcome not included or independence unknown	Observer neither independent nor blinded to intervention	Independent observer not blinded to intervention	Independent observer blinded or not necessary
J. Data validation	No Validation	Range or consistency checks	Range and consistency checks	Range and consistency checks plus external validation

Figure 7 The ICNARC Coding Method (ICM). The figure uses the example of bacterial pneumonia.



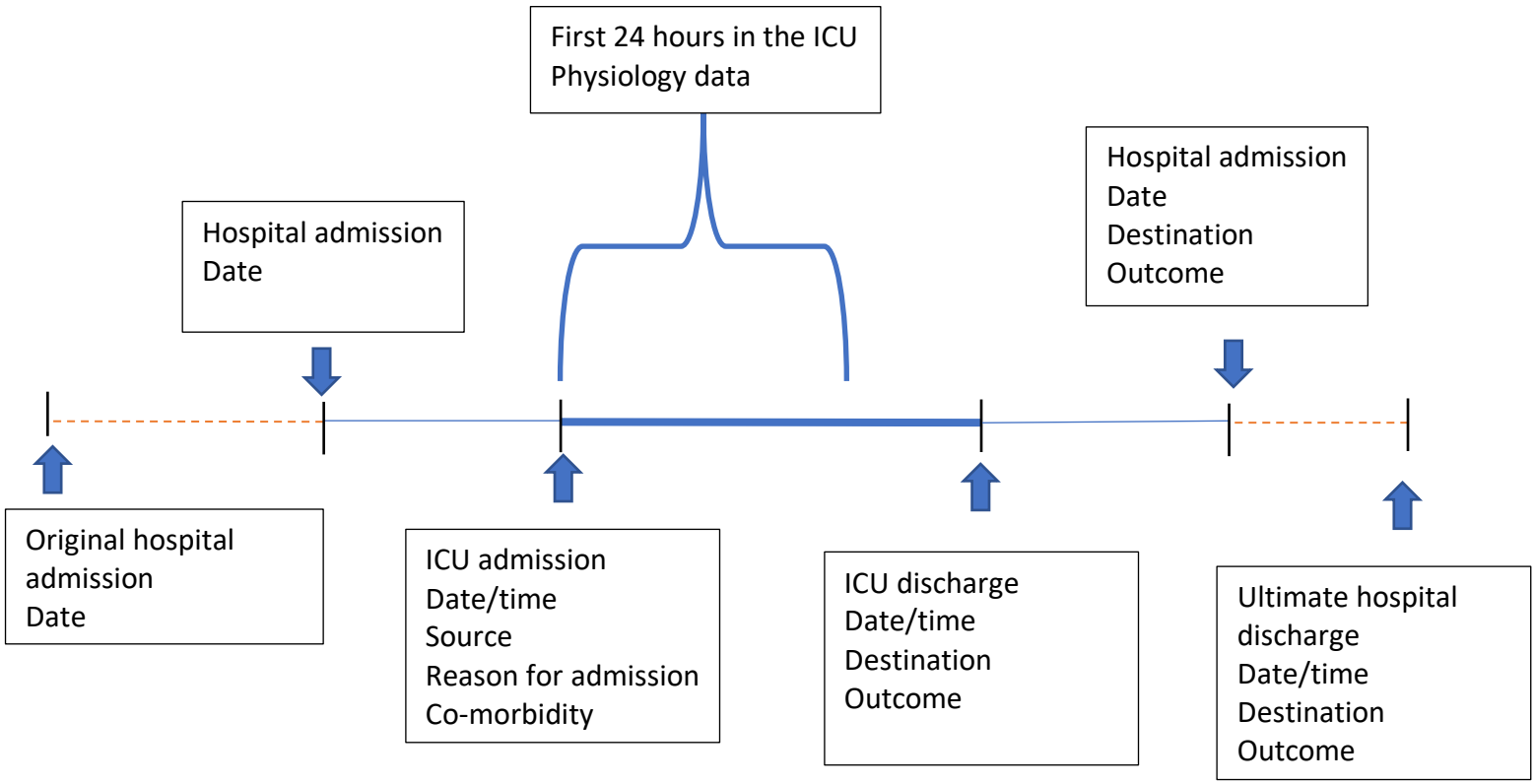
The ICM can be used to identify the primary and secondary reason for admission, based on the first 24 hours of ICU care as well as the ultimate reason for admission based on information from discharge. Admissions can be identified at any tier of the ICM, for example, we could identify all surgical patients, all patients with cardiovascular surgery, all patients with diseases of the lung, all patients with infection and all patients with viral infection. We can also identify patients using physiological criteria such as the Sepsis-3 definition[122].

The data included in the CMPD included demographic data as well as details about hospitalization. A schematic description of the timeline is described in

Figure 8. All patients are followed from the original data of hospitalization to their ultimate hospital discharge. Raw data is collected rather than categorical or aggregated data. Patient admissions are allocated a unique identifier and records are reversibly anonymized. Patients are cross linked by postcode for deprivation scoring. Physiological data are collected to calculate the Acute Physiology and Chronic Health Evaluation (APACHE II) and ICNARC risk adjustment scores[123, 124].

Data on source of admission and type of surgery is defined as per the National Confidential Enquiry into Perioperative Deaths. Elective surgery is planned surgery at a time suitable to the patient. Emergency surgery refers to surgery that should be performed immediately, simultaneously with resuscitation or as soon as possible thereafter. ICU and hospital survival are recorded. For patients transferred between hospitals and ICUs, the ultimate hospital discharge is also recorded. Length of ICU stay is calculated in hours from the dates and time of admission to ICU discharge. For patients transferred between hospitals, ultimate hospital length of stay is the total duration of the episode. Readmissions to the ICU are identified by the postcode, date of birth and gender and confirmed by the participating ICU.

Figure 8. Timeline for CMPD data collection[120]



3.2 Performance of the CMPD against the DocDat criteria

1. Representativeness

The CMPD includes about 99% of all adult general critical care ICU and HDU in England, Wales, and Northern Ireland and 100% of all general adult ICUs. (Level 4)

2. Completeness of recruitment

ICUs included consecutive admissions (Level 4)

3. Variables included

The CMPD includes all major known confounders in the form of raw data. Outcomes beyond hospital discharge are not included in the CMPD. (Level 3)

4. Completeness of variables

All variables were at least 95% complete (Level 4).

5. Collection of raw data

All continuous data was collected as raw data. (Level 4)

6. Explicit definitions

A detailed data collection manual is provided to each ICU using definitions developed after wide consultation. (Level 4)

7. Explicit rules

The CMPD has explicit rules for deciding how variables are recorded and provides training and re-training. (Level 4)

8. Reliability of coding

The quality of coding is difficult to verify, however, the ICM has been found to have high inter-rater reliability [125]. ICUs test the reliability of data by recollecting a random sample of admissions selected by ICNARC. These reliability assessments are done 2-3 times per year. (Level 3).

9. Independence of observations

The outcome variable is mortality and is not subject to gaming. (Level 4)

10. Data validation

Validation of the CMPD includes logic, range, and consistency checks, although there is no external validation. (Level 3).

The CMPD has many strengths. First, a major strength is the wide coverage. The CMPD includes all general adult ICUs in England, Wales, and Northern Ireland. Second, the explicit definitions used, the data collection manual and the precise dataset specification minimize inter-observer variability. The training provided has been effective in minimizing inter-observer variability[126]. A lack of clear definitions and instruction about the timing of data collection has been identified as a source of inter-observer variability in other clinical datasets that are commonly used such as the APACHE II and APACHE IV databases[127, 128]. The inter-rater variability for the ICM coding of reason for admission was 79% in the previous study[125].

This compares well with the inter-observer variability of the APACHE IV database of 68%[128]. Third, the CMPD collects raw data that allows for improved risk adjustment. This has led to the development of the ICNARC risk model, based on UK data. [123].

3.3 The ICNARC$_{\text{-H-2018}}$ risk adjustment model

Risk adjustment has its limitations when comparing ICU performance. The commonly used models (e.g., APACHEII, APACHE IV, SAPS II) are unable to distinguish between a sick patient on admission and a patient that gets sicker over the first 24 hours because of poor care. [129]. In 2006 ICNARC published a validation of four models-the APACHE II, APACHE III, Simplified

Acute Physiology Score (SAPS II) and the Mortality Probability Models (MPM version II)[130]. Even with recalibration these models do not fit well across different populations[130]. The ICNARC model outperforms these models in both overall fit and discrimination[131]. Variables relating to past medical history did not improve model performance. This is likely because chronic conditions are reflected in the physiologic derangements and admitting diagnosis. The variables describing chronic medical conditions are very severe. Very severe cardiac disease refers to dyspnoea, angina, or New York Heart Association Functional Classification Class IV, documented 6 months prior to admission. Severe respiratory disease refers to permanent shortness of breath with light physical activity due to pulmonary disease present six months prior to admission End stage renal disease refers to patients requiring chronic renal replacement at least 6 months prior to admission. This includes peritoneal dialysis, haemo-dialysis, or chronic hemofiltration.

The ICNARC model has been recalibrated to ensure risk predictions are as accurate as possible. Further refinements since the model was introduced in 2007 have been non-linear modelling of physiological variables and incorporating more data from the ICM hierarchical model for reasons for admission to the ICU. The latest iteration is the ICNARC $_{\text{H-2018}}$ model.

Table 4. The variables included in the ICNARC$_{H\text{-}2018}$ risk prediction model[123, 131, 132]

Physiological parameters during the first 24 -hours after ICU admission
○ Highest hear rate
○ Lowest systolic
○ Highest temperature
○ Lowest respiratory rate
○ Lowest PaO_2/FiO_2 ratio
○ Lowest arterial pH
○ $PaCO_2$ associated with lowest pH
○ Highest blood lactate
○ Total urine output
○ Highest urea
○ Highest creatinine
○ Highest sodium
○ Lowest white blood count
○ Lowest platelet count
○ Glasgow Coma Scale
Age in years
Past medical history
○ Severe liver disease- biopsy proven cirrhosis, portal hypertension, hepatic encephalopathy
○ Metastatic disease
○ Haematological malignancy
○ Severe respiratory disease
○ Immunocompromised
○ Severe cardiac disease
Dependency prior to hospitalisation assessed 2 weeks prior
○ Independent
○ Minor assistance
○ Major assistance for
○ Fully dependant for activities of daily living
Cardiopulmonary resuscitation 24-hours prior to admission
○ In-hospital
○ Out of hospital
Mechanical ventilation within 24 hours of ICU admission
Source of admission
○ Emergency department
○ Operating theatre (elective/emergency)
○ Ward
○ Other acute hospital
Primary reason for ICU admission using ICM
Interactions between physiological variables and:
○ Other physiological variables
○ Past medical history
○ Interventions (CPR, mechanical ventilation)
○ Primary reason for admission
Interactions with age and past medical history

The CMPD uses rigorous methods of data collection and validation and compares favourably to other clinical databases. The CMPD meets the DoCDat criteria of a high-quality database. Results of analysis from the CMPD are representative and permit reliable comparisons with national and international data[120].

3.4 A summary of the data included in the Case-Mix Program

Table 5 provides a brief description of the CMPD from 1 January to 31 December 2016 used in this study. The total number of ICU episodes was 1,008,023 from 231 ICUs. Episodes that were excluded were: 41,868 readmissions, 31,358 inter-ICU transfers and 1,513 patients that were both readmitted and transferred into the ICU. There were 933,284 patients. Readmission accounted for 4.3% of all critical care episodes. For patients with multiple episodes only data from the index admission was used for risk adjustment.

The mean age of patients in the CMPD was 61 years and 55% were male. About 75% of patients were functionally independent and 81.8% had no recorded severe co-morbidities. Most critical care admissions were medical patients (57%). Elective surgical patients constituted about 25% and emergency surgery the remaining 18% of ICU admissions. Less than 1%of ICU admissions require ward level (Level 0) or enhanced care (Level1). The majority of patients require some form of organ support or monitoring i.e., Level 2 (53%) or Level 3 (45%) care. The level of care required varied across admission types. Elective surgery patients predominantly required Level 2 care (80%). Only 44% of medical patient required Level 2 care and 54% required level 3 care. About 46% of emergency surgery admissions require Level 2 care and 54% required Level 3 care. Overall hospital mortality was 21.2% and varied across

admission types with medical patients having a 30% mortality, elective surgery 3.7% and emergency surgery 23.7%.

3.5 Risk standardised acute hospital mortality (RSMR)

Mortality is a significant outcome for both patients and providers and occurs frequently in the critical care population. Unadjusted mortality however does not provide meaningful insights into institutional performance. To provide a broad description of the variation in mortality across ICUs I estimated the acute hospital risk standardised mortality rate (RSMR). This is a single summary measure that includes all ICU admissions, across all diagnoses measured until acute hospital discharge. This measure does not capture those patients discharged from hospital that subsequently die shortly thereafter.

The RSMR was calculated using a logistic regression model with acute hospital mortality as the outcome variable and the $ICNARC_{H\text{-}2018}$ expected mortality as the control variable. The acute hospital $RSMR$ for ICU j was defined as the ratio of the observed to expected deaths of ICU_j multiplied by the average mortality rate[133].

$$RSMR_j(\%) = \frac{Observed\ mortality_j}{Expected\ mortality_j}\ X\ Average\ ICU\ mortality$$

I describe the results using a funnel plot to display the RSMR graphically in Figure 9. The solid purple line represents the overall average mortality. The long-dashed lines are the boundaries for the 95% control limit and short dashed lines are the boundaries for the 99.8% control limit. The total number of patients treated in each ICU is represented on the X-axis. ICUs represented by blue dot fall within the 99.8% control limit, those ICUs in green dots are outliers with better-than-expected mortality and those ICUs in red dots are outliers in having a worse than acute hospital predicted mortality. The wide dispersion highlights the wide centre level variation in performance.

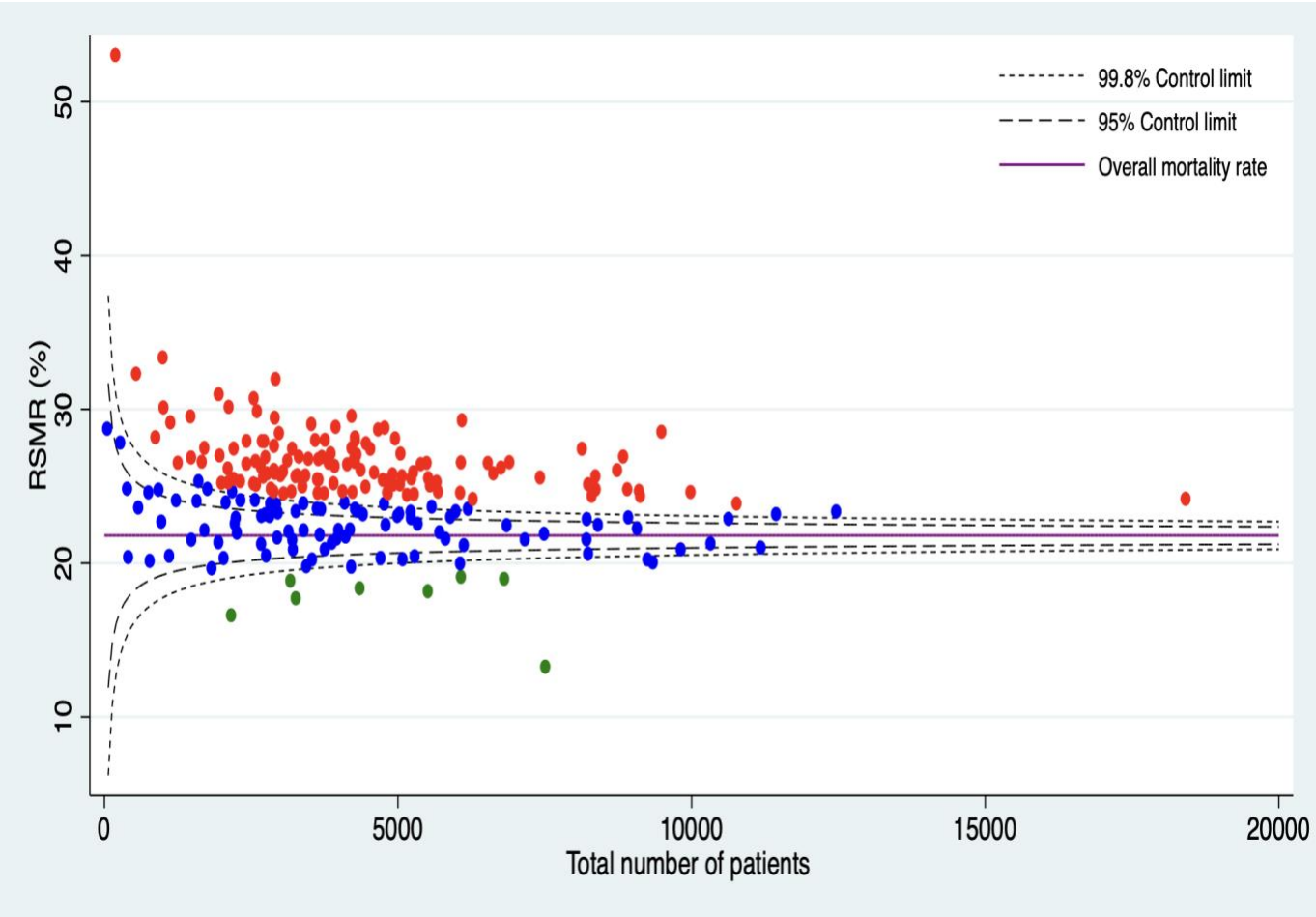
Table 5. Description of data in the CMPD (N=933284)

Variable	Value
Age(years) Mean (95% CI)	61.1 (61.0-61.1)
Race*, n (%)	
○ White	841450 (90.2)
○ Asian	31884(3.4)
○ Black	20630(2.2)
○ Other/mixed	38499(4.1)
Sex, n (%)	
○ Female	419533(45.0)
○ Male	513750(55.0)
Residence prior to admission, n (%)	
○ Home	911294(97.6)
○ Health related institution	15970(1.7)
○ Non- health related institution	6009(0.6)
Country n, (%)	
○ England	833673(89.3)
○ Wales	58185(6.2)
○ Northern Ireland	32011(3.4)
○ Isle of Man	1100(0.1)
○ Scotland	1017(0.1)
○ Channel Islands	391(0.0)
Functional status prior to ICU admission[#]	
○ Fully independent	704514(75.5)
○ Some assistance with activities of daily living	213182(22.8)
○ Fully dependant	9726(1.0)
APACHE II score	15.6(15.6-15.6)
$ICNARC_{H\text{-}2018}$ score	16.8(16.8-16.8)
$ICNARC_{H\text{-}2018}$ predicted acute hospital mortality	19.4% (19.4-19.5%)
Admission type[$]	
○ Medical	530483(56.8)
○ Elective surgery	229057(24.5)
○ Emergency surgery	173651(18.6)
Infection with 24hors of admission n, (%)	294634(31.6)
Sepsis diagnoses with 24hrs of admission	273001(29.3)
Septic shock	54419(5.8)
Requiring renal replacement therapy within 24 hrs of ICU admission	53577(5.7)

Organ failure n, (%)	
○ Cardiac	268672(28.8)
○ Respiratory	271652(29.1)
○ Renal	128303(13.8)
○ Haematological	34556(3.7)
○ Neurological	36996(4.0)
Highest level of care received in the ICU n, (%)	
○ Level 0	6529(0.1)
○ Level 1	8913(1.0)
○ Level 2	497013(53.3)
○ Level 3	424146(45.5)
ICU length of stay (hours) mean, (95%CI)	109(109,110)
Hospital length of stay (days) mean, (95%CI)	18(18-18)
ICU mortality n, (%)	136799(14.7)
Hospital mortality n, (%)	197917(21.2)

*821(0.09 %) missing ethnicity, [#]5862(0.6%) missing functional status, [$]2560(0.3%) missing highest level of care

Figure 9. Funnel plot of Risk Standardise Mortality Ratio. The solid purple line represents the overall average mortality, the long-dashed line represents the 95% control limit, and the short-dashed line represents the 99.8% control limit).



3.6 Conclusion

The CMPD is a high-quality clinical database. It is an invaluable resource for quality improvement, providing data that is that is representative and reliable in assessing the structure, process, and outcomes from critical illness. Comparing performance of any individual ICU requires comparative audit to enable benchmarking. This requires some account of patient characteristics including underlying chronic illness, severity of acute illness and other demographic factors. The most widely used outcome measure in critical care is mortality. Mortality is easily measured, not subject to gaming and patient centred. The most used form of mortality used to compare institutional performance is the risk standardised mortality ratio (RSMR). The RSMR is a ratio of the observed to expected mortality. The expected mortality is calculated using the sum of probabilities based on a risk prediction model. The $ICNARC_{H\text{-}2018}$ risk adjustment model is applied to the CMPD and compares favourably with other commonly used models in terms of overall fit and discrimination for predicting acute hospital mortality. The use of a highly calibrated risk adjustment model allows for meaningful conclusions to be drawn from observational data.

Chapter four - Literature review: The Volume-Outcome relationship

Abstract

Introduction: The relationship between caseload volume and outcome has been described across a range of medical and surgical conditions. The volume-outcome relationship has been used to support centralisation of critical care services. This relationship has, however, been inconsistently described across different cohorts of patients treated in the ICU.

Aim: A systematic review and meta-analysis was undertaken to assess the association between volume and outcome in critically ill adults.

Methods: A systematic search of MEDLINE, EMASE and Google scholar for English language articles published between 1 January 2000 and 12 December 2018. Information extracted included year of study, patient characteristics, definition of volume, outcomes (ICU hospital or fixed time mortality), type of methodology used (multilevel modelling, risk adjustment) and organisational details.

Results: Twenty-three studies were included in the meta-analysis comparing the highest volume category to the lowest volume category. The results showed a significant reduction in mortality with increased volume (OR 0.86, 95% CI 0.78-94), however, there was substantial heterogeneity ($I^2 = 98.95$). The subgroup of studies done outside the US did not find a significant volume outcome relationship (OR 0.89, 95% CI 0.74-1.04) compared with studies done in the US (OR 0.82, 95% CI 0.74-0.90).

Conclusion: There was significant association between ICU volume and mortality. A major issue identified was the lack of a gold standard in the definition of high and low volume. There were substantial heterogeneity and subgroup identified variations in this relationship depending on patient cohorts and whether the study was done in the US or outside the US.

4.1 Understanding the volume outcome relationship in the context of critical illness

Since the seminal work by Luft and colleagues in 1979 that showed improved mortality for a range of surgical procedures when the hospitals surgical caseload volume increased, this volume-outcome relationship has been the subject of extensive investigation[134] To date most studies have focused on surgical procedures and describe a favourable volume-outcome relationship[9]. However, research on the volume outcome relationship present methodological challenges. Differences in study estimates for the same procedure related to cut-off values have led to difficulties in translating the volume-outcome relationship into policy. The volume-outcome relationship has lent support to two policy ideas, namely 1) minimum volume standards and 2) centralisation[17].

Minimum volume standards require that hospital deliver services at a specified volume threshold, almost entirely related to surgical procedures. In most cases it is based on annual volumes and the usual consequence is the withholding of reimbursements if the service falls below the threshold value. Germany, Canada, the Netherlands, Switzerland, and Austria have introduced minimum volume standards for hospitals providing a small number of high-risk surgical procedures[17, 135]. The specific set of surgical procedures, as well as the volume standards for the same procedure varied substantially across countries, highlighting the limited evidence available to inform thresholds [17]. It remains unclear as to whether the introduction of minimum volume standards has had the desired effect on quality in these countries[17].

Centralisation involves the consolidation of services into fewer units to operate at a higher caseload volume[136]. In the UK the centralisation of major trauma, stroke, and paediatric services have been associated with improved patient outcomes[137-139]. Countries like

Norway, Germany, and Denmark also report improved outcomes and greater efficiency with centralisation of complex services [140] [141, 142]. The trade-off with both minimum volume standards and centralisation of services is the potential to create barriers to access, with patients having to travel longer distances to get care.

The relationship between volume and quality for non-surgical patients and critical illness is less clear. The volume-outcome relationship for patients with community acquired pneumonia is inconsistent. Amongst 3,243 hospitals participating in the National Pneumonia Quality Improvement Project, higher patient volumes of Medicare beneficiaries (>65years old) were not associated with improved mortality in community acquired pneumonia[143]. In contrast, analysis of data from Medicare administrative claims for fee for service beneficiaries found higher volumes to be associated with lower 30-day mortality[144]. This study found an attenuation of the volume outcome relationship with volumes beyond 210 patients per year providing very few marginal benefits[144]. Difference in the reported outcomes from these studies probably relate to the limitations of risk adjustment using administrative data undertaken in the latter study[143-145]. Similar findings are described for heart failure. Hospital volume was not associated with lower 30-day mortality in patients enrolled in the Get with The Guidelines-HF clinical registry[146]. In contrast, studies performing risk adjustment using administrative datasets describe a favourable hospital volume outcome relationship for heart failure[144, 147].

The volume-outcome relationship in critical care is of particular importance because of the high prevalence of critical illness and the high cost to the health system. Changes to the way the service is organised and delivered could bring substantial improvements in quality and

efficiency [71, 78, 148]. The reorganisation of critical care services to meet minimum volume standards or centralisation such that the highest risk patients are treated in ICUs within geographic regions in a tiered fashion akin to trauma networks, could match patient need with expertise and scarce resources to maximise outcomes and efficiency[148].

These policies face several challenges. First, high quality evidence in favour of the volume outcome relationship in critical care is lacking. Second, these policies require a central authority to regulate the system. Third, these policies could create barriers to access time sensitive care and increase care fragmentation. Patients with multiple diagnoses may be required to travel between several providers to access care. Fourth, developing the infrastructure to support reorganisation of services and interhospital transfers of patients may be costly. The aim of this review is to evaluate the relationship between high volume critical care services and mortality. I sought to review the methodologies applied and the organisational factors that may be mechanistically relevant.

4.2. Materials and methods

4.2.1. Search strategy

I reviewed the literature to identify studies that described the volume-outcome relationship for patients cared for in the ICU. I conducted a systematic search of MEDLINE, EMASE and Google scholar for English language articles published between 1 January 2000 and 12 December 2018. The search terms included intensive care, critical care, hospital volume, ICU volume as well as common medical conditions associated with ICU utilisation and regionalisation. All studies were combined using Rayyan, a web-based tool to manage references. References of included studies were reviewed to identify additional references.

We excluded studies prior to 1 January 2000 because of the significant changes to clinical practice and clinical practice prior to this time[149].

4.2.2 Study selection

4.2.2.1 Identifying studies

All retrieved records were reviewed and assessed. Titles and abstracts were screened for obvious exclusion. Full text articles were retrieved and reviewed to determine if they met the inclusion specification. Studies included were of adult patients receiving critical care in ICUs for conditions that would commonly be admitted to a general ICU. Studies of specific high risk-surgery patients were excluded because these reflected in part the surgeon's technical proficiency and were not considered generalisable to critically ill patients.

4.2.2.2. Search strategy

critical care [mesh] OR intensive care[mesh] OR neuro-intensive care[mesh] OR neuro-critical care [mesh] OR pulmonology/critical care[mesh] OR liver diseases/surgery[mesh] OR lung diseases/surgery[MESH] OR coronary artery bypass[mesh] OR trauma[mesh]OR angioplasty[mesh] OR myocardial infarction[mesh] OR cardiac arrest [mesh]OR sepsis OR vascular surgical procedures[mesh]OR pneumonia [mesh] OR mechanical ventilation [mesh]OR renal replacement therapy [mesh] AND utilization[subheading] AND (volume[Text word] OR frequency[Text word] OR competition[Text word] OR concentration[Text word] OR frequent [Text word] OR statistics[subheading]) AND (outcome assessment[mesh] OR outcome and process assessment[mesh] OR outcome[Text word] OR regionalization [text word]).

4.2.3. Data extraction

Data extraction was performed using a pre-specified data extraction form. Information extracted included year of study, patient characteristics, definition of volume, outcomes (ICU hospital or fixed time mortality), type of methodology used (multilevel modelling, risk adjustment) and organisational details. Risk of bias was examined using the previously established ROBINS-I framework [150].

4.2.4. Statistical analysis

The pooled odds ratio for mortality was obtained by a random effect restricted maximum likelihood estimator. Heterogeneity was evaluated using the Higgins I^2 with threshold values of 25%, 50% and 75% used to describe low, moderate and high heterogeneity [151] [152]. I used a Galbraith plot to identify outliers[151]. Given the clinical heterogeneity, I explored subgroups based on disease types of critically ill patients, namely gastro-intestinal, general, neurology, renal, respiratory and sepsis. I also considered subgroups of studies performed in the US and outside the US to explore the potential effects of the broader health system on the outcome. I evaluated publication using the Egger test and by funnel plots[153]. We performed a sensitivity analysis using the leave-one-out approach and the trim-and-fill method[154].

4.3. Results

The literature search yielded 10,995 records from which 241 full text articles were reviewed for eligibility (Figure 10). In total 28 studies were included with most being excluded because they did not include critical care patients or did not include mortality as the outcome. Table 6 provides the descriptive details of the studies included. Four studies did not provide sufficient data to be included in the quantitative analysis. A major issue identified was the lack of a gold standard in the definition of high and low volume. Initially it was planned to consider volume as a continuous variable, however, most studies reported volume as a categorical variable. There was considerable variation in the thresholds used to define volume and the number of categories used (terciles, quartiles, or quintiles). This was likely due to the specific dataset used. Consequentially, an ICU defined as high volume in one study would be considered low volume in another. For simplicity, we used the risk adjusted odds ratio for patient mortality treated in the highest volume category compared with the lowest category as reference. This approach limits the generalisability of these findings. The risk of bias assessment using the RoBINS-I is described in Table 7. Almost all studies were judged to be at serious risk of bias but not critical in any one domain. There were some studies where there was insufficient information on one or more domains to assess risk of bias. Overall, higher volume was associated with lower mortality, OR 0.86(95% CI 0.78-0.94). There were high levels of heterogeneity, $I^2 = 98.95\%$. Figure 11. The Egger test did not identify any small study effects (p= 0.302) and the non-parametric "trim and fill" method did not find any significant change in the overall effect size after imputation.

Sepsis was the only sub-group of more than one study associated with a mortality effect, OR 0.73[95% CI 0.58-0.88] but also with high levels of heterogeneity $I^2 = \ 96\%$. A second subgroup analysis found no significant volume-outcome relationship for studies performed outside the US (OR 0.89 95% CI 0.74-1.09) (

Figure 12).

We assessed publication bias visually with a funnel plot (Figure 13). The more precise studies with smaller standard errors are displayed at the top with the less precise trials at the bottom. The funnel plot did not identify any obvious publication bias. The Egger test was non-significant for small study effects(p=0.301). The trim-and -fill method imputed one study to the left of the funnel plot. This did not alter the final estimate (OR 0.85, 95% CI 0.77-0.93).

In terms of methodology, 3 studies used a multilevel model and 6 studies used generalised estimating equations to account for the clustering of patients with hospitals or ICU. The remainder of the included studies did not account for the clustering of patients.

Figure 10. Prisma Diagram of systematic literature search

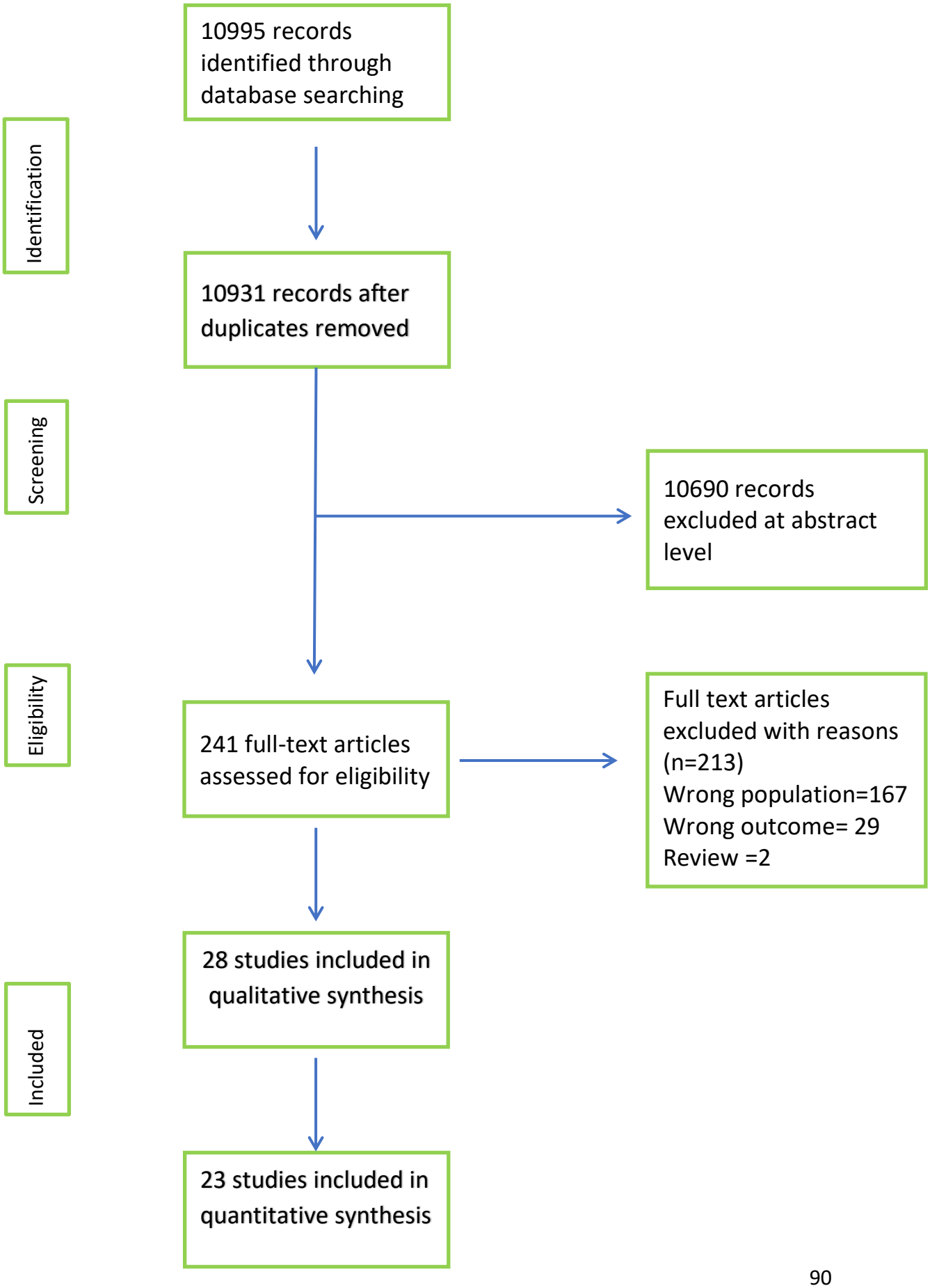


Table 6 Descriptive data of studies included in the systematic review and meta-analysis

Author	Year	Country	Definition of low and high volume	Population	Outcome	Data source
Iapichino	2004	12 European Countries	Annual Number of patients per ICU bed	ICU admissions	Hospital mortality	EURICUS-I database
Durairaj	2005	USA	Annual hospital volume Respiratory diagnosis <500 vs >1000	Medical ICU	Hospital mortality	Cleveland Health Quality Choice
Durairaj	2005	USA	Annual hospital volume GI diagnoses <400 vs >700	Medical ICU	Hospital mortality	Cleveland Health Quality Choice
Durairaj	2005	USA	Annual hospital volume Neurology diagnosis <400 vs >700	Medical ICU	Hospital mortality	Cleveland Health Quality Choice
Kahn	2006	USA	Annual volume of hospital admissions for MV <151 vs >400	Medical MV	Hospital mortality	APACHE Clinical information system
Lindenauer	2006	USA	Annual hospital volume pneumonia median 57 IQR (39-77) vs median 465 IQR 393-570	Pneumonia	30-day mortality	National Pneumonia Quality Improvement Project
Glance	2006	USA	Annual ICU case load <631 vs >1233	All ICU admissions	Hospital mortality	Project IMPACT
Needham	2007	Ontario, Canada	Annual hospital caseload for MV <20 vs >699	MV	30-day mortality	Linked administrative database of Ontario Province Canada
Peelen	2007	Netherlands	Annual ICU case load <29 vs >117	Sepsis	Hospital mortality	Dutch National Intensive Care Evaluation registry
Lecuyer	2008	France	Average annual case volume of haematology patients <12 vs >30	Respiratory failure	Hospital mortality	CUB-Rea database
Kahn	2009	Pennsylvania, USA	Annual hospital caseload for MV <100 vs >599	MV	Hospital mortality	Pennsylvania State Discharge records
Metnitz	2009	Austria	Number of patients per year per ICU bed	All ICU admissions	Hospital mortality	Austrian Centre for Documentation and Quality Assurance In Intensive Care Medicine
Reinikainen	2010	Finland	ICU beds <6 beds vs >6 beds	Sepsis	Hospital mortality	Finnsepsis Database
Powell	2010	USA	Annual ED case load of sepsis <146 vs >371	Sepsis	Hospital mortality	National Inpatient Sample
Darmon	2011	France	Annual volume of hospital admissions for MV <99 vs >282	MV	Hospital mortality	French Nationwide Database Ministry of Health
Nyugen	2011	USA	Annual case volume of RRT treated patients with AKI <10 vs >29	RRT	Hospital mortality	Project IMPACT

Nyugen	2011	France	Annual case volume of RRT treated patients with AKI <18 vs >58	RRT	Hospital mortality	CUB-Rea database
Gopal	2011	Birmingham, UK	Annual case load of MV < 125 vs >178	MV	ICU mortality	Birmingham Critical Care Network Database
Vaara	2012	Finland	Annual case volume of RRT treated patients with AKI <23 vs > 44	RRT	Hospital mortality	Finnish Intensive Care Consortium
Shahin	2012	UK	Annual ICU caseload with severe sepsis median 320 IQR (281-374 vs median 829 IQR (656-1022)	Sepsis	Hospital mortality	ICNARC
Moran	2012	Australia New Zealand	Annual volume of ICU admissions for MV per year <102 vs >800	MV	Hospital mortality	APACHE
Zuber	2012	France	Average Annual ICU case volume of septic shock in cancer patients <5 vs >12	Septic shock	ICU mortality	CUB-Rea database
Banta	2012	California, USA	Annual Hospital case load <8627 vs >19575	Sepsis	Hospital mortality	California Office of State-wide Health Planning and Development
Cooke	2012	USA	Annual ICU case load <20 vs >63	Medical MV	Hospital mortality	Veterans Affairs Inpatient Evaluation Centre
Fernandez	2013	Spain	Annual ICU volume <300 vs >799	ICU admissions	Hospital mortality	Sabadell Score clinical database
Kumamaru	2014	Japan	Annual hospital caseload of pneumonia patients per6 months <8 vs >17	Pneumonia	Hospital mortality	Japanese Diagnosis Procedure Combination database
Shahin	2014	UK	Annual number of MV patients per ICU per year <141 vs >480	MV	Hospital mortality	ICNARC
Walkey	2014	USA	Annual Hospital Sepsis volume <317 vs >604	Sepsis	Hospital mortality	University Health System Consortium (UHC) Clinical Database Resource Manager
Goodwin	2015	South Carolina, USA	Annual Hospital sepsis case load <75 vs >299	Sepsis	Hospital mortality	South Carolina State inpatient Database
Sasabuchi	2015	Japan	Annual ICU admissions <496 vs >748	General ICU	Hospital mortality	Japanese Diagnosis Procedure Combination database
Mehta	2016	California, USA	Annual hospital NIV caseload median 25 IQR (2-43) versus median 235 IQR(163-565)	NIV	Hospital mortality	California State Inpatient Database

AKI=acute kidney injury; RRT= renal replacement therapy; MV= mechanical ventilation

*Studies in red did not have sufficient data to be included in the quantitative analysis.

Table 7 Risk of Bias Assessment tool for Non-randomised Studies (RoBINS-I) tool [108]

Author	Year	Selection of Participants	Confounding Variables	Exposure Measurement	Outcome assessment	Incomplete Outcome Data	Selective Outcome reporting
Iapichino	2004	High	High	High	Low	High	High
Durairaj	2005	Low	High	Low	Low	Low	Low
Glance	2006	High	Unclear	Low	Low	Unclear	Low
Kahn	2006	Low	Low	Low	Low	Low	Low
Lindenauer	2006	Unclear	Low	Low	Low	Low	Low
Needham	2007	Low	Unclear	Low	Low	Low	Unclear
Peelen	2007	Low	Low	Unclear	Low	Low	Low
Lecuyer	2008	High	High	Unclear	Low	Unclear	High
Kahn	2009	Low	Low	Low	Low	Low	Low
Metnitz	2009	Low	High	High	Low	Low	Low
Powell	2010	High	High	High	Low	High	Unclear
Reinikainen	2010	Low	Low	Unclear	Low	Low	Low
Darmon	2011	High	Unclear	Unclear	Low	High	High
Gopal	2011	High	High	Low	Low	Unclear	Unclear
Nguyen	2011	High	Unclear	Unclear	Low	Unclear	Low
Banta	2012	Low	Low	Low	Low	Unclear	Unclear
Cooke	2012	Low	Low	Low	Low	Low	Low
Moran	2012	Low	Low	Low	Low	Low	Low
Shahin	2012	Low	Low	Low	Unclear	High	Low
Vaara	2012	High	High	Low	Low	Unclear	Low
Zuber	2012	Low	Low	Low	Low	Unclear	Unclear
Fernandez	2013	High	High	High	Low	High	High
Kumamaru	2014	High	High	High	Low	Unclear	Unclear
Shahin	2014	Low	Low	Low	Low	Unclear	Low
Walkey	2014	Low	Low	Low	Unclear	High	High
Goodwin	2015	Low	Unclear	Low	Low	Low	Low
Sasabuchi	2015	Low	Unclear	Unclear	Low	Low	Unclear
Mehta	2016	Unclear	Unclear	Unclear	Low	Low	Low

*Studies in red did not have sufficient data to be included in the quantitative analysis.

Figure 11. Forest plot odds ratio for survival comparing lowest (reference) to highest category volume.

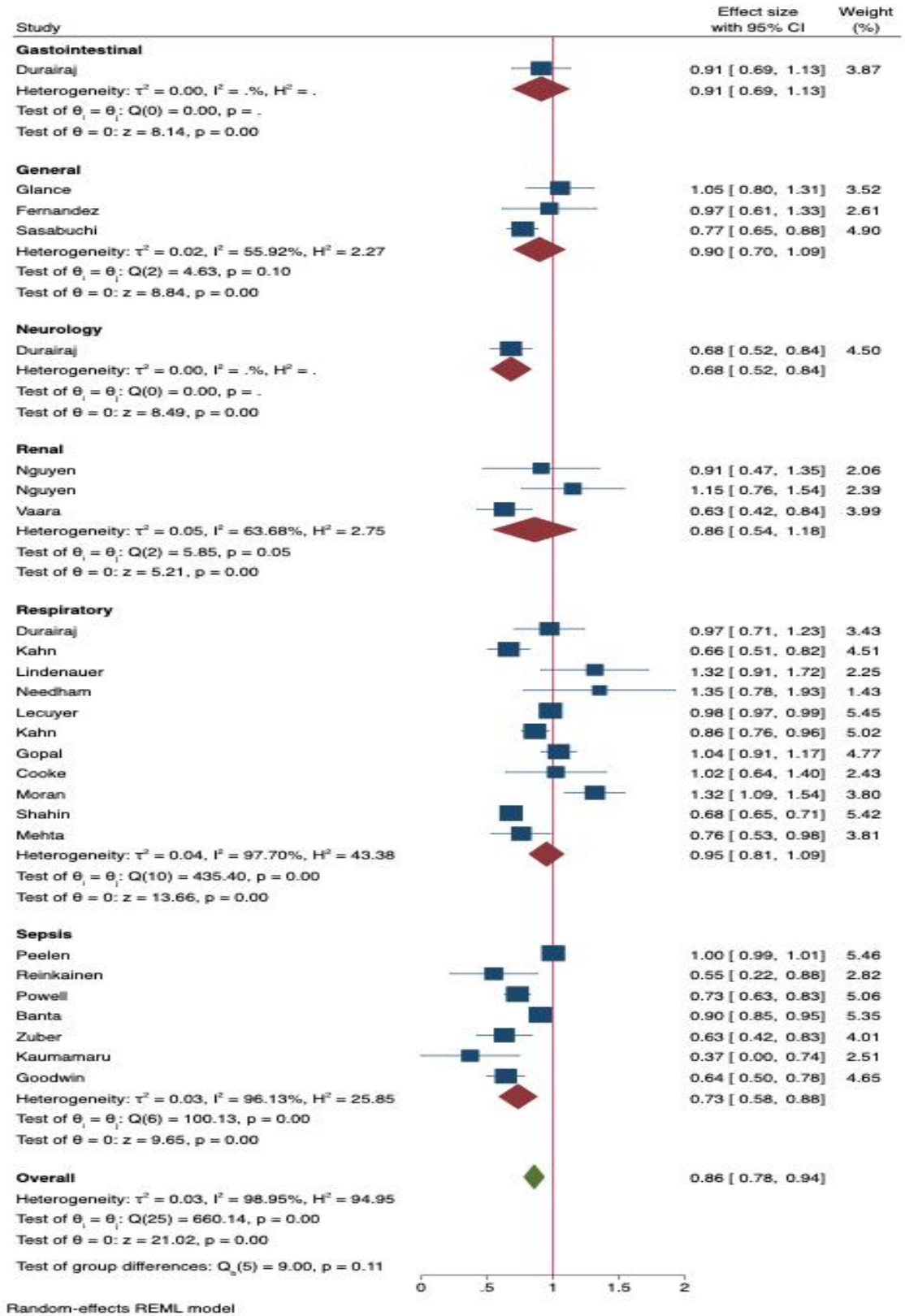


Figure 12. Subgroup analysis comparing studies done in the US to those done outside the US.

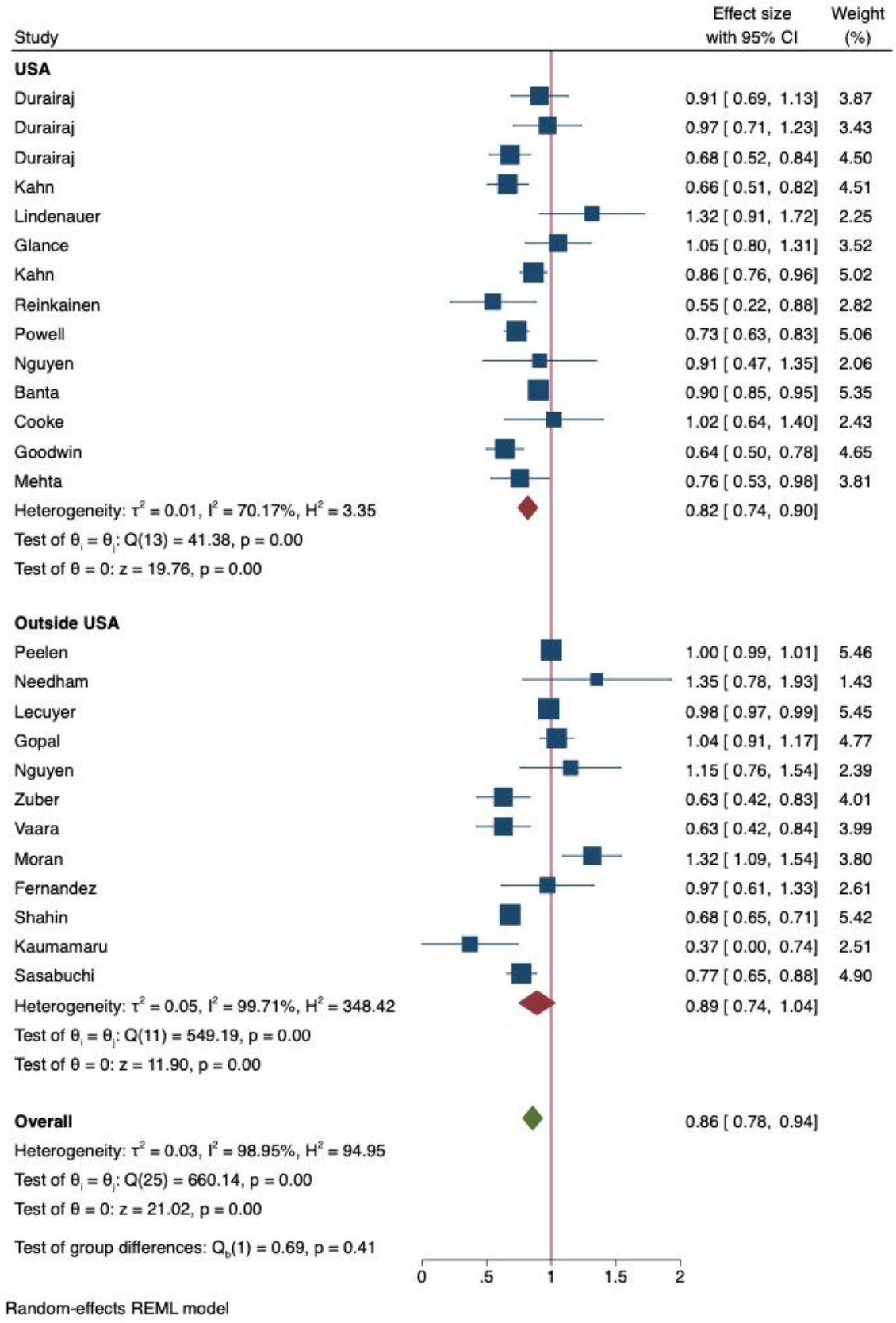
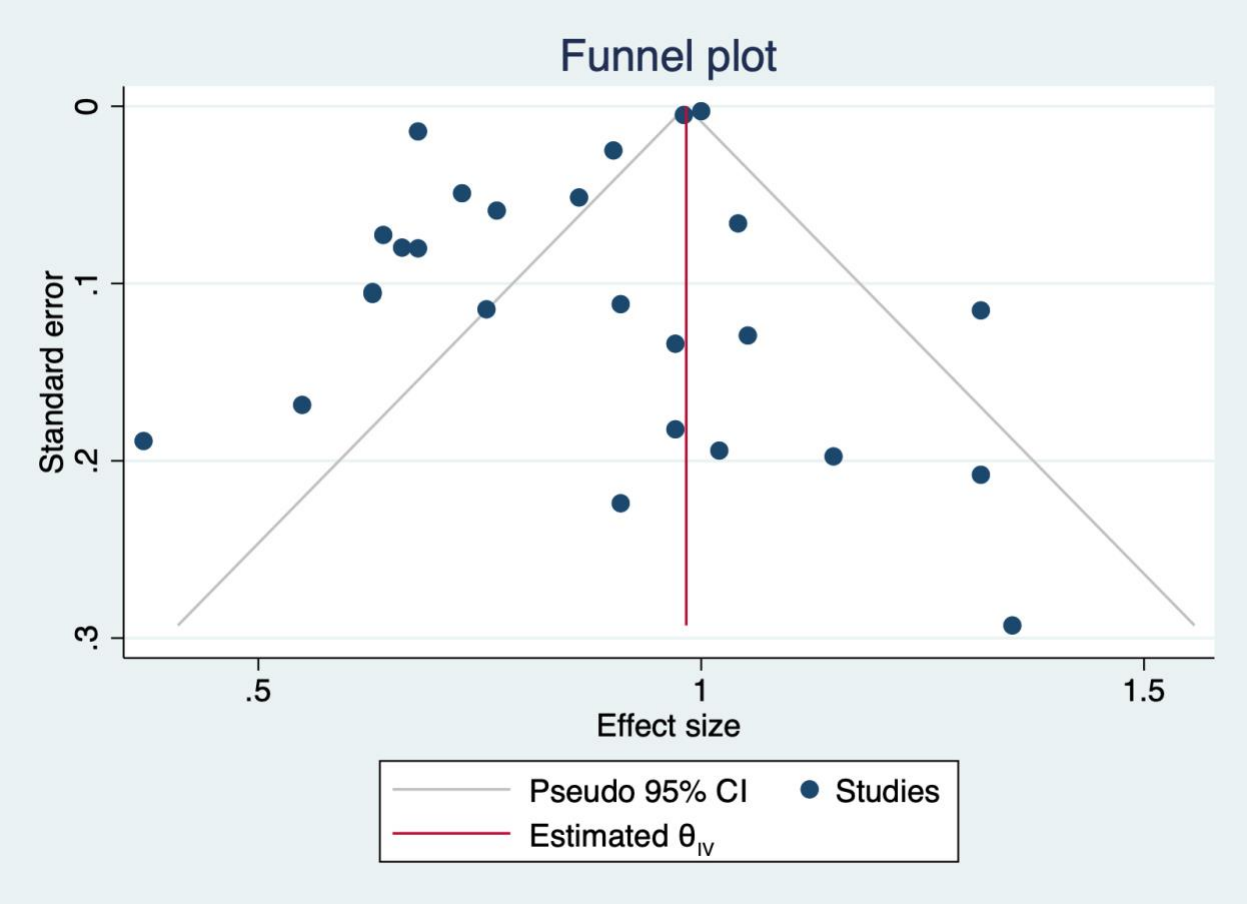


Figure 13 Funnel plot.



4.4 Discussion

This study found that higher volume was associated with a lower mortality, however, the thresholds used to define low and high volume varied across studies. Studies done outside the US, those using a clinical database, using risk adjustment, and those including academic affiliation were less likely to identify a volume-outcome relationship.

I conducted a systematic review to include only patients admitted to general ICUs, where the hospital capabilities between high and low volume ICUs were similar and where the outcome was not influenced by access to specialist services outside the ICU. We excluded the studies where transplant was only available in high volume centres and was material to survival. Similarly, we excluded studies where access to PCI or cardiac surgery was limited to high volume centres. I also excluded comparisons of volumes between specialist hospitals such as level 1 trauma centres as these were not generalisable.

In summary, critically ill patients appear to benefit from being treated in high volume ICUs, however, the definition of volume is inconsistent. Most of the included studies had significant risk of bias. The volume-outcome relationship may be accounted for by using clinical datasets with detailed risk adjustment, controlling for organisational factors such as academic affiliation, and using methods to account for the clustering of patients.

Chapter five - Paper one

Association of Intensive Care Unit Annual Sepsis Caseload with Patient Mortality from Sepsis in the United Kingdom, 2010-2016

Key points

Question: Do patients with sepsis have a lower chance of dying if they are treated in an intensive care unit (ICU) with a high sepsis volume?

Findings: In this study of 273,001 patients with sepsis from 231 ICUs in the UK, a higher annual sepsis caseload volume was associated with a significantly lower in-hospital mortality. This association showed no significant interaction with illness severity. A lower volume threshold of 215 treated patients was identified above which acute hospital mortality decreased significantly. Yet, 39% of patients were treated in ICUs below this threshold volume.

Meaning: This study demonstrates that sepsis patients in the UK have a higher chance of survival if treated in larger ICUs.

Abstract

Importance: Sepsis has a high burden of inpatient mortality. Intensive care units (ICUs) that have more experience treating sepsis patients are associated with lower mortality.

Objective: To assess the relationship between volume of sepsis patients receiving care in UK ICUs and mortality.

Design: Retrospective repeated cross-section analysis.

Setting: Adult patients with sepsis from 231 UK ICUs between 2010 and 2016

Participants: A total of 273001 adult patients with sepsis drawn from a clinical database. Demographic and clinical data was extracted from the ICNARC Case-mix Program Database.

Exposures: Patients with sepsis admitted to the ICU

Main outcome measures: Hospital mortality after ICU admission

Results: The mean age of patients was 63.3 years (95% CI 63.2-63.4) and 54.9% of the cohort were male. The mean ICNARC score was 20.6 (95%CI 20.6-20.7). Septic shock accounted for 19.3% of patients and 54.3% of patients required mechanical ventilation. The median annual sepsis volume was 242 (IQR [177-334]. The study identified a significant volume-outcome relationship for sepsis. The logistic regression model found a statistically significant reduction in hospital mortality for patients admitted to ICUs in the highest quartile of sepsis volume compared to the lowest quartile (OR 0.89, 95% CI 0.82-0.96, p=0.002). With volume modelled as a restricted cubic spline, larger ICUs were associated with lower hospital mortality (p=0.001). We identified a lower volume threshold of 215 patients above which acute hospital mortality decreased significantly and note that 39% of patients were treated in ICUs operating below this threshold volume. The volume-outcome relationship was not affected when allowing for illness severity.

Conclusion and relevance: This study demonstrates that sepsis patients in the UK have a higher chance of survival if treated in larger ICUs. The benefits of high sepsis volume were not related to the severity of the sepsis episode.

5.1. Introduction

Sepsis is a dysregulated host response to infection that results in organ dysfunction[94]. It is amongst the leading causes of death worldwide and the global burden of sepsis is expected to increase as populations age [93]. The World Health Assembly has urged member states and other stakeholders to strengthen efforts to prevent, diagnose and treat sepsis[155]. Patients with sepsis require high-cost interventions in intensive care units (ICUs) where, even with prompt treatment, they face a high probability of death[93]. One strategy to reduce mortality might be to treat sepsis patients in larger, high-volume ICUs. This paper aims to provide evidence of this contention using data from three countries of the United Kingdom.

Since the seminal report by Luft et al. in 1979, there has been growing evidence that patients receiving treatment for complex conditions have lower mortality when treated in institutions with a high-volume caseload compared with institutions with low volume [9, 156-162]. Other major benefits beyond outcome are the potential for lower costs through economies of scale and more efficient use of staff and other resources. The major concerns are the potential for fragmentation of care, the need to transport patients away from their local hospital and the possibility of overwhelming high-volume centres. Prior studies of the volume-outcome relationship for patients with sepsis have shown conflicting results [163-166]. This literature is subject to the following limitations, which we address fully in this study[164, 165, 167-169].

First, many volume-outcome studies in sepsis patients were undertaken in the US where there is a complex system of health care funding and the observed benefits attributed to volume may to some extent reflect unmeasured disparities in access to care as well as socioeconomic disparities [149, 170, 171] [5, 172, 173]. Studies undertaken in countries such as Canada, Finland or the UK with single-payer, publicly funded healthcare systems, have not

demonstrated a consistent volume-outcome relationship[165, 169, 174]. Second, comparisons between high- and low-volume specialist and non-specialist services do not disentangle the effects of specialisation and volume. In our study, we consider a cohort of patients requiring emergency care in a publicly funded health system in general, non-specialist ICUs.

Third, a major limitation of the existing volume-outcome literature is the lack of a gold standard in defining volume. Examining quartiles does not improve our general understanding of the volume-outcome relationship as ICUs considered high-volume in one study may fall within a lower quartile in another because the quartiles are specific to each dataset. In our study we employ restricted cubic splines that allow flexibility in describing the functional form of volume in regression models. In using the full range of data, these methods provide a more accurate description of the relationship between volume and mortality, with the additional ability to identify optimal volume thresholds. Fourth, many studies include a relatively small number of ICUs with a narrow spectrum of volumes leaving them underpowered to detect a small, but statistically and clinically meaningful volume-outcome relationship.

Finally, most studies use secondary administrative data collected for other uses. Such data have inherent limitations in both the identification of sepsis and the characteristics of both patients and ICUs. Our study employed a large clinical database of patients with sepsis admitted to all general ICUs in the UK allowing us to perform detailed risk adjustment and identify ICU specific characteristics

We evaluated the volume-outcome relationship in sepsis patients for several reasons[94, 175]. Sepsis, while relatively common and clinically identifiable, has not attracted a great deal of attention in the volume-outcome literature. Sepsis requires time critical interventions provided almost exclusively within the ICU, allowing direct attribution of effects to ICU treatment. UK ICUs are unable to risk select low risk patients because sepsis is an emergency condition and patients are taken to the nearest hospital, often by the ambulance service. Our empirical findings suggest treatment benefits could be made through a concentration of ICU facilities, similar to the successful policy adopted by the National Health Service (NHS) in some areas with respect to the treatment of stroke[176]. We chose mortality as the outcome because sepsis is associated with significant mortality, and this outcome is not subject to gaming or manipulation.

5.2 Methods

5.2.1 Data

We analysed data from the Case-Mix Program Database (CMPD), a national clinical database of all adult patients admitted to ICUs in England, Wales, and Northern Ireland. The CPMD contains all adult admissions to all general ICU and is co-ordinated by the Intensive Care Audit & Research Network (ICNARC). Details of the validation of the CMPD have been previously published [120, 121, 177-179]. Approval for the Case Mix Programme was obtained under Section 251 of the NHS Act 2006 (approval number PIAG 2–10(f)/2005). The London School of Economics waived the requirement for informed consent because this research involves secondary analysis of an established dataset of anonymized data. We report this study as per Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines[180].

5.2.2 Patient selection

All adult admissions with sepsis to 231 general ICUs in England, Wales and Northern Ireland between 1 January 2010 and 31 December 2016 were included. The sepsis cohort of ICU admissions was identified using Sepsis-3 definitions [94]. We considered the index critical care admission for sepsis as infection with a Sequential Organ Failure (SOFA) score of 2 or greater. Septic shock was an infection with a cardiovascular component of the SOFA score of 2 or greater or a serum lactate concentration greater than 18mg/dL as per the Sepsis-3 consensus definition[94] Patients less than 16 years old and patients for whom all physiologic data were missing and patients that stayed in the ICU for less than 8 hours were excluded.

5.2.3 Exposure

The exposure was defined as the annual caseload volume of the year of ICU admission, so for example, if a patient was admitted to the ICU in 2010, we used the sepsis volume of 2010 for that ICU. In the initial analysis we follow the common approach of categorising ICU volumes into quartiles which might be justified here given that we are analysing the complete set of general ICUs in England, Wales, and Northern Ireland. Our preferred analyses specified volume as a continuous variable and uses restricted cubic splines to identify the best fitting model.

5.2.4 Study outcome

The primary outcome was death before discharge from the acute hospital. Patients that were transferred between ICUs were excluded from the analysis of mortality but included in the

estimation of ICU volumes. This was done to avoid confounding results with outcomes from differing ICUs. For patients that were readmitted to the ICU, only the first admission was included in the mortality analysis.

5.2.5. Analysis

The risk adjusted association between ICU volume and acute hospital mortality was assessed using a mixed-effects logistic model in a three-level hierarchical structure based on individual patients nested in years nested within ICUs. This mixed-effects approach was adopted to clearly identify the ICU volume-outcome effect, while giving adequate control for the within ICU variation over time. Control variables included in the model were age, gender (base category = female), prehospital dependence (base category = no dependence), race (base category = white), comorbidities (categories: severe respiratory disease, severe cardiac disease, end-stage renal disease, severe liver disease, metastatic disease, haematological malignancy and immunocompromised), socioeconomic deprivation as measured by the Index of Multiple Deprivation, severity of illness as measured by the $ICNARC_{2018}$ score, annual bed occupancy rate and academic affiliation (base category = non-university). Full details are provided in the Appendix.

Annual ICU sepsis volume was initially analysed as a categorical variable as with earlier studies. Categorisation is a popular method for volume-outcome studies but has disadvantages. The categories are determined by the distribution of the data, so the cut-off points are arbitrary and study-specific, limiting generalisation. There is also significant loss of information through categorisation, with all ICUs in the same category assumed to have the same mortality risk.

Our subsequent analysis therefore defines volume as a continuous variable, and we specify restricted cubic splines to allow for the non-linear relationship between volume and mortality. In making the model more flexible, potential overfitting is avoided, while the interpretability of the modelled relationship is retained. Restricted cubic splines can identify local features and provide stable estimates at the tails of data, making the spline model reliable in identifying a local marginal treatment effect. We fitted models with 3, 4, 5 and 6 knots and used information criteria and likelihood ratio tests to select the model with 3 knots, this being the most parsimonious [181]. We use a Wald test to assess the overall association between volume and mortality and present the results graphically with confidence intervals. We specify values of ICU volume at midpoints on the knots to provide a comparison with the quartile model[182]. Details, including various specification tests are included in the Appendix. Significance was defined as $P<0.05$. Data analysis was performed using Stata (version 16.0, StataCorp LP).

We used the three-level hierarchical logistic regression model to account for the clustering of patients within ICUs across years. This approach also estimates random intercepts for each ICU which are interpreted as the latent ICU-level variation[183]. A more detailed discussion is included in the appendix. We evaluated the significance of the between-ICU variation using a median odds ratio (MOR) [184].

5.2.5.1 Sub-group analysis

We hypothesised that sicker patients face lower mortality risk if treated in a high-volume ICU. To test this, we performed an interaction test of volume and illness severity using the mortality risk predicted by the $ICNARC_{2018}$ score. We performed sensitivity within this sub-group analysis by altering the definition of more severely ill. We subsequently defined the sicker admissions as those with septic shock, with an expected mortality >30% predicted by the ICNARC model or receiving mechanical ventilation or renal replacement therapy. Second, we analysed non-surgical patients with sepsis to ensure that the observed outcome was not driven by surgical patients with sepsis.

5.2.5.2 Sensitivity analysis

We used fractional polynomials as an alternate specification of volume as a continuous variable to test the sensitivity of our results to the specification on the volume-outcome relationship [185]. Fractional polynomials are global functions and may obscure local features, particularly at the tails of the data distribution and may therefore be less useful than cubic splines in identifying a threshold volume, particularly at low volumes [186].

We then performed a quantitative bias assessment to assess the influence of unmeasured covariates [187, 188]. We did so using E-values, which measure the minimum association that an unmeasured covariate would require with both ICU volume and mortality, conditional on the measured covariates, to fully explain the empirically determined volume-outcome relationship[188].

We check that volume is exogenous. In our model exogeneity would require that ICU volume is uncorrelated with the ICU-level random effect [189]. A more detailed discussion is provided in the Appendix.

Lastly, we perform a sensitivity analysis to check if the results remain consistent after removing the extremes of volumes. We excluded the lowest annual volumes (<20 sepsis patients per year and the highest (>740 sepsis patients per year). The results are included in the appendix.

5.3 Results

5.3.1 Descriptive statistics

Of the 305,748 ICU admissions that met the Sepsis-3 criteria between 2010 and 2016, 32,747 (10.7%) were excluded from the mortality analysis. This included 19,809 readmissions, 12,296 patients transferred between ICUs and 642 patients that were both readmissions and inter-ICU transfers. Descriptive statistics for the sample of 273,001 patients with sepsis treated within general ICUs from the period 2010 to 2016 are presented in **Error! Reference source not found.**, eTable 1 and patient flow in eFigure 1 in the appendix for paper one. Median age for the patient population was 66 years (IQR 53-76 years). Most patients (80.1%) had no significant medical co-morbidity. There were 1.8% patients recorded as having severe cardiac disease, 4.6% with severe respiratory disease, 1.9% with end stage kidney disease, 2.2% with liver disease and 8.8% were immunocompromised. The mean $ICNARC_{2018}$ predicted mortality was 29.7% (95% CI 29.6% to 29.8%). Mechanical ventilation was used in 54.3% of patients,

19.3% had a diagnosis of septic shock, and 8.6% had received renal replacement therapy within 24 hours of admission.

The unadjusted hospital mortality rate was 31.9% (95% CI 31.8% to 32.1%). Hospital mortality was 33.3% in the lowest quartile compared with 30.7% in the highest quartile (Table 8)

Table 8.Characteristics of patients admitted to the ICU with between 2010 and 2016 and across quartiles of annual ICU case load of sepsis.

Variable	**Total Sepsis**	**Quartile I**	**Quartile II**	**Quartile III**	**Quartile IV**	p-value
		12-177 cases	178-242 cases	243-334 cases	335-744 cases	
	N=273001	N=68952	N=69269	N=68289	N=66491	
Age in years						<0.001
<54	68947(25.2)	17022(24.9)	16650(24.4)	17453(25.5)	17232(25.2)	
54-66	69264(25.3)	17322(25.3)	17021(24.9)	17110(25.0)	17007(24.8)	
67-76	68289(25.0)	18011(25.4)	18256(25.8)	17674(25.0)	16894(23.9)	
>76	66491(24.4)	16592(25.4)	17337(26.5)	16052(24.6)	15358(23.5)	
Male	148149(54.2)	37226(54.0)	37326(53.9)	37280(54.6)	36317(54.6)	0.006
Female	124852(45.7)	31726(46.0)	31943(46.1)	31009(45.4)	30174(45.4)	
Ethnicity						<0.001
White	248275(91.0)	63059(91.5)	64504(93.2)	62712(91.9)	58000(87.2)	
Asian	9438(3.5)	2472(3.6)	1779(2.6)	2114(3.1)	3073(4.6)	
Black	5504(2.0)	1304(1.9)	1092(1.6)	1036(1.5)	2072(3.1)	
Mixed/other	9617(3.5)	2070(3.0)	1848(2.7)	2353(3.4)	3346(5.0)	
Comorbidities						
Cardiac	4857(1.8)	1390(2.0)	1032(1.5)	11318(1.9)	1117(1.7)	<0.001
Respiratory	12498(4.6)	3187(4.6)	2870(4.2)	2863(4.2)	3578(5.4)	<0.001
ESRD	5171(1.9)	1002(1.5)	967(1.4)	1297(1.9)	1905((2.9)	<0.001
Liver	6030(2.2)	1208(1.8)	1285(1.9)	1468((2.2)	2069(3.1)	<0.001
Metastatic cancer	6598(2.4)	1610(2.4)	1509(2.2)	1709(2.5)	1770(2.7)	<0.001
Haematological malignancy	9763(3.6)	2349(3.4)	2178(3.2)	2551(3.8)	2685(4.1)	<0.001
Immunocompromised	24035(8.8)	5884(8.6)	5553(8.1)	6287(9.3)	6311(9.5)	<0.001
Level of dependency prior to acute hospitalization						<0.001
Independent	184850(68.0)	47150(68.7)	47545(68.9)	44925(66.1)	45230(68.3)	
Some assistance	81913(30.1)	20220(29.5)	20233(29.3)	21851(32.1)	19609(29.6)	
Total dependence	5071(1.9)	1262(1.8)	1223(1.8)	1214(1.8)	1372(2.1)	

Usual residence prior to hospitalization						<0.001
Home	264730(97.0)	66816(96.9)	67200(97.0)	66286(97.1)	64428(96.9)	
Work or non-health related institution	564(0.2)	144(0.2)	132(0.2)	143(0.2)	145(0.2)	
Nursing home, hospice, or health related institution	6756(2.5)	1781(2.6)	1716(2.5)	1646(2.4)	1613(2.4)	
No fixed address	951(0.4)	211(0.3)	221(0.3)	214(0.3)	305(0.5)	
IMD quintile						<0.001
I	69728(25.7)	15507(22.7)	15654(22.7)	17144(25.3)	21423(32.5)	
II	58496(21.6)	15047(22.0)	15574(22.6)	13632(20.1)	14243(21.6)	
III	53199(19.6)	14075(20.6)	14345(20.8)	13200(19.4)	11579(17.5)	
IV	47306(17.5)	12864(18.8)	12623(18.3)	12095(17.8)	9724(14.7)	
V	42400(15.6)	107776(15.8)	10743(15.6)	111833(17.4)	9048(13.7)	
Admission type						<0.001
-Medical	204524(74.9)	52890(76.7)	51067(73.7)	50163(73.4)	50404(74.9)	
-Elective surgery	11780(4.3)	3167(4.6)	2825(4.1)	2710(4.0)	3078(4.6)	
-Emergency surgery	56671(20.8)	12886(18.7)	15368(22.2)	15409(22.6)	13008(19.6)	
APACHE II score, mean(95% CI)	18.4(18.4-18.4)	18.5(18.4-18.5)	18.3(18.2-18.3)	18.5(18.4-18.5)	18.5(18.5-18.6)	<0.001
ICNARC score, mean(95% CI)	21.0(20.9-21.0)	21.3(21.2-21.4)	21.1(21.0-21.1)	20.9(20.9-21.0)	20.4(20.4-20.5)	<0.001
ICNARC predicted probability of death, mean% (95% CI)	29.7(29.6-29.8)	30.7(30.5-30.9)	29.8(29.6-30.0)	29.5(29.3-29.7)	28.8(28.6-29.0)	<0.001
Renal failure in the first 24 hours	23573 (8.8)	6154(9.1)	6253(9.2)	5866(8.7)	5300(8.1)	<0.001
Mechanical ventilation	145041 (53.1)	38278 (55.5)	36994 (53.4)	36035 (52.8)	33734 (50.7)	<0.001
Septic shock	54419 (19.9)	14458(21.0)	13912 (20.1)	13016 (19.1)	13033 (19.6)	<0.001
ICU length of stay in hours, median IQR	90(42-189)	93(41-200)	88(41-186)	90(42-187)	88(42-186)	<0.001
Hospital length of stay in days, median IQR	14(7-28)	14(7-29)	14(7-27)	14(7-28)	15(7-30)	<0.001
ICU mortality	62277(22.8)	16156(23.4)	16245(23.5)	15567(22.8)	14309(21.5)	<0.001
Hospital mortality	86728 (31.9)	22789(33.3)	22381(32.5)	21263(31.3)	20295(30.7)	<0.001

For categorical variables a Chi squared test was used. The null hypothesis was that there is no difference in the distribution of responses to the outcome across comparison groups. For continuous variables we use the ANOVA to analyse the differences in means between

Of the 231 ICUs, 122 (52.8%) were in non-university hospitals, 39(16.9%) were university affiliated and 70 (30.3%) were university based. The median number of ICU beds was 8[IQR 6-10] in the lowest quartile of ICU volume compared with 23 [IQR 18-28] ICU beds in the highest quartile (Table 9, eTable 2 and eFigure 2).

Table 9. Characteristics of 231 Intensive care units across quartiles of annual sepsis volume

Variable	Total	Quartile I	Quartile II	Quartile III	Quartile IV	p-value
ICU beds, median (IQR)	13(9-18)	8(6-10)	11(9-13)	15(12-17)	23(18-28)	<0.001
Occupancy %, median (IQR)	73.5(67.9-79.5)	67.5(59.4-74.2)	71.7(66.9-77.0)	74.0(70.5-79.6)	78.6(74.8-82.9)	<0.001
Sepsis volume, median (IQR)	242(177-334)	136(112-160)	214(197-228)	280((260-302)	415(378-483)	<0.001
Non-sepsis volume, median (IQR)	497(346-747)	288(220-369)	432(343-521)	572(461-706)	918(713-1176)	<0.001
Total volume, median (IQR)	742(533-1087)	427(353-516)	646(552-737)	856(732-997)	1348(1173-1614)	<0.001

IQR= interquartile range

For categorical variables a Chi squared test was used. The null hypothesis was that there is no difference in the distribution of responses to the outcome across comparison groups. For continuous variables we use the ANOVA to analyse the differences in means between groups.

5.3.2. Regression Analysis

The logistic regression model found a statistically significant reduction in hospital mortality for patients admitted to ICUs in the highest quartile of sepsis volume compared to the lowest quartile (OR 0.89, 95% CI 0.82-0.96, P=0.002) (Table 10) (Figure 14)(eFigure 3). This would

mean that patients with sepsis would have 11% lower odds for mortality when treated in the highest quartile of volume compared with the lowest quartiles of sepsis volume. To reduce instability at the extremes restricted cubic splines are constrained to be linear before the first knot and after the last knot. Interpreting the effect of volume on mortality when volume is specified as a restricted cubic spline is challenging because effect changes as the volume changes. Table 10 describes the adjusted odds ratio at the midpoint of each spline as a representation of the average effect within the spline.

Table 10 Odds ratio acute hospital mortality specifying ICU sepsis volume as categorial and a restricted cubic spline.

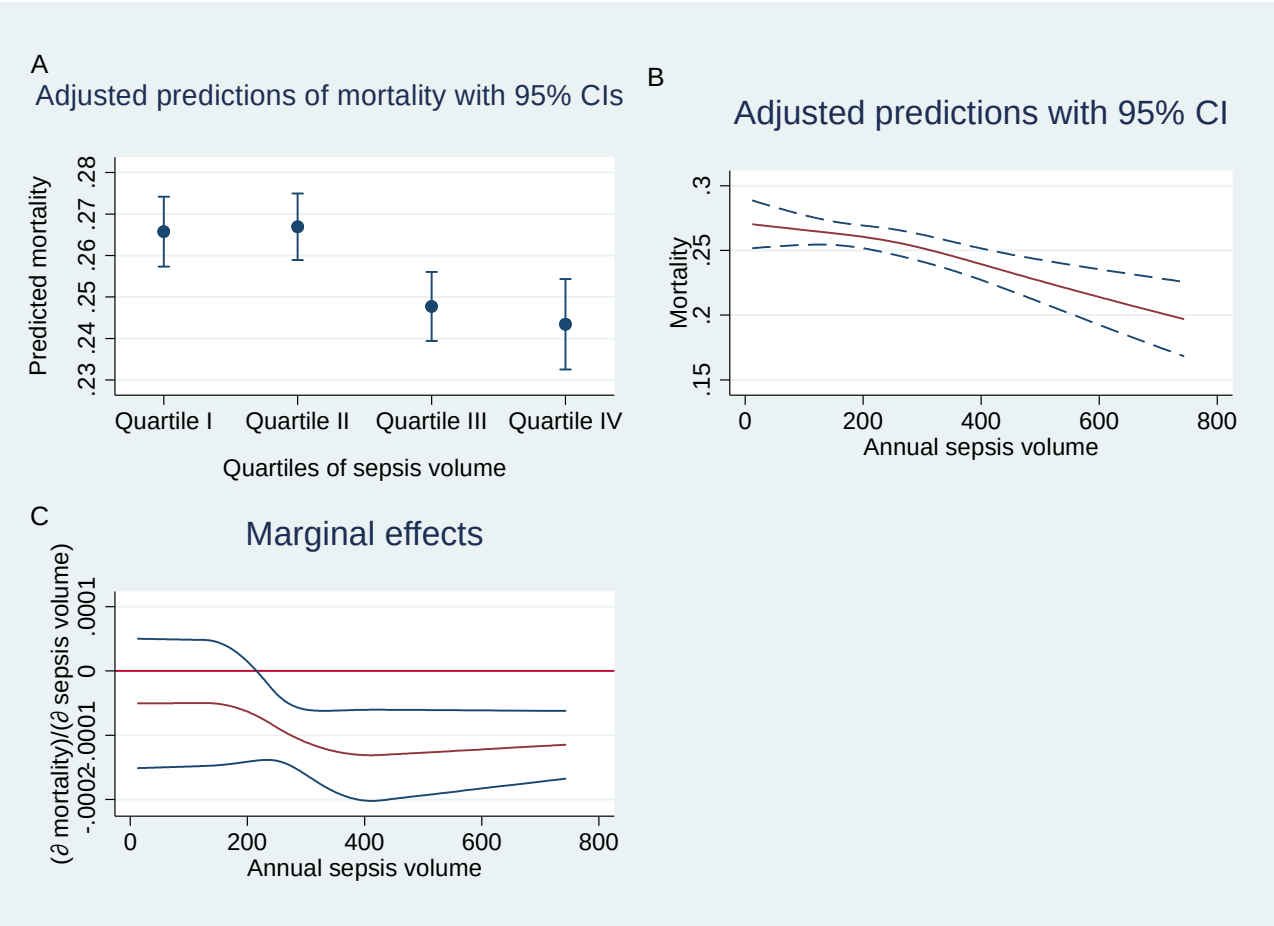
Model	**OR**	**95% CI**	**P value**
Categorical			
Quartile I (12-177cases)	1.0	Reference	
Quartile II (178-242 cases)	1.01	0.96-1.05	0.799
Quartile III (243-334 cases)	0.91	0.86-0.96	0.001
Quartile IV (335-744 cases)	0.89	0.82-0.96	0.002
*Restricted cubic Splines **			
Midpoint origin and knot 1(63 patients)	1.0	Reference	
Midpoint knot 1 and knot 2 (184 patients)	0.97	0.91-1.03	
Midpoint knot 2 and knot 3 (335 patients)	0.90	0.82-0.99	
Midpoint knot 3 and maximum (589 patients)	0.75	0.66-0.86	

*Per 50 patients with sepsis

With volume modelled as a restricted cubic spline, larger ICUs were associated with lower hospital mortality(P<0.001) (Panel B in Figure 14 and eTable 3). The adjusted prediction shows the predicted mortality against volume fitted as a restricted cubic spline with the average values for other covariates (Panel B in Figure 14). The marginal effect refers to the predicted change in mortality per unit change in ICU volume and in non-linear models varies with the point of estimation. The coefficient is the first derivative and the point at which it crosses the 95% confidence interval crosses the zero line is the threshold value. The restricted cubic spline specification identified a lower threshold of about 215 patients above which there was a favourable volume-outcome relationship (Figure 14). Above this volume threshold there was a significant reduction in mortality. About 39 % of patients with sepsis were treated in ICUs below this threshold value and 72% of ICUs operate below this threshold value. We could not identify an upper threshold value.

Figure 14. Relationship between ICU sepsis volume and acute hospital mortality. Data are adjusted probabilities and 95% confidence intervals.

(A) Sepsis volume as quartiles Predicted mortality: Quartile I (12-177cases) 26.5% (95%CI 25.7% to 27.4%); Quartile II (178-242 cases) 26.7% (95%CI 25.9% to 27.5%); Quartile III (243-334 cases) 24.8% (95%CI 23.9% to 25.6%); Quartile IV (335-744 cases) 24.3% (95%CI 23.2% to 25.4%) (B) Sepsis volume as a restricted cubic spline with 3 knots showing predicted mortality with 95% confidence interval. Knot 1 is at 127 patients, knot 2 is at 242 patients and knot 3 is at 428 patients. (C) Marginal effects of sepsis volume showing predicted change in mortality for a change in sepsis volume using a 3-knot restricted cubic spline. This analysis identifies a threshold of 215 patients above which there is a favourable volume outcome relationship.



A, Markers represent adjusted probabilities and whiskers indicate 95% CIs. B, Shaded area indicates 95% CI. C, Shaded area indicates 95% CI; dashed vertical line indicates the threshold at which an increase in volume resulted in a significant reduction in estimated mortality.

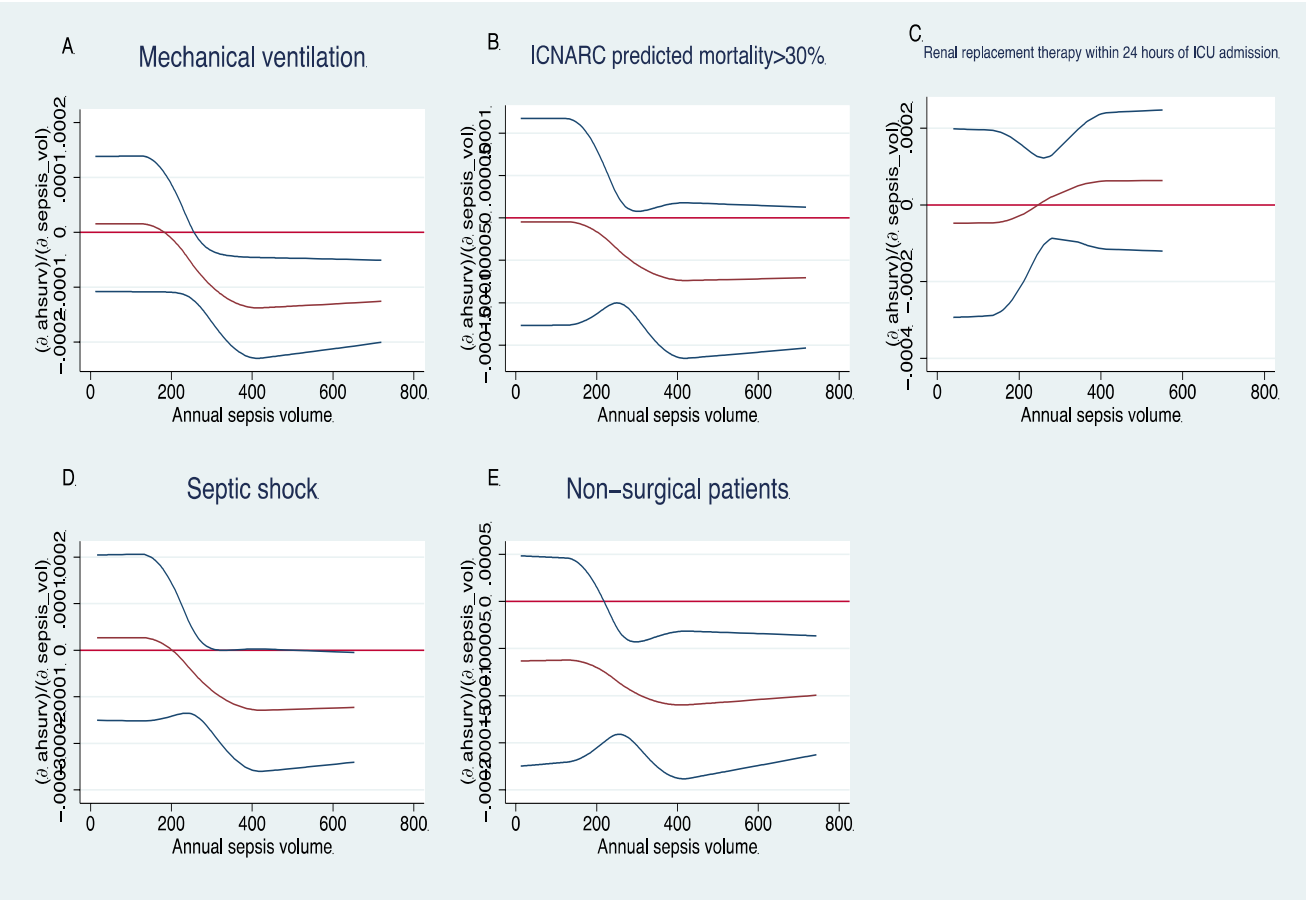
The between-ICU practice variation is derived from the mixed-effects model using estimated intercepts as a measure of latent quality. The MOR for hospital mortality was 1.27 (95% CI 1.23-1.30), implying significant unexplained between-ICU practice variation. The variance within the same ICU across the study period did not change significantly implying that an individual's ICU's performance in terms of mortality was stable over time (eFigure 4).

5.3.2.1 Subgroup analyses

There was no significant interaction between volume and severity of illness as described by the $ICNARC_{2018}$ score (P=0.570). Additionally, subgroup analyses of those patients defined as severely ill through receiving mechanical ventilation, having a predicted mortality greater than 30%, receiving renal replacement therapy within 24 hours of ICU admission and patients with septic shock also did not identify an enhanced volume-outcome relationship (**Error! Reference source not found.**). The subgroup of non-surgical sepsis patients demonstrated a similar volume-outcome relationship to the entire cohort (Figure 15 and eFigure 5).

Figure 15. Subgroup analysis of marginal effects of ICU volume

(A)patients receiving. mechanical ventilation, (B) ICNARC predicted mortality>30%, (C) patients receiving renal replacement therapy within 24 hours of ICU admission, (D) septic shock and (E) Non-surgical patients.



Shaded areas indicate 95% CIs. ICNARC, Intensive Care National Audit & Research Centre.

5.3.2.2 Sensitivity analysis

The inverse relationship between volume and mortality remained statistically significant for the fractional polynomial model (P<0.001) (eFigure 6, eTable 4, and eTable 5). The quantitative bias analysis returned a E-value of 1.31 and the lower confidence limit is 1.17 (eFigure 7 and Appendix). The lack of statistical significance in the between- and within-cluster effects for ICU volume imply a lack of correlation in the ICU volume and the ICU random effect, in support of the assumption that ICU sepsis volume is exogenous (eTable 6). The analysis that excluded the extremes of sepsis volume identified a similar threshold value for the volume-outcome relationship. Details are provided in the appendix.

5.4 Discussion

Our study has several strengths. In terms of completeness, coverage, and representativeness of the data, this is one of the largest studies of a volume-outcome relationship. By including all general ICUs in England, Wales, and Northern Ireland, it covers the entire adult sepsis population treated in these countries over the study period [120]. This study also used a granular clinical database with a standardised data collection process, a highly validated risk adjustment model developed for UK ICUs and used the international consensus Sepsis-3 definition to identify patients with sepsis [120, 123] [94, 149].

Our results, which return a strong volume-outcome relationship, are robust with the positive association being consistent across the categorical and non-linear specifications of ICU volume. The sepsis volumes included in this study exceeds the spectrum of volumes described in other published studies, thereby improving the power to detect even a small volume-

outcome relationship [149, 165]. The potential for selection bias is limited by using a cohort of patients with sepsis treated in publicly funded general ICUs within the UK NHS which covers the whole population.

The study also identified a lower volume threshold of 215 patients above which there was a statistically significant reduction in mortality. This threshold is estimated based on our preferred empirical specification using a 3-knot restricted cubic spline regression which also controlled for a rich set of covariates to model the volume-outcome relationship. There was no significant interaction between volume and severity of illness. The study found significant ICU practice variation not explained by patient or hospital characteristics, implying that sample selection was not distorting the volume-outcome relationships described. The within-ICU variation remained unchanged across years, suggesting that higher performing ICUs maintain their good performance over time.

Our findings are based on a large population of ICUs observed over time. A recent meta-analysis of smaller observational studies found an overall positive effect of volume, however there was significant heterogeneity [163]. Many previous studies that did not account for the clustered nature of the data have produced upwardly biased estimates of the volume-outcome relationship[190, 191]. The hierarchical structure of our analysis may account for the more modest effect seen in our study compared with other publications [163] .

5.4.1. Limitations

This study nevertheless has weaknesses. The study uses observational data that may be subject to unmeasured confounding. We evaluated the potential for unmeasured

confounding using E-values. In our study this was a threshold risk ratio of 1.17. While the E-value is modest, we think that, given the detailed clinical data recorded in the CMPD, significant unmeasured confounding is improbable. If an omitted variable is correlated with an included covariate, then the omitted variable does not result in much bias. E-value assumes the distribution of unmeasured confounders is as unfavourable as possible and represents the most conservative scenario.

As is typical of this literature, we use the contemporaneous volume as the exposure. This does not distinguish between the static scale effects of volume and the cumulative learning-by-doing effects. Additionally, the dataset does not have details on processes of care specific to sepsis such as timing of first dose of antibiotics. We are therefore unable to clearly establish the underlying mechanism of the volume-outcome relationship for sepsis.

5.5 Conclusion

This study identified a significant volume-outcome relationship for sepsis mortality and a lower volume threshold associated with this improvement in mortality. Further research is required to better understand the mechanism through which the volume-outcome relationship operates for this group of patients.

Acknowledgements

Role of funder

The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Chapter six: Learning by doing or economies of scale in sepsis: does experience influence performance?

Abstract

Objective: This study aims to distinguish between the two mechanisms of the volume-outcome relationship, namely learning-by-doing and scale economies measured by lagged and contemporaneous volume.

Data sources: The ICNARC Case-Mix program database was used to identify patients with sepsis admitted to the intensive care unit in the United Kingdom between 2010 and 2016. Patients with sepsis were identified by the Sepsis-3 consensus criteria.

Study design: The patient was the primary unit of analysis. We fitted a multilevel logistic regression model of patients nested in intensive care units over quarters to assess the effects of static scale effects and the learning-by-doing effect on acute hospital mortality. Patient and ICU characteristics were included for risk adjustment.

Principal findings: Our study identified a cohort of 273,001 patients with sepsis admitted to 231 ICUs in the UK. Our study finds that in comparison with contemporaneous volume, lagged volume had a stronger association with acute hospital mortality. This finding was consistent across alternate specifications of learning-by-doing.

Conclusion

The mechanism for the volume-outcome relationship is learning-by doing and not the static economies of scale.

6.1 Introduction

The volume-outcome relationship has been a commonly invoked policy initiative aimed at improving the quality of healthcare. This inverse relationship between the caseload volume of patients treated and patient mortality has been described across many health settings and in many countries[11, 192, 193]. Despite the large body of literature demonstrating this favourable relationship, most studies have focused on differentiating the effects of selective referral and the true effects of volume. In comparison there are few studies evaluating the underlying mechanism of the volume-outcome relationship[193, 194]. Consequently, policies that encourage centralisation or minimum volume standards have never firmly been established[135, 195-197]. Whilst economists have an interest in improving scale efficiency, a number of countries have implemented pro-competition reforms with the view that increasing the number of providers, which potentially reduces volumes, would lead to better health outcomes [198]. Resolving this tension between policies of centralisation of services and allowing providers to accrue experience over time requires an understanding of any underlying mechanism of the volume-outcome relationship[199, 200].

The term learning-by-doing was introduced by Arrow in 1962 to refer to institutional learning and describes the improvements in outcomes by experience[1]. The simple idea is that the team acquires experience by performing tasks repeatedly and that care improves because of the cumulative effect of skills gained through patients treated in the past. The knowledge is gained by a production of experience as activity increases, and has been termed a learning-curve [1]. Learning-by-doing would result in increases in quality attributed to increases in knowledge. The acquisition of knowledge is termed learning and is a product of experience[1].

Even when volume remains unchanged between years, improvements in quality attributable to experience still occur. Learning takes place through activity and can be observed as improving performance so that today's volume improves tomorrow's outcomes. One policy implication is to that you need to keep a team together.

The alternate mechanism of the volume outcome relationship is the static scale effect. Here economies of scale are driven by indivisibilities of critical investments which impact patients' outcomes and reflect contemporaneous gains provided through consolidation or centralisation. In the economies of scale mechanism, the caseload volume of today affects the contemporaneous outcomes.

From a policy perspective, the underlying mechanism by which volume leads to positive patient outcomes matters. If the static volume is the underlying mechanism, then any ICU in which volume is concentrated will lead to better outcomes. In the learning by doing mechanism, shifting volume from one ICU to another would reduce opportunities for learning in the transferring ICU which is losing volume. The learning-by-doing mechanism makes ICU consolidation less attractive than the static scale mechanism.

There is evidence that supports both contemporaneous static economies of scale as well as learning by doing cumulative volume and the predominant effect may be related to the type of patient cohort studied[1, 15]. The literature thus far has focused on coronary revascularisation and elective surgery and there is no specific work on critically ill patients. Gaynor et al. examined the effects of scale economies and learning-by-doing in cardiac surgery concluding that the benefits of volume were primary due to static scale

economies[199]. A major limitation of this work was that the year lag structure used to detect a learning by doing effect meant that learning over shorter time periods was not observable. A study of coronary angioplasty found the contemporaneous annual hospital procedure volume to be associated with improved outcomes and found no evidence of learning-by doing[201]. This study specified learning-by-doing as the cumulative volume since 1984 and before the current year.[202]. The study did not describe a specific rationale for choosing 1984 as the starting point. A major critique of this study was that the learning curve was poorly identified by cumulative volume as the innate level of knowledge and experience between hospitals in 1984 was unknown as 73% of hospitals were already performing angioplasty by 1984, 5% of hospitals started after 1990 and a further 6% discontinued after 1994[202]. Similarly, a study of cardiac procedures in the US found no learning effect and that all the observed benefits of volume were attributable to static scale[19]. Survival after cardiac procedures is high, making it difficult to demonstrate a learning effect when using mortality as an outcome. In contrast, a study of advanced cancer surgery in Sweden found learning-by-doing to be significant[203]. This study found a larger effect from lagged volume than from contemporaneous volume [203]. The study found a larger benefit in patients with colon cancer compared with breast cancer and concluded that this might represent an additional benefit to patients with higher complexity of disease.

A systematic review of the underlying mechanism of the volume-outcome relationship found that most studies did not explore the underlying learning mechanism[194]. The current literature on the mechanism of the volume-outcome relationship has several limitations. First, almost all the evidence of the mechanism of the volume outcome relationship is derived from elective surgical populations and predominantly cardiac and cancer related surgery. In

these populations, there may be selective referral to higher performing centre. In contrast to elective surgery, patients with sepsis are usually taken to their nearest ICU removing selection related to ICU quality or severity of illness. Second, all of the previous studies use a fixed effects approach to account for time invariant institution level unobserved heterogeneity[18]. The inclusion of an institutional fixed-effect means that the regression estimates the effects of changes to volume within the hospital rather than the effects of changes to cumulative volume across hospitals on mortality [18]. The fixed-effects specification requires an adequate number of institutions with significant variation in volume to detect a learning-by-doing effect. Previous studies that have failed to show a learning-by doing effect may therefore not have been suitably powered[18].

Sepsis is the syndrome of life-threatening organ dysfunction that occurs in response to infection[94]. Sepsis is a major public health concern and is the leading cause of morbidity and mortality globally[93]. The recent Global Burden of Disease Study estimates 48.9 million incident cases of sepsis causing 11 million deaths worldwide, representing 20% of all global deaths[93]. The reported incidence of sepsis is increasing, reflecting ageing populations with more co-morbidities, and increasing social deprivation[93]. The Seventeenth World Health Assembly recognised the importance of strong, functional health systems, including access to intensive care services and health system organizational strategies to improve outcomes from sepsis[204]. Centralisation of care for sepsis has often been proposed to improve patient outcomes by capitalising on the volume-outcome relationship[196, 205]. Centralisation is predicated on the assumption that the volume-outcome relationship operates through the static scale effect. If the volume-outcome relationship operates through the learning-by-doing mechanism, then patient outcomes would improve by the volume of patients treated over time, making system-wide centralisation unnecessary. The unsettled question of the

underlying learning mechanism in the volume-outcome relationship therefore has clear implications for all stakeholders in the health system.

We make several contributions to the literature on the volume-outcome relationship. First, this study measures volume at a higher frequency than previously done (quarter instead of year), the quarterly time lag being more likely to detect learning than year lags given the temporal instability of ICU teams. Second, our study includes a large number of ICUs over several years and employs a mixed-effects logistics regression model. This approach is more sensitive to detecting a learning-by-doing effect than previous fixed-effects approaches. Precise estimation using the fixed-effect approach requires data from a sufficient number of institutions over a significant number of time periods to observe sufficient variation in volume. Previous studies have contained small sample sizes in terms of number of institutions, which may therefore be underpowered to detect a learning-by-doing effect [18]. Third, we control for a rich set of patient and ICU characteristics to minimise the risk of omitted variable bias. Lastly, this paper ties the less commonly reported literature on the underlying mechanism of the volume-outcome relationship, namely economies of scale and learning-by-doing, with mortality and provides useful information on how reconfiguring service lines may improve underperforming lower volume ICUs.

This study is organised as follows: In section 2, we describe the data. In section 3, we describe the empirical strategy. In section 4 we present the results of the main analysis and sensitivity analysis and in section 6 we present the discussion of the results. In section 7 we present our conclusions and recommendations.

6.2 Method

6.2.1 Data

Data was extracted from the Intensive Care National Audit and Research Centre Case Mix Program database which is a clinical database that covers all adult ICUs in England, Wales, Northern Ireland, United Kingdom[121]. Trained data collectors extracted detailed physiological, diagnostic, and sociodemographic data from consecutive adults admitted to ICUs in the United Kingdom participating in the Case-Mix Program database between 1 January 2010 and 31 December 2016[120]. Approval for the collection and use of patient identifiable data in the CMP was obtained under Section 251 of the National Health Service Act of 2006. We report an observational cohort study, as per Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines[206].

6.2.2. Exposure

The exposure was defined as by the quarterly sepsis volume with contemporaneous quarterly volume being the measure of the static scale effects and the lagged quarterly volume identifying the learning-by-doing effect.

6.2.3. Study outcome

The primary outcome was death before discharge from an acute hospital. Patients who were transferred between ICUs were excluded from the analysis of mortality but included in the estimation of ICU volumes. This was done to avoid confounding results with outcomes from different ICUs. For patients who were readmitted to the ICU, only the first admission was included in the mortality analysis. We chose the patient as the level of analysis. An ICU-level analysis of sepsis volume and mortality would smooth out variability in outcomes across

patients[201]. In the case of no observed learning effect, it would be unclear as to whether this is due to data aggregation or a true absence of a learning effect.

6.2.4. Empirical strategy

Baseline characteristics and unadjusted outcomes for the cohort were tabulated using standard summary statistics. We used a multivariate hierarchical logistic regression model to assess the association between volume and acute hospital mortality. Our model recognises that individual patients are clustered in quarters and nested in ICUs and provides a consistent estimate of the standard errors for clustered data.

Patient-level covariates include age, gender, ethnicity, functional status, co-morbidities, and sociodemographic status as measured by the Index of Multiple Deprivation. We measure patient level severity of illness using the $ICNARC_{2018}$ score. We include dummy variables for the presence of severe co-morbidities involving 7 organs systems. Functional status was categorised by the degree of assistance needed with activities of daily living. ICU characteristics are quarterly caseload volume, academic affiliation (non-university, university, university-affiliated) and quarterly throughput. Quarterly throughput is defined as the number of ICU admissions per ICU bed.

In the first step we will focus on overall learning curves in hospital mortality. We describe the basic model of Benkard with the important difference that the model presented here will involve multiple ICUs where the initial experience is unknown [207]. The simplest specification of the volume-outcome relationship is:

$$y_{ijq} = \beta_1 V_{jq} + \beta_2 V_{jq-1} + \sum_{m=1}^{M} \delta_m X'_{ijq} + \sum_{n=1}^{N} \varphi_n Z'_{jq} + \xi_j + \varepsilon_{ijq}$$

The dependant variable y_{ijq} is patient-level mortality of patient i in ICU j in quarter q. X'_{ijq} is a vector of patient characteristics and Z'_{jq} is a vector for ICU level characteristics. V_{jq} refers to quarterly ICU sepsis volume and captures the effects of static scale. The coefficient of V_{jq-1} is the sepsis volume in the preceding quarter and describes the ICU-level learning-by-doing effect. The ξ_{jq} captures to the ICU effect and ε_{ijq} is a classical error term.

We can expand the learning-by-doing component by including four lags of sepsis volume.

$$y_{ijq} = \beta_1 V_{jq} + \sum_{q'=q-1}^{q-4} \beta_{q'} V_{jq'} + \sum_{m=1}^{M} \delta_m X'_{ijq} + \sum_{n=1}^{N} \varphi_n Z'_{jq} + \xi_j + \varepsilon_{ijq}$$

The individual weights of $\beta_{q'}$ are called lag weights and they collectively constitute the lag distribution from $q' = q-1 \dots q-4$, with the full set of quarters being $q = \{q, q'\}$. The lags in volume estimate the effects of learning over time. If there was learning-by-doing and knowledge was passed on from one period to the next, we would observe a larger coefficient with each succeeding time period i.e., $\beta_{q-1} > \beta_{q-2} > \beta_{q-3} > \beta_{q-4}$. This is because learning-by-doing allows patients treated in the current time period to benefit from experience gained in the preceding time periods.

The ICU effect in the quarter $q-1$ makes it unnecessary to know the ICU's entire production history. We separate out the volume-outcome effects into its static and dynamic components. Instead of using cumulative learning treating all past periods as the same, we use lags of the previous quarters' volumes of sepsis. We compare the relative size of the coefficients. If the static scale economies are the main mechanism for the volume outcome relationship, then the coefficients of the lagged volumes would be small i.e. $(\beta_1 > \beta_{q-1} + \beta_{q-2} + \beta_{q-3} + \beta_{q-4})$. If the learning-by-going is important then the coefficients on the lagged volume would

be a larger proportion of the total effect. This would imply that experience gained in the past impacts the outcomes of the present. If the contemporaneous volume accounts for a larger proportion of the effect, then it would mean there are benefits to static scale. This would imply that any ICU high volume ICU would improve outcomes and that there would be benefits to indivisibilities of investments in infrastructure, favouring consolidation of critical care services.

Quarterly volumes are correlated over time, meaning that V_{jq} is correlated with V_{jq-1}, as are V_{jq-1} and V_{jq-2} correlated as well as $V_{jq-2}$and V_{jq-3}. High level correlation between regressors, referred to as multicollinearity, leads to unreliable coefficient estimates with large variances and standard errors. This leads to lag distributions in which the sequence of lag coefficients bounces between large and small and even sometimes positive and negative. We describe the distribution of the correlation coefficients between the volume lags. A weaker correlation between volume lags would support a low risk of multicollinearity.

6.2.5. Sensitivity analysis

The main variable of interest is in capturing the learning mechanism. In the primary analysis we used quarterly lags to identify a linear learning-by-doing-effect. We undertook several sensitivity analyses to identify other specifications of the learning-by-doing mechanism. First, we used monthly sepsis volume to detect any learning that may occur over shorter time periods. Second, we specify quarterly volume as a simple square root form to identify a non-linear learning-by-doing relationship.

6.3 Results

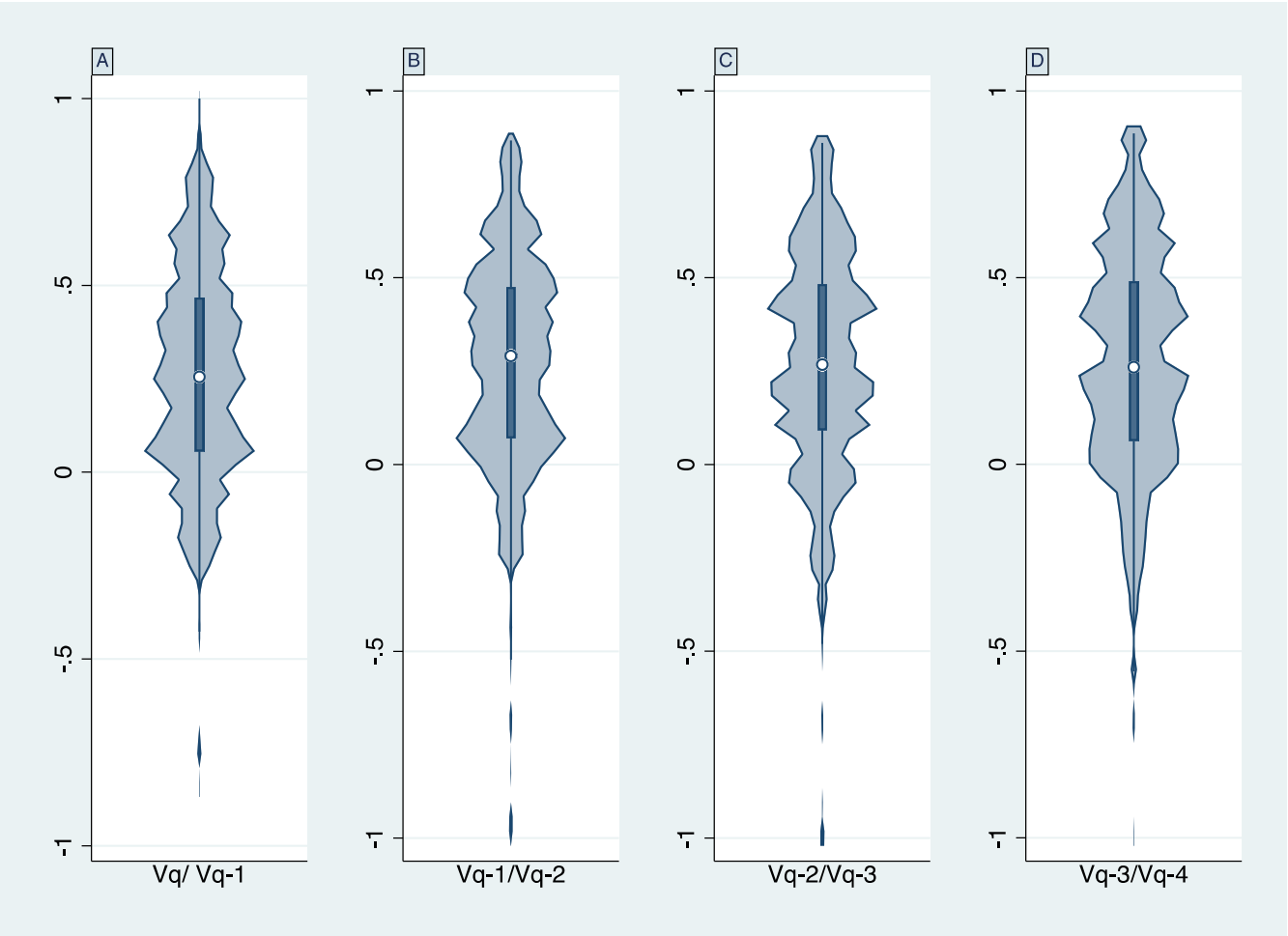
The study population of adult sepsis patients admitted to the ICU was 273,001. The median quarterly sepsis volume was 63 IQR [46-86]. **Error! Reference source not found.** summarises the patient characterises across quartiles of quarterly caseload volumes from 2010 to 2016. The mean age of patients was 63 (95% CI 63-63) years. A minority of patients had a severe comorbidity (20.3%) and most patients were functionally independent (68.0%). The mean $ICNARC_{2018}$ predicted risk of acute hospital mortality was 21.0% (95%CI 21.0b -21.0). Across quartiles of sepsis caseload volume, patients treated in the lowest quartile had higher acute severity of illness scores ICNARC score 21.3[21.2-21.4] compared with 20.5[20.5-20.6] in the highest quartile, $p<0.001$. Lower sepsis volume ICUs treated more patients with no chronic comorbidities compared with higher volume ICUs where 81.3% in the lowest quartile had no comorbidities compared with 76.3% in the highest quartile. On average, higher volume ICUs operate at higher occupancy and had higher throughput of patients.

Table 11.Patient characteristics across quartiles of quarterly sepsis volume

	Total	Quartile I [4-46]	Quartile II [47-63]	Quartile III [64-86]	Quartile IV [87-226]	P value
Age in years						
≤53	68357(25.0)	17490(24.4)	16787(24.4)	16785(25.3)	17295(26.4)	<0.001
54-66	68460 (25.1)	17957(25.0)	16947(24.6)	16702(25.2)	16854(25.6)	
67-76	70835(25.9)	18890(26.3)	18016(26.2)	17348(26.1)	15809(25.1)	
≥77	65339(23.9)	17452(24.3)	17094(24.8)	15573(23.5)	15520(23.0)	
Age in years	63[63-63]	64[64-64]	64[64-64]	63[63-63]	63[63-63]	<0.001
Male sex	148149(54.3)	38734(54.0)	37114(53.9)	36287(54.7)	36014(54.6)	0.004
Ethnicity						
White	277787(92.3)	73742(92.3)	69996(92.2)	68379(92.1)	65670(86.8)	<0.001
Asian	10723(3.5)	2682(3.4)	2225(2.9)	2310(3.1)	3506(4.6)	
Black	6192(2.0)	1328(1.7)	1353(1.8)	1066(1.4)	2445(3.2)	
Mixed/other	10855(3.5)	2178(2.7)	2261(3.0)	2375(3.2)	4041(5.3)	
ADLs						
Independent	184850(68.0)	49229(68.6)	46876(68.1)	43940(66.2)	44805(67.9)	<0.001
Some assistance	81913(30.1)	20927(29.2)	20416(29.7)	20985(31.6)	19585(29.7)	
Fully dependent	5071(1.9)	1325(1.9)	1252(1.8)	1187(1.8)	1307(2.0)	
ICNARC score mean (95% CI)	21.0[21.0-21.0]	21.3[21.2-21.4]	21.1[21.0-21.2]	21.0[20.9-21.0]	20.5[20.5-20.6]	<0.001
APACHE II, mean (95% CI)	18.4[18.4-18.5]	18.4[18.4-18.5]	18.3[18.3-18.4]	18.4[18.4-18.5]	18.6[18.6-18.7]	<0.001
Occupancy %, (95% CI)	72[72-73]	68 [68-68]]	72[72-72]	74[74-74]	77[77-77]	<0.001
ICU beds mean, (95% CI)	15[15-15]	9[9-9]	12[12-12]	16[16-16]	24[24-24]	<0.001
Quarterly throughput	5.0[5.0-5.0]	4.3[4.3-4.3]	5.1[5.1-5.1]	5.2[5.2-5.2]	5.3[5.3-5.3]	<0.001
Co-morbidities n(%)						
None	217655(79.7))	58354 (81.3)	55821 (81.1)	53,156 (80.0)	50324 (76.3)	<0.001
Cardiac disease	4857(1.8)	1398(2.0)	1151(1.7)	1243(1.9)	1065(1.6)	<0.001
Respiratory disease	12498(4.6)	3100(4.3)	3002(4.3)	2923(4.4)	3473(5.3)	
ESKD	5171[1.9)	953(1.3)	1075(1.6)	1230(1.9)	1913(2.9)	<0.001
Liver disease	6030(2.2)	1213(1.7)	1315(1.9)	1428(2.2)	2074(3.1)	<0.001
Metastatic cancer	6598(2.4)	1677(2.3)	1529(2.2)	1620(2.4)	1772(2.7)	<0.001
Hematologic malignancy	9763(3.6)	2377(3.3)	2235(3.3)	2341(3.5)	2810(4.3)	<0.001
Immunocompromised	24035(8.8)	6012(8.4)	5706(8.3)	5803(8.7)	6514(9.9)	<0.001
Septic shock	54419(19.9)	14961(20.8)	13907(20.2)	12703(19.1)	12848(19.5)	<0.001
ICU LOS (hrs)	163[162-164]	169[168-171]	159[157-160]	163[162-165]	159[157-161]	<0.001
Hospital LOS (days)	23[23-23]	23[23-23]	22[22-22z]	23[23-23]	24[24-24]	<0.001
ICU mortality	62277(22.8)	16868(23.5)	15919(23.1)	15196(22.9)	14294(21.7)	<0.001
Hospital Mortality	86728(31.8)	23650(32.9)	21944(31.9)	20932(21.6)	20202(30.6)	<0.001

Abbreviations: ADLs= Activities of daily living, ESKD = end stage kidney disease, LOS= length of stay. Data were missing for age n=10 (0.0%); ethnicity n=191 (0.1%); ADLS n=1167(0.4%); comorbidities n=1137 (0.4%); ICU mortality n=7(0.0%); hospital mortality n=1419(0.5%)

Figure 16.Violin plot of the distribution of the correlation coefficient between (A) (A) $\frac{Volume_{\ q}}{Volume_{\ q-1}}$, (B) $\frac{Volume_{\ q-1}}{Volume_{\ q-2}}$, (C) $\frac{Volume_{\ q-2}}{Volume_{\ q-3}}$, and (D) $\frac{Volume_{\ q-3}}{Volume_{\ q-4}}$.



The median correlation coefficient between $Volume_{\ q}$ and $Volume_{\ q-1}$ was 0.255 [IQR 0.056-0.467], median correlation coefficient between $Volume_{\ q-1}$ and $Volume_{\ q-2}$ was 0.290 [IQR 0.070-0.474] , median correlation coefficient between $Volume_{\ q-2}$ and $Volume_{\ q-3}$ was 0.267 [IQR 0.091-0.482] and the , median correlation coefficient between $Volume_{\ q-3}$ and $Volume_{\ q-4}$ was 0.260 [IQR 0.062-0.491]. These vales suggest that a weak to moderate correlation between lagged volumes and low risk of multicollinearity.

The distribution of the correlation coefficients between successive lags is described by the violin plots in **Error! Reference source not found.**. The median correlation coefficient between V_q and V_{q-1} was 0.255 [IQR 0.056-0.467], median correlation coefficient between V_{q-1} and V_{q-2} was 0.290 [IQR 0.070-0.474] , median correlation coefficient between V_{q-2} and V_{q-3} was 0.267 [IQR 0.091-0.482] and the , median correlation coefficient between V_{q-3} and V_{q-4} was 0.260 [IQR 0.062-0.491]. These vales suggest that a weak to moderate correlation between lagged volumes and low risk of multicollinearity.

In **Error! Reference source not found.** and **Error! Reference source not found.** we present the contemporaneous and lagged effects of volume on hospital mortality. We show that lagged effect is significant in each of the models. The sum of the total effect represented by the F-test is more reliably estimated across models. The F-test for the combined effect of the lagged volume remains significant. The change in the level of significance for the contemporaneous volume when the lagged volumes are added to the model implies some collinear effects. We would therefore interpret these results and identifying the dominant effect of learning-by-doing compared with economies of scale.

In **Error! Reference source not found.** we further explore the relative relationship between the contemporaneous and lagged volume. We first present the coefficients for each of the contemporaneous and lagged volumes. The sum of the results shows a consistent direction of effect across models. We then present an analysis of the contemporaneous volume relative to the total effect. This ratio suggests that the contemporaneous volume accounts for the smaller component of the total effect of volume. The results suggest a significant learning-by-doing effect in comparison with the contemporaneous economies of scale effect.

Table 12.Odds of acute hospital mortality with increasing number of lagged sepsis volume

	(1)		(2)		(3)		(4)		(5)	
	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
$Volume_q$	0.98***	0.98, 0.99	1.00	0.99, 1.01	1.01	0.99, 1.02	1.01	1.00, 1.02	1.00	0.99, 1.02
$Volume_{q-1}$			0.99**	0.98, 1.00	1.00	0.99, 1.01	1.00	0.99, 1.01	1.00	0.99, 1.01
$Volume_{q-2}$					0.98***	0.97, 0.99	0.98**	0.97, 0.99	0.99**	0.97, 1.00
$Volume_{q-3}$							0.99*	0.98, 1.00	0.99	0.98, 1.00
$Volume_{q-4}$									1.00	0.99, 1.02
ADLs										
Independent	1.00		1.00		1.00		1.00		1.00	
Some assistance	1.09***	1.06, 1.11	1.08***	1.06, 1.11	1.09***	1.06, 1.12	1.08***	1.06, 1.11	1.09***	1.06, 1.11
Total assistance	1.37***	1.26, 1.48	1.35***	1.24, 1.46	1.36***	1.25, 1.48	1.35***	1.24, 1.47	1.35***	1.24, 1.48
Ethnicity										
White	1.00		1.00		1.00		1.00		1.00	
Asian	0.99	0.93, 1.05	0.99	0.93, 1.05	1.00	0.93, 1.06	0.99	0.93, 1.06	1.00	0.93, 1.06
Black	0.84***	0.78, 0.92	0.84***	0.77, 0.91	0.85***	0.78, 0.93	0.84***	0.77, 0.92	0.84***	0.76, 0.92
Mixed/other	0.99	0.93, 1.05	0.98	0.92, 1.05	0.99	0.93, 1.06	0.98	0.91, 1.04	0.97	0.91, 1.04
Age in years	1.01***	1.01, 1.01	1.01***	1.01, 1.01	1.01***	1.01, 1.01	1.01***	1.01, 1.01	1.01***	1.01, 1.01
Male	1.11***	1.09, 1.13	1.12***	1.09, 1.14	1.12***	1.09, 1.14	1.12***	1.09, 1.14	1.12***	1.09, 1.14
ICNARC score	1.05***	1.05, 1.05	1.05***	1.05, 1.05	1.05***	1.05, 1.05	1.05***	1.05, 1.05	1.05***	1.05, 1.05
Comorbidity										
Severe respiratory disease	1.07**	1.02, 1.12	1.07**	1.02, 1.12	1.06*	1.01, 1.12	1.07*	1.01, 1.12	1.07*	1.01, 1.13
Very severe cardiovascular	1.37***	1.27, 1.47	1.38***	1.28, 1.48	1.37***	1.27, 1.48	1.38***	1.28, 1.49	1.40***	1.29, 1.51
ESKD	1.35***	1.26, 1.45	1.34***	1.24, 1.44	1.32***	1.23, 1.42	1.32***	1.23, 1.43	1.34***	1.24, 1.44
Severe liver disease	1.55***	1.45, 1.66	1.55***	1.45, 1.66	1.54***	1.44, 1.65	1.54***	1.43, 1.65	1.53***	1.42, 1.65
Metastatic disease	1.18***	1.11, 1.26	1.18***	1.11, 1.26	1.20***	1.12, 1.28	1.18***	1.10, 1.27	1.17***	1.09, 1.25
Haematological malignancy	1.11***	1.05, 1.18	1.11***	1.05, 1.17	1.12***	1.05, 1.19	1.11***	1.05, 1.18	1.10**	1.04, 1.17
Immunocompromised	1.17***	1.12, 1.21	1.16***	1.11, 1.20	1.15***	1.10, 1.20	1.15***	1.10, 1.20	1.15***	1.10, 1.20
Quarterly throughput	1.02**	1.01, 1.03	1.01*	1.00, 1.03	1.00	0.99, 1.02	1.00	0.99, 1.02	1.01	0.99, 1.02
Academic affiliation										
Non-university	1.00		1.00		1.00		1.00		1.00	
University affiliated	0.98	0.90, 1.07	0.98	0.90, 1.07	0.98	0.90, 1.07	0.98	0.90, 1.07	0.98	0.89, 1.07
University	1.04	0.97, 1.12	1.05	0.97, 1.13	1.05	0.98, 1.14	1.05	0.97, 1.13	1.05	0.97, 1.14
IMD	1.00***	1.00, 1.00	1.00***	1.00, 1.00	1.00***	1.00, 1.00	1.00***	1.00, 1.00	1.00***	1.00, 1.00
AIC		232558.12		221261.94		210757.87		201056.27		189384.57
BIC		232798.75		221511.92		211017.09		201324.66		189661.76
F-statistic			8.03		21.19		22.02		7.33	
p-value			0.0046		<0.001		<0.001		0.0068	

P<0.05=*. P<0.01=**P<0.001=***

Table 13.Coefficients from logistic regression model showing coefficients, the proportion of the contemporaneous volume to the total volume effects and the ratio of lagged coefficients.

Lag depth	Model 1	Model 2	Model 3	Model 4	Model 5
$\boldsymbol{0}$	-0.0167	-0.0356	0.0065	0.0105	0.0043
1		-0.1355	-0.0028	-0.0024	-0.0004
2			-0.0231	-0.0173	-0.0150
3				-0.0105	-0.0092
4					0.0047
Total	-0.0167	-0.1711	-0.1944	-0.0197	-0.0154
Ratio $\frac{\boldsymbol{\beta_1}}{\sum_{n=0}^{\infty}\boldsymbol{\beta_n}}$	1.00	0.21	0.03	0.53	0.28
$\frac{\boldsymbol{\beta_2}}{\boldsymbol{\beta_3}}$			0.1212	0.1387	0.0267
$\frac{\boldsymbol{\beta_3}}{\boldsymbol{\beta_4}}$				1.6476	1.6305
$\frac{\boldsymbol{\beta_4}}{\boldsymbol{\beta_5}}$					-1.9574
F -test		8.03	21.19	22.02	7.33
		0.0046	<0.001	<0.001	0.0068

P<0.05=*. P<0.01=**P<0.001=***

The absence of a consistent $\frac{\beta_n}{\beta_{n-1}} > 1$ suggest some degradation of the institutional learning-by-doing effect.

6.3.1. Sensitivity analysis

We undertook an analysis using the months as the time epoch for the learning-by-doing effect. The assumption is that learning may occur over a shorter period than quarterly. Monthly lags did not identify a shorter time window for learning. The second sensitivity analysis explored a non-linear relationship with learning. There alternate specifications of learning did not alter the main results and are described in **Error! Reference source not found.**.

Table 14. Sensitivity analyses. Odds ratio and 95% confidence interval for acute hospital mortality using the alternate specification of (a) monthly lags to evaluate a shorter time window for the learning-by doing effect and (b) using a simple square-root of the quarterly lagged volumes to evaluate the non-linear learning effect.

(a) Month sepsis volume lag										
	(6)		(7)		(8)		(9)		(10)	
	OR	95%CI	OR	95%CI	OR	95%CI	OR	95%CI	OR	95%CI
$Volume_{\ m}$	0.96***	0.94,0.98	0.97*	0.95,1.00	0.98	0.95,1.01	0.99	0.97,1.02	0.99	0.97,102
$Volume_{\ m-1}$			0.99	0.97,1.01	0.99	0.97,1.01	1.00	0.98,1.02	1.00	0.98,1.02
$Volume_{\ m-2}$					0.98	0.96,1.00	0.99	0.97,1.01	0.99	0.97,1.01
$Volume_{\ m-3}$							0.99	0.97,1.01	1.00	0.97,1.02
$Volume_{\ m-4}$									0.97	0.95,0.99
F- test						5.52	5.26		9.12	
p-value						0.0188	0.0218		0.0025	
(b) Square root of quarterly sepsis volume										
	(11)		(12)		(13)		(14)		(15)	
	OR	95%CI	OR	95%CI	OR	95%CI	OR	95%CI	OR	95%CI
$Volume_q$	0.97***	0.96,0.99	0.99	0.97,1.01	1.01	0.99,1.03	1.02	1.00,1.04	1.01	0.98,1.03
$Volume_{q-1}$			0.98*	0.96,1.00	1.00	0.98,1.01	1.00	0.98,1.01	1.00	0.98,1.02
$Volume_{q-2}$					0.96***	0.94,0.98	0.97**	0.95,0.99	0.98**	0.96,0.99
$Volume_{q-3}$							0.98*	0.97,1.00	0.98	0.97,1.00
$Volume_{q-4}$									1.01	0.99,1.02
F-test			6.52		18.84		21,07		6.25	
p-Value			0.0107		<0.001		<0.001		0.0124	

P<0.05=*. P<0.01=**P<0.001=***

6.4. Discussion

This study evaluates whether the volume-outcome relationship arises primarily from scale economies or due to learning-by-doing. The study found a significant learning-by doing effect as proxied by the lagged quarterly sepsis caseload volumes for patients with sepsis treated in the ICU between 2010 and 2016.

This study contrasts with previous studies on learning-by doing versus scale effects by Gaynor et al. and Ho et al. [15, 201] focused on cardiac procedures. There are two aspects of sepsis and critical care that may explain the difference with cardiac procedures. The first might be related to the routinisation of the procedure[208]. It is likely that learning by doing differs across conditions and may be more significant for sepsis than for cardiac procedures. Sepsis is a time-critical complex disease that requires the collective skill of the entire treating team. Sepsis afflicts a wide spectrum of patients with varied comorbidities and requires a variety of interventions. This is unlike cardiac surgical patients who have similar risk factor profiles and require one of two procedure valve and coronary artery surgery. Outcomes for routine procedures would therefore likely depend on scale effects more than experience.

The second issue is that accumulated expertise is likely to not be completely retained between time periods. This may reflect the cyclical nature of emergency medical teams. As doctors and nurses leave the service, the ICU loses the benefits of their accumulated experience[209]. Data from other healthcare contexts are consistent with the idea that the depreciation of experience is related to the staff turnover [210]. Learning in the context of surgical patients may be largely dependent on the experience of individual surgeons and

therefore more sensitive to staff turnover. Surgical cohorts such as cardiac surgery, caesarean section, Whipple's procedure, and abdominal aortic repair surgery all describe a degree of depreciation[19, 210, 211]. Cardiac surgery has a low turnover of staff, so has low depreciation of experience[210, 212]. In contrast Whipple's procedure and caesarean section describe high levels of depreciation of experience[208, 211]. This is consistent with evidence of knowledge depreciation from other industries. For example, Benkard described increases in the labour hours required to produce airplanes during lulls in production. [207]. The inconsistent coefficients for the learning-by-doing effect with regards to sepsis might relate to the depreciation of organisational learning and the low temporal stability of ICU teams. The depreciation of knowledge for complex medical treatments has important implications for patient care. It is important that ICUs maintain caseload volumes over time to preserve institutional knowledge.

This study has limitations. Firstly, one of the challenges of including both contemporaneous and lagged volumes is the likely multicollinearity. ICUs that treat many sepsis patients in one quarter are likely to treat a large number the following quarter. We use data from 231 ICUs and assume that the large sample size will contain sufficient variation between ICUs which would weaken the collinearity between the contemporaneous and lagged volume. We undertook a robustness check that found a weak to moderate correlation between lagged volumes and low risk of multicollinearity. Secondly, this study does not contain details about the compliance to evidence-based processes of care. Previous literature suggests higher volume hospitals have higher adherence to processes of care such as antibiotic administration and venous thrombo-embolism prophylaxis than lower-volume hospitals[213]. Whether this is due to the static scale effects of volume or the experience gained form learning by doing is

unclear. The adherence to processes of care also do not fully explain the volume-outcome relationship [194]. We therefore contend that conclusions of this study would not be substantially altered by controlling for compliance with evidence-based processes of care.

6.5. Conclusion and recommendations

This study supports the idea that the dominant mechanism by which volume leads to improved outcomes is through dynamic learning-by-doing as opposed to the static scale effects. ICUs tend to improve by caring for a large volume patients distributed over time. Centralisation has been proposed to improve ICU efficiency and quality, however, this may not fully leverage the benefits of the learning-by-doing mechanism[196]. Another argument against centralisation is the potential to introduce socioeconomic inequalities in access to critical care[214]. Patients may therefore be better served by ICUs organised to achieve minimum volume standards without centralisation.

Chapter seven: The association between ICU specialisation and mortality in the UK

Abstract

Importance: Proponents argue that increasing specialisation of critical care services would improve quality, but these theoretical benefits remain unproven.

Objective: To empirically determine whether increasing Intensive Care Unit (ICU) specialisation reduces mortality for critically ill patients.

Design: Retrospective cohort analysis of patients (≥ 16 years) admitted to 231 ICUs in the United Kingdom from 2010 to 2016.

Setting: All 231 ICUs in the United Kingdom

Participants: Adult patients (≥ 16 years) patients with critical illness admitted to ICUs in the UK

Exposure: ICU specialisation, which is defined as the share of patients within different categories of disease.

Main Outcome(s) and Measure(s): We used a multivariate hierarchical logistic regression model to evaluate the association between ICU specialisation and acute hospital mortality

Results: Of a total of 933,284 patients admitted to 231 ICUs, 513750(55%) were male and the median age was 65(interquartile range [IQR] 50-75) years. Hospital mortality was 21.2%. The median sepsis specialisation was 0.30 [IQR 0.23-0.37], cardiac specialisation 0.05 [IQR

0.01-0.09], neurosurgery 0.02 [IQR 0-0.06], trauma 0.12 [IQR 0.08-0.12], medical specialisation was 0.27 [IQR 0.20-0.34, elective surgery 0.13 [IQR 0.06-0.22], and emergency surgery 0.06 [IQR 0.04-0.10]. Overall, there was no reduction in the odds for mortality when patients were treated in a specialist ICUs, OR 1.06 (95%CI 1.0-1.08, p<0.001) for sepsis, OR 1.01(95%CI 0.99-1.04, p=0.292) for cardiac, neurosurgery (OR 1.01 95% CI 0.98-1.04), trauma OR 1.03(95% CI1.01-1.05, p=0.010), medical OR 1.03(95% CI 1.01-1.06, p=0.007), elective surgery OR 0.96(95%CI 0.94-0.99, p=0.002) and emergency surgery 1.00(95%CI 0.97-1.03, p=0.867). These findings were consistent across alternate definitions of specialisation and within subgroups of the most severely ill patients.

Conclusions and Relevance: Speciality ICUs do not have significantly lower hospital mortality for critically ill patients in the UK after adjusting for patient characteristics and caseload volume. This has relevance for policymakers, payers and clinicians interested in the future organisation of critical care services in that while there may be benefits from high volume ICUs, there is no compelling evidence demonstrating added value from specialist ICUs.

7.1 Introduction

The potential for efficiency gains from specialisation was first proposed by Adam Smith more than 230 year ago [31, 215, 216]. Later the 19th century economist David Ricardo suggested there may be gains from relative specialisation[216]. In healthcare, where inefficiency is a major concern and specialisation has been long been advocated to improve quality and lower costs, but empirical evidence is scarce[217-219]. Overall health care costs are dominated by hospital expenditure[220]. Within the envelope of hospital costs, critical care services are an integral and costly component, accounting for up to 40% of hospital spending [61]. As a consequence of increasing case complexity and an aging population, high income countries like the UK are projected to have exponential increases in demand for critical care[221]. This makes identifying strategies that make more efficient use of critical care services imperative. If there are gains from specialisation redesigning critical care services towards specialist intensive care units (ICUs) could theoretically promote value to both patients and providers [222]. Patients in a specialist ICUs could conceivably be exposed to more standardised care and thus benefit from focused clinical expertise, equipment, and other resource. Staff in the ICU benefits from having fewer competing operational objectives and fewer variable costs[223].There are also potential harms to specialisation. Specialised ICUs could require an inter-ICU transport network to sort patients into the appropriate specialist ICU, families may face increased travel times, care might be fragmented care between local and specialist services and there may be delays in care as patients wait for access to specialist ICUs.

The aim of this paper is to compare the effects of ICU specialization among disease groups in terms of acute hospital mortality. We aim to solve is to determine the optimal types of patients in the ICU using patient level data covering several years.

The remainder of this paper has the following structure: section 2 describes the institutional context. Section 3 describes related literature Section 4 discusses the model, section 5 the data we use in the analysis. Section 6 presents the results. Section 7 discusses the implications of the results and concludes.

7.2 Institutional context

This study was conducted in the 231 ICUs in England, Wales, and Northern Ireland where the National Health Service (NHS) is funded by general taxation. Patients face no charges for hospital care including critical care. All doctors employed in the NHS are salaried and have no share of hospital profits and no financial interest in recommending any treatment [224]. Critical care services deliver care in specialised wards called ICUs but even within these specialised wards there is variety with ICUs offering a range of services for patients with sepsis, cardiac, neurosurgery, trauma, medical, elective, and emergency care. The diagnosis types of patients appear in different proportions and combinations across ICUs. Patients receiving ICU care do not choose their ICU and are often transported via ambulance to their nearest hospital hence ICUs do not select patients based on their risk [225].

7.3 Literature review

Specialisation in healthcare has been investigated in a number of departments with inconsistent results[226].

Our study adds to the existing literature in the following ways. Firstly, most of the current literature on this topic has focused on specialisation at the hospital level[50, 227]. The evidence regarding specialist hospitals does not consistently suggest improvement in quality or costs [50, 136, 227]. Emerging evidence has suggested that the benefits to specialization may aggregated to the focal segmented service level more than the hospital level[228-230] [228, 231]. This is consistent with the idea of a "plant-within-a-plant" proposed by Skinner in the manufacturing sector[230].

Some studies have examined specialisation among hospital departments but largely focused on elective surgical diagnoses particularly cardiac, neurosurgery and orthopaedic surgery[227, 232-236]. Many of these studies fail to show the benefits of specialisation after adjustment for caseload volume and patient characteristics[50, 227, 228, 232, 237]. A systematic review of specialised hospitals does not consistently improve cots or quality [238].

Second, most of the existing evidence for specialisation relies on analyses from selected, small for-profit physician owned hospitals operating in the US hospital sector[238]. Our study makes use of a large, nationally representative clinical dataset with detailed risk adjustment that includes all ICUs in a publicly funded health system[239]. This addresses the previous concerns of physician-owned speciality services that may be motivated by financial incentives to treat healthier and wealthier patients[240].

Considering the ICU as a multiproduct firm, focused activities within an operating unit may have synergistic effects the lead to improved outcomes for all patients within the unit[241]. The ideal composition of patients admitted to the ICU remains uncertain and in general the literature is scarce. The largest study of specialist ICUs including a diverse population of

critically ill patients did not find any improved survival from admission to a specialist ICU designated by diagnostic group[47]. The study included a high proportion of cardiac surgical patients which limits its generalisability[47]. Our study adds to our understanding by focusing on service level specialisation and includes patients with a diverse set of diseases commonly seen in ICUs.

7.4 Methods

7.4.1. Data

We report an observational cohort study, as per Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines[206]. Trained data collectors extracted detailed physiological, diagnostic, and sociodemographic data from consecutive adults admitted to ICUs in the United Kingdom participating in the Case-Mix Program database between 1 January 2010 and 31 December 2016[120]. The data undergo extensive local and central validation. Approval for the collection and use of patient identifiable data in the CMP was obtained under Section 251 of the National Health Service Act of 2006.

7.4.2. Exposure: Specialization measures

There is no regulatory requirement to obtain the title of specialist ICU nor is there a gold standard measure of specialization. In the medical literature, specialization refers to the channelling of resources and labour towards a narrow focus on certain diseases or diagnostic groups that can be considered service lines[242]. Deriving information about the extent to which the ICU is specialist is not directly observable and must be constructed from data.

The patients in the ICU can categorised into diagnostic groups. We used the ICNARC Coding Method (ICM) to identify diagnostic categories of patients. The ICM is a standardized method

that records the reason for ICU admission that uses a five-tier hierarchical system to classify patients. Details about the ICM are provided in the appendix for chapter seven[34]. We identify the following seven types of patients based on a specific disease profile (1. Sepsis, 2. cardiac (non-sepsis) 3. trauma 4. neurological 5. medical (non-sepsis, cardiac, trauma, respiratory) 6. emergency surgery (non-sepsis, trauma, cardiac, respiratory) and 7. elective surgery (non-sepsis, cardiac, respiratory, trauma. The most common approaches in quantifying specialization uses the share of the various diagnostic groups [232, 243]. We describe both the absolute and relative specialisation.

7.4.2.1 Absolute specialisation

Absolute specialisation in each diagnostic category is ($S_{kjm} = n_{kjm}/N_{jm}$), where S is the share of patients in diagnostic category k in ICU j in month m. For sepsis specialization $k = sepsis$ and for neurosurgery specialization, $k = neuro\ surgery$.

The value of S_{kjm} ranges from zero to one. Zero would mean that there are no patients from a diagnostic group k in ICU j in month m and a value of 1 would mean that all the patients in that ICU for the same period are of that diagnostic group k. This approach does not compare specialisation in diagnostic group k to a benchmark of the share of k in other ICUs. The main benefit of using absolute specialisation is that it can provide a measure of change within the same ICU over time.

7.4.2.2 Relative specialisation

Relative specialisation describes the deviation from average share of a diagnostic category k across all ICUs in month m. This accounts for the fact that some diagnostic groups account for a large or small share in all ICUs.

The relative specialisation (RS) is calculated as follows:

$$RS = \left[\frac{n_{kjm}}{N_{jm}}\right] / [\frac{n_{km}}{N_m}]$$

where n_{kjm} is the total number of patients of a diagnostic category k in an ICU j in month m, N_{jm} refers to the total number of patients admitted to ICU j in month m. n_{km}refers to all patients in a diagnostic group k admitted in month m. N_m refers to all patients admitted to all ICUs in month m. The RS has a value of 1 if the share of the diagnostic category k is the same as the average share across all ICUs in that month m. Values greater than 1 imply higher levels of relative specialisation. A value of zero implies no patients of diagnostic group k.

Relative specialisation provides a measure of comparative advantage of one ICU over another. As an example, if the absolute S_{kjm} is low but the relative share may be above the average then it indicates higher levels of specialization than the average ICU. Relative specialisation may show changes in specialization when the reference level changes even when the absolute specialisation remains constant. Relative specialisation is a measure of specialization relative to other ICUs and can been seen as a measure of centralization.

7.4.3. Outcome measures

All patients were followed up until ultimate discharge from acute hospital. We used acute hospital mortality after ICU admission as a measure of quality. We excluded interhospital transfers. For ICU readmissions, we included the co-variates from the index admission but outcome from the entire episode.

7.4.4. Statistical analyses

Baseline characteristics and unadjusted outcomes for the cohort were tabulated using standard summary statistics. We used a multivariate hierarchical logistic regression model to assess the association between ICU specialisation and acute hospital mortality. Our model recognises that individual patients clustered in months and nested in ICUs and provide a consistent estimate of the standard errors for clustered data. A more detailed description is provided in the appendix for chapter seven.

Hierarchical models overcome both the atomistic fallacy of individual risk factor epidemiology and the ecological fallacy of aggregated data[244]. This approach aims to answer the question of whether higher level contexts such as specialisation impact on individual patients and in what magnitude?

To account for the hierarchical structure of the data, patients i (level 1) within the same month m (level2) are within the ICU j (level 3), a 3-level hierarchical logistic regression model is specified as follows:

$$logit = \left\{Pr\left(y_{ijm} = 1|X'_{ijm}, u_{jm}^{(2)}, u_j^{(3)}\right)\right\}$$

$$= \beta_0 + \beta_1 S_{kjm} + \beta_2 Vol_{jm} + \beta_3 ICNARC_{ijm} + \beta' \Omega'_{ijm} + \beta' \Phi_j + u_{jm}^{(2)} + u_j^{(3)}$$

The outcome variable, acute hospital mortality, y_{ijm} given a set of covariates, X'. The exposure is level of specialisation for diagnostic group k in the ICU S_{kjm} for specialisation. In the model , , X' includes monthly caseload volume Vol_{jym} , and patient level severity of illness as measured by the ICNARC$_{2018}$ score, $ICNARC_{ijm}$. Here Ω'_{ijm} is a vector of all other

patient-level covariates, Φ'_j is a vector for other ICU-level covariates, $u_{jm}^{(2)} \sim N(0, \psi^{(2)})$ is the random intercept for the month with the ICU j and $u_j^{(3)} \sim N(0, \psi^{(3)})$ is the random intercept varying over ICUs. The random intercepts $u_{jm}^{(2)}$ and $u_j^{(3)}$ are assumed to be independent of any covariates (exogenous). The random intercept $u_j^{(3)}$ represents unexplained variation in mortality across ICUs, and the random intercept $u_{jm}^{(2)}$ represents unexplained variation between months in the same ICU. The hierarchical model accounts for the clustering of the data within months and ICUs and provide a more accurate estimate of standard errors for clustered data than fixed effects models[245]. Hierarchical models have lower units of aggregation nested within higher units. The hierarchical model framework can overcome a group level omitted variable. The addition of a group level average covariate $\overline{X_j}$ will absorb all correlations between X and the group level random effects. This approach is equivalent to a fixed effects approach. Hierarchical models can also test for heterogeneity in coefficients across groups. Hierarchical models allow a framework to account for clustering, omitted group-level variables and heterogeneity of coefficients between ICUs.

Patient-level covariates include age, gender, ethnicity, functional status, co-morbidities, and sociodemographic status as measured by the Index of Multiple Deprivation. We measure patient level severity of illness using the $ICNARC_{2018}$ score and dummy variables for the occurrence of cardiac, respiratory, renal haematologic and neurologic failure[246]. The $ICNARC_{2018}$ is a score from 0-100 that includes physiological parameters, chronic medical conditions, level of dependency prior to admission, source of admission, primary reason for admission and the receipt of mechanical ventilation and cardiopulmonary resuscitation[246]. We include dummy variables for the presence of severe co-morbidities involving 7 organs

systems. Functional status was categorised by the degree of assistance needed with activities of daily living. ICU characteristics included in the model are monthly caseload volume, academic affiliation (non-university, university, university-affiliated) and monthly throughput. Throughput was the number of ICU admissions in a month per ICU bed.

7.4.4.1. Sub-group analysis

We conducted two subgroup analyses to identify groups of patients most likely to benefit from ICU specialisation. Firstly, we assessed the effects of specialisation on hospital mortality within each diagnostic group. In this analysis we consider the benefits of specialisation when patients are treated in the ideal specialist ICU for their disease type. i.e., sepsis patients in a sepsis specialised unit, cardiac patients in a cardiac specialised unit, neurosurgical patients in a neurosurgical unit, medical patients in a medical ICU and elective and emergency surgical patients in the respective specialised units. Secondly, we conducted an analysis of the subgroup of the most severely ill patients, defined as the highest quartile of $ICNARC_{2018}$ score.

7.4.4.2. Robustness check

We evaluate the robustness of the primary results to alternate specifications of specialization. We used an exponent of share to test the potential non-linear relationship of specialisation with mortality and used relative difference as an alternate measure of relative specialisation. Further details are provided in the appendix for chapter seven.

7.4.4.3. Sensitivity analysis

We undertook several sensitivity analyses. First, we added time fixed effects to the empirical specification to capture technological improvements assuming all ICUs in the same year

applied the same technologies. The primary specification of leaving out the year fixed effects and using the hierarchical structure makes a less restrictive assumption of assuming that each ICU correlated with itself over time. Second, we considered ICU mortality instead of acute hospital mortality as the outcome. This approach favourably allocates patients that survive to ICU discharge but subsequently die before hospital discharge. This approach offers a more optimistic analysis of the benefits of ICU specialisation and an absence of mortality benefit would support the primary analysis. We considered the benefits of specialisation for the subgroup of the highest quartile of severity illness treated in the ideal specialist ICU i.e., sepsis patients in the highest quartile of ICNARC score treated in the sepsis specialised ICU, cardiac patients in the highest quartile of ICNARC score treated in cardiac specialised units, neurosurgical patients in the highest quartile of ICNARC score treated in the neurosurgical ICUs and trauma patients in the highest quartile of ICNARC score treated in trauma specialised ICUs. Lastly, we considered the effects of ICU specialisation on the non-specialised cohort within the ICU. These results are included in the supplement. In all analyses, P values less than .05 were considered significant.

7.5. Results

In total there were 933,284 patients admitted to the ICU between 2010 and 2016 (eFigure 1, Appendix for chapter seven). Patient characteristics are shown in **Error! Reference source not found.** and the supplementary appendix for chapter seven, eFigures 2-7. The mean age was 61.1 year (95% CI 61.0-61.1years) and 45% were female. Generally, most patient admitted to the ICU had no comorbidities (82.2%) and were fully independent (75.5%). The unadjusted hospital mortality for all ICU patients is 21%. Sepsis (29.2%) and general medical patients

27.4%) account for the largest overall share of patients. Cardiac surgery (6.4%) and neurosurgery (5.3%) contribute the smallest share of patients.

On average, sepsis patients tend to be older with higher severity of illness and more comorbidities than other disease types. Sepsis patients had longer ICU and hospital stays and higher mortality. Elective surgical patients account for 16.2% of patients. These patients have a low acuity of acute illness but a higher burden of chronic disease. Elective surgical patients have the shortest ICU and hospital stay and a low acute hospital mortality (2.7%) (**Error! Reference source not found.**).

There was considerable variability of shares of patients across ICUs (**Error! Reference source not found.** and **Error! Reference source not found.**). The median absolute sepsis specialisation was 0.30 [IQR 0.23-0.37], cardiac 0.05[IQR 0.01-0.09], neurosurgery 0.02[IQR 0-0.06], trauma 0.13[IQR 0.08-0.17], medicine 0.27[IQR 0.21-0.34], elective surgery 0.13[IQR 0.06-0.22] and emergency surgery 0.6[IQR 0.04-0.010]. In terms of relative specialisation, the median for sepsis was 0.98[IQR 0.76-1.22], cardiac 0.72[IQR 0.16-1.50], neurosurgery 0.41[IQR 0-1.17], trauma 0.95[IQR 0.63-1.30, medicine 0.96[IQR 0.74-1.2], elective surgery 0.81[IQR 0.36-1.42] and emergency surgery 0.88[0.48-1.37].

Table 15. Patient characteristics across speciality groups

Variable	Total* N=933284 (100%)	Sepsis N=273001 (29.3%)	Cardiac surgery N=59844 (6.4%)	Trauma N=121882 (13.1%)	Neurosurgery N=49684 (5.3%)	Medical N=256092 (27.4%)	Elective surgery N=158064 (16.9)	Emergency surgery N=69280 (7.4%)
Age(years), n(%)								
≤ 50	245157 (26.3)	57587 (21.1)	6986 (11.7)	55537 (45.6)	16717 (33.7)	80820 (31.6)	25168 (15.9)	17094 (24.7)
51-65	240390 (25.8)	72191 (26.4)	14566 (24.3)	24481 (20.1)	16276 (32.8)	65861 (25.7)	46006 (29.1)	15322 (22.1)
66-75	223081 (23.9)	70718 (25.9)	19548 (32.7)	17624 (14.5)	10569 (21.3)	55396 (21.6)	46653 (29.5)	15111 (21.8)
≥76 years	224617 (24.0)	72495 (26.6)	18744 (31.3)	24235 (19.9)	6122 (12.3)	54001 (21.1)	40234 (25.5)	21747 (31.4)
missing	39(0.0)	10 (0.0)	0(0.0)	5(0.0)	0(0.0)	14(0.0)	3(0.0)	6(0.0)
Dependency								
Fully independent	704514 (75.5)	184850 (67.7)	47979 (80.2)	100795 (82.7)	42752 (86.1)	187655 (73.3)	130823 (82.8)	53438 (77.1)
Some assistance	213182 (22.8)	81913 (30.0)	11395 (19.0)	18998 (15.6)	6256 (12.6)	63313 (24.7)	26151 (16.5)	15195 (21.9)
Fully dependent	9726(1.0)	5071(1.9)	185(0.3)	539(0.4)	292(0.6)	3004(1.2)	605(0.4)	373(0.5)
missing	5862(0.6)	1167(0.4)	285(0.5)	1550(1.3)	384(0.8)	2120(0.8)	485(0.3)	247(0.4)
Ethnicity								
White	841450(90.2)	248275(90.9)	55420(92.6)	111003(91.1)	43941(88.4)	224350(87.6)	144735(91.6)	63125(91.2)
Asian	31884(3.4)	9438(3.5)	1056(1.8)	2907(2.4)	1739(3.5)	12031(4.7)	3928(2.5)	2212(3.2)
Black	20630(2.2)	5504(2.0)	752(1.3)	1950(1.6)	1110(2.2)	8062(3.2)	2835(1.8)	1407(2.0)
Mixed/other	38499(4.1)	9617(3.5)	2569(4.3)	5957(4.9)	2817(5.7)	11436(4.5)	6314(4.0)	2487(3.6)

Missing	821(0.1)	167(0.1)	4.7(0.1)	65(0.1)	87(0.2)	213(0.1)	252(0.2)	49(0.1)
Age in years	61.1[61.0,61.1]	63.3[63.2,63.4]	67.3[67.2,67.5]	53.7[53.6,53.8]	56.3[56.2, 56.5]	58.6[58.5,58.7]	64.9[64.8,65.0]	63.3[63.2,63.5]
Sex								
Female	419533(45.0)	124852(45.7)	15254(27.2)	51535(42.3)	20832(41.9)	115858(45.2)	75751(47.9)	36369(52.5)
Male	513750(55.0)	148149(54.3)	43590(72.8)	70347(57.7)	28852(58.1)	140234(54.8)	82313(52.1)	32921(47.5)
Missing	1(0.0)	0(0.0)	0(0.0)	0(0.0)	0(0.0)	0(0.0)	0(0.0)	0(0.0)
APACHE II score	15.6[15.6-15.6)	18.4(18.4-18.5)	13.9(13.9-14.0)	13.1(13.1-13.2)	12.2(12.2-12.3)	17.1(17.1-17.2)	12.1(12.1-12.2)	13.8(13.7-13.9)
ICNARC score	16.8(16.8-16.8)	21.0(20.9-21.0)	13.9(13.8-14.0)	15.4(15.3-15.4)	12.1(12.0-12.1)	19.2(19.2-19.3)	10.4(10.4-10.5)	14.3(14.3-14.4)
Co-morbidities								
Any-comorbidities	164266 (17.6)	54209 (19.9)	5816 (9.7)	8680 (7.1)	5270 (10.6)	48528 (19.0)	34865 (22.1)	13311(19,2)
Severe respiratory disease	23917(2.6)	12498(4.6)	915(1.5)	1363(1.1)	549(1.1)	6575(2.5)	1783(1.1)	1139(1.6)
Severe cardiac disease	15851(1.7)	4847(1.8)	1596(2.7)	1182(1.0)	236(0.5)	5773(2.3)	1830(1.2)	958(1.4)
ESKD	18549(2.0)	5171(1.9)	802(1.3)	872(0.7)	233(0.5)	8897(3.5)	1757(1.1)	1289(1.9)
Chronic liver disease	24184(2.6)	6030(2.2)	237(0.4)	998(0.8)	283(0.6)	11056(4.3)	1573(1.0)	4439(6.7)
Metastatic disease	30325(3.3)	6598(2.4)	758(1.3)	1704(1.4)	1870(3.7)	5635(2.2)	12274(7.8)	3057(4.4)
Haematologic malignancy	18188(2.0)	9763(3.6)	402(0.7)	791(0.7)	373(0.8)	5616(2.2)	1203(0.8)	673(1.0)

Immunocompromised	69851(7.5)	24035(8.8)	2086(3.5)	3427(2.8)	2756(5.6)	15471(6.0)	21226(13.4)	4034(5.8)
Organ failure								
Cardiac	268672(28.8)	75565(27.7)	19138(31.0)	28703(23.6)	15420(31.0)	92916(36.3)	37687(23.8)	14118(20.4)
Respiratory	271652(29.1)	120432(44.1)	12983(21.7)	31893(26.2)	8286(16.7)	75866(29.6)	24514(15.5)	13507(19.5)
Renal	128303(13.8)	51727(19.0)	4911(8.2)	8935(7.3)	1515(3.1)	54130(3.8)	6173(3.9)	5938(8.6)
Haematological	34556(3.7)	15508(5.7)	1219(2.0)	2769(2.3)	645(1.3)	12067(4.7)	1671(1.1)	2299(3.3)
Neurological	36996(4.0)	8944(3.3)	609(1.0)	6838(5.6)	2289(4.6)	18621(7.3)	465(0.3)	741(1.1)
#IMD quintile								
Quintile I	224182(24.0)	69728(25.5)	13211(22.1)	32990(27.1)	12429(25.0)	65762(25.7)	28955(18.3)	14638(21.1)
Quintile II	197276(21.1)	58496(21.4)	12230(20.4)	26030(21.4)	10568(21.3)	55775(21.8)	31257(19.8)	14269(20.6)
Quintile III	184984(19.8)	53199(19.5)	12441(20.8)	23256(19.1)	9760(19.6)	49750(19.4)	33092(20.9)	14319(20.7)
Quintile IV	167939(18.0)	47306(17.3)	11524(19.3)	20318(16.7)	8967(18.1)	43820(17.1)	32386(20.5)	13325(19.2)
Quintile V	150502(16.1)	42400(15.5)	10057(16.8)	17120(14.1)	7448(15.0)	38428(15.0)	31273(19.8)	12345(17.8)
Missing	8401(0.9)	1872(0.7)	381(0.6)	2168(1.8)	512(1.0)	2557(1.0)	1101(0.7)	384(0.6)
ICU mortality n, (%)	136799(14.7)	62277(22.8)	4530(7.5)	10703(8.8)	3010(6.0)	56391(22.0)	1529(1.0)	4532(6.5)
Hospital mortality n, (%)	197917(21.2)	86728(31.8)	7053(11.8)	16576(13.6)	5286(10.6)	78710(30.7)	4657(3.0)	8798(12.7)
ICU LOS hrs, [95%CI]	109[109,109]	163[162,164]	84[84,84]	109[108,110]	95[94,97]	108[107,108]	56[55,56]	80[79,81]
Hospital LOS, days[95%CI]	18[18,18]	23[23,23]	16[16,16]	19[19,19]	20[20,20]	16[16,16]	14[14,15]	21[20,21]

60 patients have missing data so unclassified

51140 patients have more than one primary diagnosis group.

LOS= length of stay, IMD= Index of Multiple Deprivation; CI= confidence interval

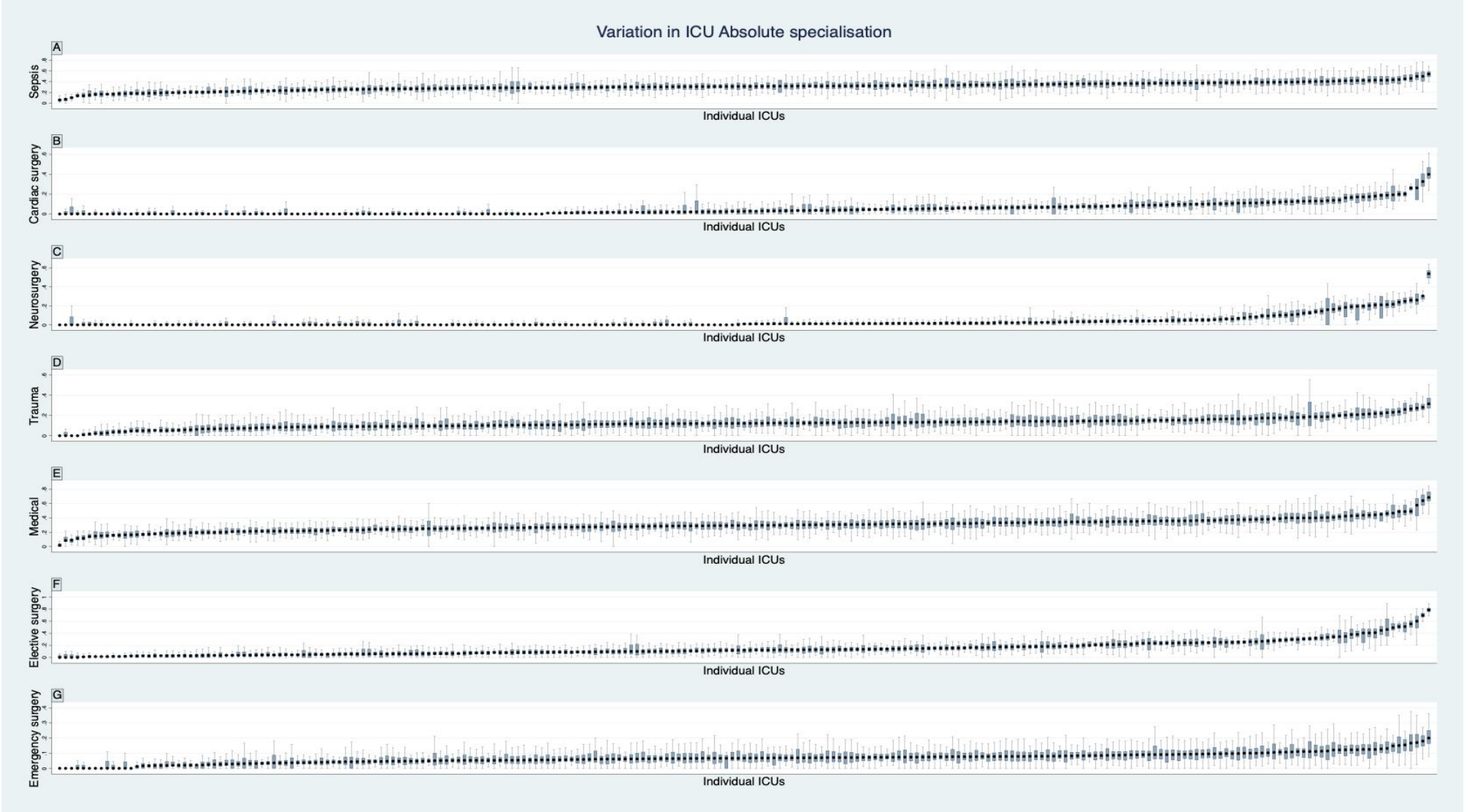
Sepsis specialised ICUs operate at lower total caseload volume (**Error! Reference source not found.**). Increasing sepsis specialisation was associated with fewer ICU beds, The highest quartile of sepsis specialization also had the lowest number of ICU beds and lower throughput (**Error! Reference source not found.**). ICUs with higher shares of sepsis patients were more likely to be non-university affiliated (p=0.012). In contrast ICUs with increasing cardiac, neurosurgical, trauma and elective surgery specialisation operated at higher volumes and had more ICU beds.

Table 16 Distribution of ICU characteristics across quartiles of specialisation. (Mean and [95% CI] unless otherwise stated.)

Variable	Monthly total volume	Throughput	Occupancy	ICU beds	Academic affiliation		
					Non-university	University affiliated	University
Total	80[80,80]	5.2[5.2,5.2]	0.73[0.73,0.73]	16[16,16]	122(52.8)	39(16.9)	70(30.3)
Sepsis							
Quartile I (0-0.23)	109[108,109]	5.6[5.6,5.6]	0.73[0.73,0.73]	21[21,21]	19(15.6)	6(15.4)	22(31.4)
Quartile II(0.23-0.29)	81[81,81]	5.3[5.3,5.3]	0.73[0.73,0.73]	16[16,16]	31(25.4)	8(20.5)	19(27.1)
Quartile III(0.30-0.37)	70[70,70]	5.2[5.2,5.2]	0.72[0.72,0.72]	14[14,14]	42(34.4)	8(20.5)	12(17.1)
Quartile IV (0.37-0.85)	60[60,60]	4.8[4.8,4.8]	0.72[0.72,0.72]	13[13,13]	30(25.5)	17(43.6)	17(24.3)
P value	<0.001	<0.001	<0.001	<0.001			0.012
Cardiac							
Quartile I(0-0.01)	56[56,56]	5.1[5.1,5.1]	0.70[0.70,0.70]	11[11,11]	41(33.6)	13(33.3)	21(30.0)
Quartile II(0.01-0.05)	85[85,85]	5.1[5.1,5.2]	0.73[0.73,0.73]	17[17,17]	37(30.3)	13(33.3)	17(24.2)
Quartile III(0.05-0.09)	95[95,96]	5.2[5.2,5.2]	0.74[0.74,0.74]	19[19,19]	26(21.3)	8(20.5)	14(20.0)
Quartile IV (0.09-0.61)	85[84,85]	5.4[5.4,5.5]	0.73[0.73,0.73]	16[16,16]	18(14.8)	5(12.8)	18(25.7)
P value	<0.001	<0.001	<0.001	<0.001			0.577
Neurosurgery							
Quartile I(0,0)	53[53,53]	5.1[5.1,5.1]	0.70[0.70,0.70]	11[11,11]	66[54.1]	17(43.5)	23[32.9]
Quartile II(0.01,0.02)	78[78,78]	5.5[5.5,5.5]	0.73[0.73,0.73]	15[15,15]	20[16.3]	4(10.2)	11[15.7]
Quartile III(0.02,0.06)	82[82,83]	5.4[5.4,5.4]	0.72[0.72,0.72]	15[15,15]	26[21.3]	10(25.6)	14[20.0]
Quartile IV (0.06,0.63)	119[118,119]	5.1[5.1,5.1]	0.77[0.77,0.77]	24[24,24]	10[8.2]	8(20.5)	22[31.4]
P value	<0.001	<0.001	<0.001	<0.001			0.003
Trauma							
Quartile I(0,0.07)	70[70,70]	5.3[5.3,5.3]	0.71[0.71,0.71]	14[14,14]	26(21.3)	9(23.1)	28(40.0)
Quartile II(0.08, 0.13)	77[77,77]	5.2[5.2,5.2]	0.72[0.72,0.72]	15[15,15]	32(26.2)	9(23.1)	13(18.6)
Quartile III(0.13,0.17)	87[87,88]	5.2[5.2,5.2]	0.74[0.74,0.74]	17[17,17]	30(24.6)	9(23.1)	13(18.6)
Quartile IV (0.17,0.71)	86[86,86]	5.1[5.0,5.1]	0.74[0.74,0.74]	18[18,18]	34(27.9)	12(30.8)	16(22.9)
P value	<0.001	<0.001	<0.001	<0.001			0.202
Medicine							
Quartile I(0,0.21)	101[101,102]	5.5[5.5,5.5]	0.73[0.73,0.73]	19[19,19]	17(13.9)	10(25.6)	20(28.6)
Quartile II(0.21,0.27)	85[85,85]	5.3[5.3,5.3]	0.73[0.73,0.73]	16[16,16]	22(18.0)	8(20.5)	11(15.7)
Quartile III(0.27,0.34)	72[72,72]	5.1[5.1,5.1]	0.73[0.73,0.73]	15[15,15]	36(29.5)	11(28.2)	20(28.6)
Quartile IV (0.34,0.84)	62[62,62]	4.9[4.9,4.9]	0.72[0.72,0.72]	13[13,13]	47(38.5)	10(25.6)	19(27.1)
P value	<0.001	<0.001	<0.001	<0.001			0.224
Elective surgery							
Quartile I (0,0.06)	68[68,68]	4.7[4.7,4.7]	0.73[0.73,0.73]	15[15,15]	31(25.4)	9(23.1)	26(37.1)
Quartile II (0.06,0.13)	70[70,70]	5.1[5.1,5.1]	0.73[0.73,0.73]	14[14,14]	34(27.9)	17(43.6)	10(14.2)
Quartile III (0.13,0.22)	83[83,83]	5.4[5.4,5.4]	0.73[0.73,0.73]	16[16,16]	36(30.0)	6(15.4)	9(12.9)
Quartile IV (0.22,0.89)	100[99,100]	5.6[5.6,5.6]	0.72[0.72,0.72]	19[19,19]	21(17.2)	7(18.0)	25(35.7)
P value	<0.001	<0.001	<0.001	<0.001			<0.001
Elective surgery							
Quartile I (0,0.03)	78[78,78]	5.0[5.0,5.0]	0.72[0.72,0.72]	16[16,16]	28[23.0]	9[23.0]	30[42.9]
Quartile II (0.3,0.06)	87[87,87]	5.2[5.2,5.2]	0.73[0.73,0.73]	17[17,17]	25[20.5]	11[28.2]	21[30.0]
Quartile III (0.06,0.10)	85[84,85]	5.3[5.3,5.3]	0.73[0.73,0.73]	16[16,17]	31[25.4]	9[23.1]	10[14.3]
Quartile IV (0.10, 0.46)	71[71,71]	5.3[5.3,5.3]	0.72[0.72,0.72]	14[14,14]	38[31.2]	10[25.6]	9[12.9]
P value	<0.001	<0.001	<0.001	<0.001			0.009

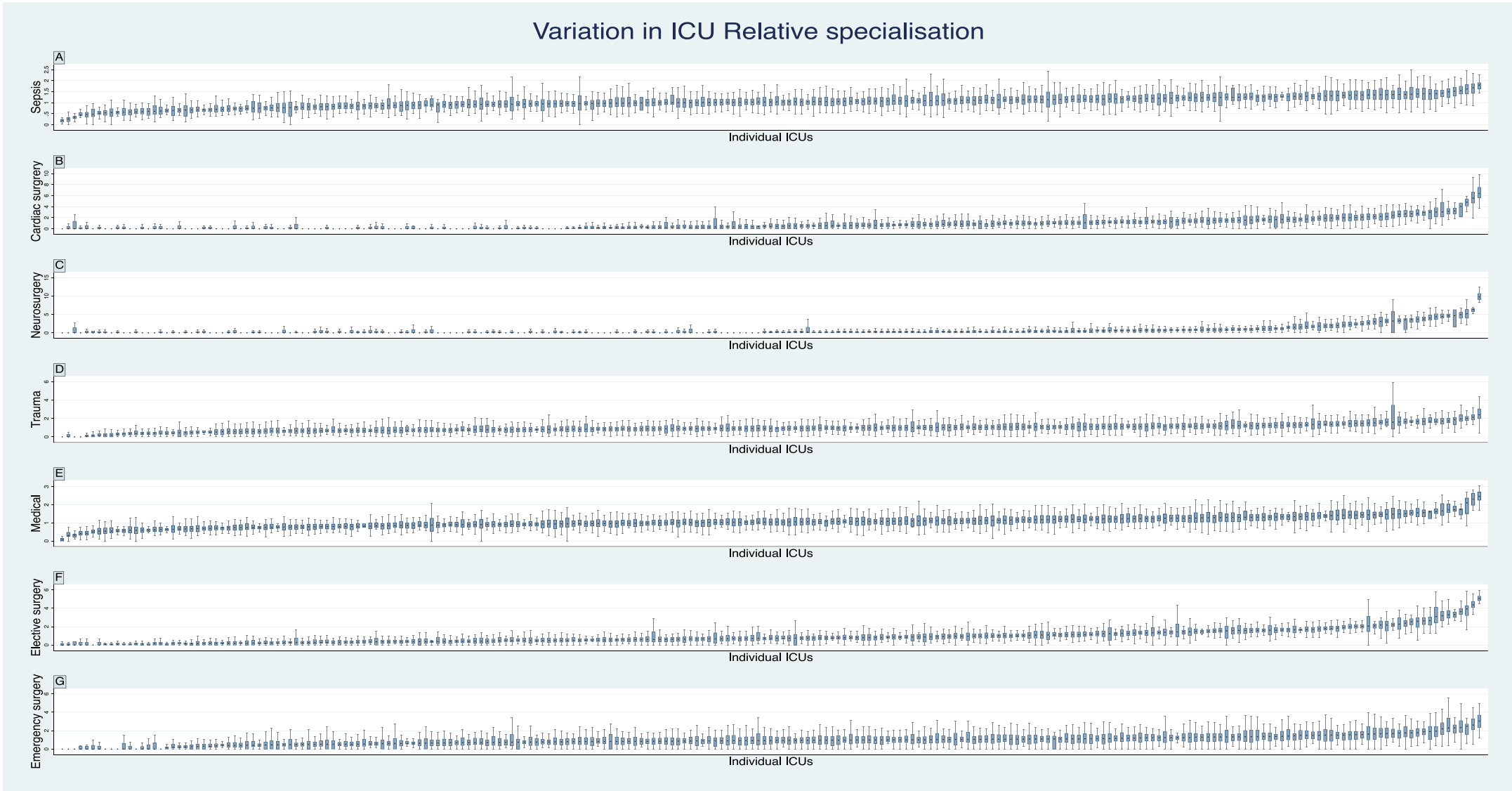
After controlling for covariates, an increase in sepsis specialisation was associated with higher patient cute hospital mortality. An absolute and relative increase in sepsis specialisation by 10% was associated with a 6% (OR 1.06 95% CI 1.04-1.08, p<0.001) and a 1% (OR 1.01, 95%CI 1.00-1.02, p=0.003) increase in acute hospital mortality respectively. There were no mortality benefits for patients treated in specialist cardiac and neurosurgical ICUs (**Error! Reference source not found.**).

Figure 17. Variation in absolute specialisation across Intensive Care Units



A. Absolute specialisation in sepsis patients. B. Absolute specialisation in cardiac surgery patients. C. Absolute specialisation in neurosurgery patients. D Absolute specialisation in trauma patients. E. Absolute specialisation in medical patients. F. Absolute specialisation in elective surgery patients. G. Absolute specialisation in emergency surgery patients. Navy diamond indicates median, blue rectangle indicate interquartile range and black line indicates full range

Figure 18.Variation in relative specialisation across Intensive Care Unit



A. Relative specialisation in sepsis patients. B. Relative specialisation in cardiac surgery patients. C. Relative specialisation in neurosurgery patients. D. Relative specialisation in trauma patients. E. Relative specialisation in medical patients. F. Relative specialisation in elective surgery patients. G. Relative x specialisation in emergency surgery patients. Navy diamond indicates median, blue rectangle indicate interquartile range and black line indicates full range.

Table 17. Results of regression showing association between specialisation measures and acute hospital mortality

	Absolute			Relative		
	OR	**95%CI**	**P value**	**OR**	**95%CI**	**P value**
Specialization measures						
• Sepsis	1.06***	1.04, 1.08	<0.001	1.01**	1.00, 1.02	0.003
• Cardiac	1.01	0.99, 1.04	0.292	1.00	1.00, 1.00	0.917
• Neurosurgery	1.01	0.98, 1.04	590	1.00	1.00, 1.00	0.557
• Trauma	1.03**	1.01, 1.05	0.010	1.00**	1.00, 1.01	0.009
• Medical	1.03**	1.01, 1.06	0.007	1.01	1.00, 1.01	0.055
• Elective surgery	0.96**	0.94, 0.99	0.002	0.99***	0.99, 1.00	<0.001
• Emergency surgery	1.00	0.97, 1.03	0.867	1.00	1.00, 1.00	0.941
Monthly total volume	0.97***	0.96, 0.98	<0.001	0.96***	0.95, 0.97	<0.001
Functional state						
• Independent(reference)	1.00			1.00		
• Some assistance	1.12***	1.10, 1.13	<0.001	1.12***	1.10, 1.13	<0.001
• Total assistance	1.53***	1.43, 1.63	<0.001	1.53***	1.43, 1.63	<0.001
Ethnicity						
• White	1.00			1.00		
• Asian	0.95**	0.91, 0.98	0.006	0.94**	0.91, 0.98	0.005
• Black	0.83***	0.78, 0.87	<0.001	0.83***	0.78, 0.87	<0.001
• Mixed/other	0.99	0.96, 1.03	0.732	0.99	0.96, 1.03	0.720
Age in year	1.01***	1.01, 1.01	<0.001	1.01***	1.01, 1.01	<0.001
Male sex	1.11***	1.10, 1.13	<0.001	1.11***	1.10, 1.13	<0.001
ICNARC score	1.06***	1.06, 1.06	<0.001	1.06***	1.06, 1.06	<0.001
Co-morbidities						
• Severe respiratory disease	1.13***	1.09, 1.17	<0.001	1.13***	1.09, 1.17	<0.001
• Severe cardiovascular disease	1.30***	1.24, 1.36	<0.001	1.30***	1.24, 1.36	<0.001
• End-stage renal disease	1.25***	1.20, 1.31	<0.001	1.25***	1.20, 1.31	<0.001
• Severe liver disease	1.37***	1.32, 1.42	<0.001	1.37***	1.32, 1.42	<0.001
• Metastatic disease	1.20***	1.15, 1.24	<0.001	1.20***	1.15, 1.24	<0.001
• Haematological malignancy	1.17***	1.12, 1.23	<0.001	1.17***	1.12, 1.23	<0.001
• Immunocompromised	1.07***	1.04, 1.10	<0.001	1.07***	1.04, 1.10	<0.001
Organ failure						
• Shock	1.02*	1.00, 1.04	0.011	1.02*	1.00, 1.04	0.014
• Respiratory failure	1.20***	1.18, 1.22	<0.001	1.20***	1.18, 1.22	<0.001
• Acute renal failure	1.18***	1.16, 1.20	<0.001	1.18***	1.16, 1.20	<0.001
• Haematological failure	1.26***	1.22, 1.30	<0.001	1.26***	1.22, 1.30	<0.001
• Neurological failure	1.04*	1.00, 1.07	0.041	1.04*	1.00, 1.07	0.047

IMD	1.00***	1.00, 1.00	<0.001	1.00***	1.00, 1.00	<0.001
Monthly throughput	1.00	1.00, 1.01	0.344	1.00	0.99, 1.01	0.383
Academic affiliation						
• Non-university	1.00			1.00		
• University affiliated	0.99	0.91, 1.07	0.756	0.99	0.91, 1.08	0.822
• University	1.04	0.97, 1.12	0.251	1.04	0.97, 1.12	0.280

Each 1 -unit change is equivalent to a 10% change in absolute or relative share: OR= odds ratio, CI= confidence interval, IMD= Index of Multiple Deprivation.
P<0.05=*. P<0.01=**P<0.001=***

7.5.1. Sub-group analysis

The first subgroup analysis was of diagnosis specific patients treated in the diagnosis appropriate ICU. This would mean that sepsis patients would derive the greatest benefit from specialist sepsis ICUs, cardiac patients from specialist cardiac ICUs, neurosurgical patients from specialist neurosurgical ICUs. The subgroup analysis was consistent with the primary results in that no specific disease specific patient derived any benefits from the "ideal" ICU specialisation, both in absolute and relative measures of specialisation (**Error! Reference source not found.**). Sepsis patients did not benefit from specialised sepsis ICUs, cardiac patients dd not benefit from specialised cardiac ICUs, neurosurgical patients did not benefit from specialist neurosurgical ICUs. The same was true for trauma, medicine, elective, and emergency surgery patients. The second subgroup analysis was of the highest quartile of illness severity. This analysis did not find any benefit from ICU specialisation. The results are described in the supplementary Appendix for chapter seven.

7.5.2. Robustness checks

We used alternate specifications of specialisation to assess the robustness of our analysis. Using alternate measures of specialisation did not alter the results from the primary analysis. The details are described in the Appendix three.

7.5.3. Sensitivity analysis

We undertook three sensitivity analyses. Firstly, the addition of time fixed effects did not change the primary results and is included in the supplement. Secondly, ICU specialisation was also not associated with lower ICU mortality, consistent with the primary analysis using hospital mortality. We included a subgroup analysis of the highest quartile of severity of illness treated in the ideal speciality ICU. There was no association with lower hospital mortality in all groups of patients except for the most severely ill trauma ICU patients treated in trauma specialised ICUs. Lastly, we included an analysis of the effects of specialist ICUs on non-specialist patients. We did not identify any favourable spill over effects of specialisation. These results are detailed in Appendix three.

Table 18 . Subgroup analysis that examines the relationship between acute hospital mortality and "ideal" ICU specialisation.

Subgroup	N	Ideal Absolute specialisation			Ideal Relative specialisation		
		OR	95%CI	P value	OR	95%CI	P value
Sepsis	258,422	1.02	99,1.06	0.112	0.99	0.98, 1.00	0.104
Cardiac	57,268	0.95	0.86,1.04	0.297	1.00	0.99, 1.00	0.138
Neurosurgery	45,621	1.08	0.97, 1.21	0.166	1.00	1.00, 1.01	0.073
Trauma	113,121	1.04	0.98,1.11	0.180	1.00	0.99,1.01	0.746
Medical	228,770	1.00	0.96,1.03	0.810	1.00	0.99,1.01	0.913
Elective surgery	152,905	1.09	0.99,1.20	0.074	1.01	1.00,1.02	0.091
Emergency surgery	66,615	0.94	0.85, 1.04	0.196	1.00	0.99,1.00	0.329

P<0.05=*. P<0.01=**P<0.001=***

This analysis examines the relationship between acute hospital mortality and "ideal" ICU specialisation. i.e., acute hospital mortality for sepsis patients treated in the sepsis specialist ICU, patients with cardiac disease treated in the cardiac specialised ICU, patients with neurosurgical disease treated in the neurosurgery specialised ICU, patients with medical disease treated in a medicine specialised ICU, and patients requiring elective or emergency surgery treated in the respective "ideal" specialised ICU. This analysis did not identify any other benefit to ICU specialisation.

7.6. Discussion

In this representative study of adult ICUs in the UK, ICU specialisation was not associated with improved outcomes for patients. These results are robust to alternate specifications of specialisation, as well as across subgroups of patients and control for caseload volume and severity of illness.

Our findings are consistent with previous studies that did not show benefits from disease specific ICU specialisation and support the idea that critically ill patients require similar treatment regardless of underlying disease. [47, 50, 247]. While most of the preceding literature largely focused on small, single cohorts of surgical patients, our study adds to the exiting literature by including a nationally representative cohort of critically ill patients treated in a publicly funded health system. Additionally, our study included a diverse group of patients including a substantial proportion of non-surgical patients, making the results more generalisable. These results have important implications for the organization of critical care services. In countries like the UK and the US where there may be an interest in expanding critical care services[196]. Under the value-based paradigm, the goals of improving patient outcomes are unlikely to be advanced by creating specialist ICUs after controlling for volume.

Our study has several limitations. We do not describe the underlying mechanism by which concentrating high complexity patients operates. It may be because of the available resources, technologies, and staffing. We would expect that these would be correlated with the various measures of specialisation used. The absence of these covariates would bias the results in favour of specialization. The failure to show reductions in mortality by omitting these variables reinforces our conclusions that specialist ICUs do not offer any mortality benefit. There is potential confounding by omitted covariates and misspecification of the

exposure. To account for these, we undertook detailed covariate adjustment and several robustness checks. The results remain consistent across these analyses.

Our study has several strengths. This study includes a large proportion of patients with a diverse set of diagnoses and expands on this literature that has thus far focused on surgical diagnoses. We include a nationally representative sample of critical care services in the UK provided in publicly owned hospitals in the National Health Service. We also use granular clinical data to undertake risk adjustment and create several measures of ICU specialisation in the absence of a gold standard. Lastly, we add to the emerging literature on service-level specialisation in contrast to the hospital level.

7.7. Conclusion

The ICU is a complex system and there are relatively few studies exploring the optimal organisation of critical care services. As critical care services grow, policy makers and hospital administrators are faced with the choice of building more general or specialist ICUs. This study finds that patients treated in specialist ICUs do not have lower hospital mortality after controlling for caseload volume.

Chapter eight: Conclusion

This thesis contributes to the evidence on the effectiveness of policies to reorganize services to improve patient outcomes. Since its inception, the ICU has been considered an organisation intervention, where the sickest patients are collocated to benefit from the expertise of the team[248, 249]. Over time our understanding of the organisation of critical care services has evolved helping us define the modern critical care. The COVID-19 pandemic has brought home the real-world consequences to the impacts of ICU organisation. Media reports describing survival depending on the availability of ICU beds and the interwoven narratives of staff and equipment shortages and the wide variation in hospital, regional, and country level variation in survival amongst countries with broadly similar investments in health, point to potential inefficiencies in ICU organisation [250] [251].

The thesis proposes alternate organisational models to meet these challenges. Chapter one introduces the volume-outcome relationship, and the potential underlying mechanism namely learning-by-doing compared with static scale economies. The thesis then introduces specialisation as a related model of care. Chapter two contextualises critical care services within the hospital and provides some international context for the generalisability of this work. Chapter three provides a detailed description of the ICNARC CMPD clinical dataset. Chapter four proceeds to systematically assess the current evidence base for the volume-outcome relationship and its methodological limitations.

Chapter five investigates the volume-outcome relationship in a cohort of patients with sepsis. The study found a significant association between the sepsis case volume in an ICU and

hospital mortality from sepsis. This association was consistent across the categorical and nonlinear specifications of ICU volume. The study also identified a lower volume threshold of 215 patients treated per year, above which there was a statistically significant reduction in mortality. There was no significant interaction between case volume and severity of illness. There was also no significant difference in the volume outcome relationship with surgical compared with medial patients with sepsis. The study found that significant ICU practice variation was not explained by patient or hospital characteristics, implying that sample selection was not distorting the associations described. The within-ICU variation remained unchanged across years, suggesting that higher-performing ICUs maintained good performance over time.

Chapter six explores the underlying mechanism by which the volume-outcome relationship operates. The two competing potential mechanisms are the static effects of economies of scale and the dynamic effects of learning-by-doing. Distinguishing between these mechanisms is important to inform policy decisions about centralisation of services. The study found dynamic learning-by-doing effects to be more important than the static scale effect. This finding suggests that patients would be better served by organising ICUs to achieve minimum volume standards previously identified without centralisation of services. ICU teams are temporally unstable, and the study examines the learning-by-doing effect by considering two time intervals, quarterly and monthly, rather than the annual intervals used in most empirical studies. The study suggests that there may be some depreciation of learning, supporting the idea of keeping teams together to retain team learning and memory.

Chapter seven explores the gains from specialisation in terms of acute hospital mortality in the ICU across several diagnostic groups of patients covering several years. The study explores the primary hypothesis that increasing specialisation in a particular diagnostic group improves outcomes for all patients in that ICU. The study did not find a consistent benefit from specialisation across seven diagnostic groups of patients for patients in those ICUs after controlling for volume and disease severity. A secondary hypothesis is that increasing specialisation in a focal segment improves outcomes for patients within that speciality. The study did not find benefits of specialisation for patients within the diagnosis-specific specialisation. Lastly, the study explored the hypothesises that increasing ICU specialisation will benefit the sickest patients within that speciality. The only group of patients for which there may be some benefit is the highest risk trauma patients. For all the other diagnostic groups, no benefit to specialisation was identified. A possible explanation for these findings is that ICU patients often require treatments that converge across diagnostic groups. Narrowing the focus of ICUs along diagnostic groups would not be advantageous. Therefore, whilst high volume ICUs reduce mortality, narrowing the focus of clinical care through specialisation is unlikely to confer additional benefit to critically ill patients.

8.1 Discussion of the findings and implications for research

Chapter five addresses some of the limitations of the current literature on the volume outcome relationship in critical illness. First, a major limitation of the existing literature is the absence of a standard for defining volume. The most common approach has been to divide the volume into quartiles, however, examining quartiles does not improve the general understanding of the association between volume and mortality because quartiles are

specific to the dataset from which they are derived. An ICU could fall within a high-volume quartile in one study but the low-volume quartile in another[163]. To address this, we used restricted cubic splines that allowed flexibility in fitting the regression models. This approach also allows us to detect the optimal volume thresholds. Second, many previous studies included ICUs with a narrow set of volumes., making them underpowered to detect a small but significant volume outcome relationship. Instead, this study uses a nationally representative dataset that included a wide spectrum of ICU volumes, overcoming this limitation. Third, most studies use secondary administrative data collected for other uses. Such data have inherent limitations in both the identification of sepsis and the clinical characteristics of patients and ICUs. Our study used the Case-Mix Program Database which is a clinical database of patients admitted to all general ICUs in the UK. This rich dataset allowed us to identify patients with sepsis by applying the international consensus definition as well as to perform detailed risk adjustment and identify ICU-specific characteristics. In terms of completeness and representativeness, this is one of the largest studies to evaluate the volume-outcome relationship for sepsis. The study included all adult sepsis ICU admissions in England, Wales, and Northern Ireland.

The benefit of using a cohort of sepsis patients is that ICUs in the UK are unable to make risk-based selection of patients with a low risk of mortality because patients with sepsis are taken to their nearest hospital. The study evaluated the potential effects of unmeasured confounders using E-values, adding further evidence to the robustness of the primary analysis[188].

Chapter six explores the underlying mechanism by which volume affects outcome. Much of the preceding literature has focused on addressing the issue of endogenous selection through selective referral in the context of elective surgery. Selective referral is not a substantial threat

to identification in the setting of sepsis because these patients do not choose their ICU. Sepsis is a life-threatening illness that requires time critical intervention and patients are usually brought to the nearest hospital via ambulance without the opportunity to select their ICU. There is considerably less literature addressing the question of whether it the accumulation of knowledge by experience, namely learning-by -doing, from the static economies of scale effects. Many studies used fixed effects in a small number of centres to control for time invariant centre-level heterogeneity[18]. This approach is likely to be underpowered in detecting a learning by doing effect. It is also likely that within a small number of centres the annual volumes are likely to be highly correlated introducing the problem of multicollinearity. Many learning-by-doing studies consider annual lags but these may not fully capture institutional learning because of the turnover of staff within this period. We contribute to the literature by considering quarterly and monthly volumes from 231 ICUs and used a hierarchical model to account for the clustered nature of the data.

Chapter seven compares the effects of ICU specialisation in a diverse range of diagnostic groups in terms of acute hospital mortality. Whilst the idea of competitive advantage through a focused factory over a more complex factory has been described in industries outside of healthcare, the gains from specialist service lines in healthcare are less clear[3]. Most of the current literature is centred around small for-profit physician-owned specialist hospitals in the US hospital sector. These studies are subject to endogenous selection through cherry picking because these hospitals may be incentivised to treat healthier and wealthier patients[240]. This thesis addresses these concerns by making use of a nationally representative clinical dataset in a publicly funded health system. The study finds no consistent reduction in mortality from ICU specialisation after controlling for volume. Whilst there is no consensus on the best measure of specialisation, the share of patients has

frequently been used to describe levels of specialisation[42, 232, 252]. This study employs several alternate specifications of specialisation and undertook a series of sensitivity analyses, all of which support the primary analysis.

The study has limitations. We do not control for the available resources or technology, assuming that these may be correlated with the level of specialisation. The omission of these variables would bias the results in favour of specialisation. The absence of such a specialisation benefit therefore is further support for the primary analysis.

8.2. Implications for policy

Variation in the quality of care between ICUs and workforce shortages have been highlighted by the Covid-19 pandemic, necessitating system level changes in the organisation of the critical care services[253]. This study uses a large cohort of patients with well-defined critical illness syndromes to explore questions about ICU organisation and outcomes. In terms of policy implications, chapter five identified a volume-outcome relationship for patients with sepsis. A lower volume threshold at which mortality improved was identified. Volume-outcome relationships for high risk elective surgery has previously been described and policies recommending minimum volume standards have been implemented in many countries [254] [8, 17, 255]. Chapter five argues the case in favour of minimum volume standards for sepsis to reduce mortality. For example, more than 38.3% of patients with sepsis were treated in ICUs that operated below the threshold of 215 patients per year and more than 72% of ICUs operate below this threshold. Implementing this minimum volume standard would involve significant service reconfiguration.

Centralisation of care is another potential strategy to improve care. Concentrating expertise and resources in a few high volume centres has been proposed in trauma, stroke, neonatal

care as well as high-risk surgery[256] [257-259]. Sepsis is an acute condition for which aggregating patients in a small number of centres has been proposed[205, 260]. There may also be cost savings by more efficient use of resources. This process would require creating tiers defined by explicit triage criteria, professional competencies, hospital accreditation and outcome surveillance[260]. There are concerns that centralising sepsis patients may lead to worse outcomes. Concentrating patients into fewer ICUs would lead to longer travel times to access care for a sepsis, which is a time sensitive condition. There may be and strain of resources at the receiving ICU because most large ICUs already operate at high occupancy and there may be declining marginal productivity. There is also the possibility of increasing care fragmentation with patients receiving some care at their local hospital, particularly rehabilitation care after critical illness and increased travel costs for families. If the static scale economies were the dominant mechanism, then it would justify the investments in centralising sepsis patients beyond minimum volume standards.

Chapter six finds that learning-by-doing is the dominant mechanism through which the volume-outcome relationship operates and therefore quality is improved through experience. This argues in favour of keeping teams together to retain institutional learning. Shifting patients from one ICU to another would reduce opportunities for learning in the transferring ICU which is losing volume. The balance between the benefits and harms of centralising patients can be struck by maintaining minimum volume standards for sepsis but also not centralising sepsis services beyond achieving these standards. This study addresses a key knowledge gap in our understanding of the volume-outcome relationship in critical care.

Chapter seven looks at the value of specialist service lines within critical care. The study finds no consistent benefit from specialisation, instead argues for the potential for knowledge spill

overs between different service lines within the ICU. Knowledge gained from caring for one group of patients appears to improve care for other groups of patients in the ICU. Critically ill patients require the management of specific organ failures, regardless of the primary diagnosis. These treatments include mechanical ventilation, renal replacement therapy and the management of shock and rehabilitation. As an example, patients with sepsis and patients with trauma would require similar processes of care in managing lung injury or renal failure. The optimal scope of patients would therefore seem to extend beyond single service lines, although this not precisely estimated in this study. A broader range of activities seems to increase the performance of the ICU. The relatedness of the diversification might be relevant in explaining this observation. These studies provide evidence for policy makers to make important decisions about the future organisation of critical care services.

8.3 Further research

This thesis examines policy relevant issues regarding the reorganisation of critical are services in the UK. However, several research and policy relevant questions remain unanswered.

First, although this thesis includes some details about institutional characteristic's such as size and academic affiliation it does not explore the how to best use the available critical care staff to maximise returns for patients. In a recent workforce report, the Faculty of Intensive Care Medicine cited a chronic under-provision of sufficient numbers of critical care doctors in the UK and suggested a reconfiguration of both staff and beds was required[261]. The data included in this thesis does not describe any workforce characteristics so does not address any policy related to workforce reorganisation.

Second, more empirical work is needed in understanding the longer-term implications of these policies of centralisation, minimum volume standard and specialisation. This thesis uses acute hospital mortality and tracks patients between hospitals until the until end of their acute hospitalisation. However, acute mortality probably underestimates the true burden of sepsis and critical illness generally. Patients with critical illness are more likely to be re-hospitalised in the weeks and months following discharge. The average rate of 30-day rehospitalization for sepsis survivors is between 19% and 32%[262]. This increases to about 63% by 1 year[262]. Sepsis is associated with cognitive and functional decline and an increased risk of cardiovascular events such as coronary artery disease, stroke acute myocardial infarction and for up to ten years after a diagnosis of sepsis[263, 264]. A comprehensive national dataset of patients with critical illness that includes longer term mortality and the longer term sequalae of critical illness would be helpful in appreciating the full benefits of any policies aimed at improving the quality of critical care and would inform any subsequent economic evaluation.

Third, this thesis is focused on the quality dimensions of critical care service reorganisation and does provide any economic evaluation related to these policies including the increased costs to access care for patients and their carers. Patients may have to travel further to access providers if low volume local providers are aggregated into higher volume providers. This increased burden will impact patients in lower sociodemographic groups. Therefore, the impact of policies such as minimum volume standards on equity of access needs to be considered. Most datasets to not contain sufficient information about unmet need and barriers to health access. Data from stroke and heart attack centres suggest that minimum volume standards could increase inequity for minority communities and widen disparities for non-urban populations[265, 266]. Lastly, critical care services are central to delivering a wide

range of acute services and the wider implications of critical care restructuring with regards to health access needs further study. For example, relocating critical care services from a hospital would result in that hospital being unable to provide high-risk surgery. Therefore, any critical care organisation needs to cognizant of the wider system implications.

Going beyond the scope of this thesis, undoubtedly the evolution of technology will likely have a significant role in addressing workforce and ICU bed supply shortages. Critical care services could be delivered through ICU telemedicine and regional outreach programs that go beyond the borders of the ICU. Telemedicine in critical care refers to the exchange of medical information typically between ICU doctors in a high-volume centralised service and a remote provider. Telemedicine facilitates the delivery of healthcare from experienced providers to remote locations. Regional outreach programs involve collaboration, benchmarking and sharing experience and quality improvement resources within a regional network. This may lead to improved outcomes by some network effect either, through knowledge spill overs between critical care services or quality improvement through benchmarking.

In summary the policy and research recommendations from this work are as follows:

- Patients with sepsis have a higher chance of survival if treated in ICUs with an annual caseload greater than 215 patients. The benefits of being treated in a high-volume ICU was not related to the severity of illness.
- About 39% of sepsis patients were treated in ICUs below this threshold annual volume. More than 72% of ICUs operated at volumes below this thresholds volume. Implementing minimum volume standards for patients with sepsis in the UK would therefore require significant structural changes.

- Learning-by-doing is the dominant mechanism through which the volume-outcome relationship operates for patients with sepsis. This argues for polices that keep clinical teams together for as long as possible. Beyond meeting minimum volume standards, further centralisation of services may not fully leverage the learning-by-doing mechanism.
- Demand for critical care services continue to grow. Policy makers and hospital administrators are faced with the choice expanding general or specialist critical care. The organisation of critical care services into specialist ICUs is unlikely to improve mortality.

References

1. Luft HS, Hunt SS, Maerki SC: **The volume-outcome relationship: practice-makes-perfect or selective-referral patterns?** *Health Serv Res* 1987, **22**(2):157-182.
2. Birkmeyer JD, Skinner JS, Wennberg DE: **Will volume-based referral strategies reduce costs or just save lives?** *Health Aff (Millwood)* 2002, **21**(5):234-241.
3. Christian CK, Gustafson ML, Betensky RA, Daley J, Zinner MJ: **The volume-outcome relationship: don't believe everything you see**. *World J Surg* 2005, **29**(10):1241-1244.
4. Birkmeyer JD, Siewers AE, Finlayson EV, Stukel TA, Lucas FL, Batista I, Welch HG, Wennberg DE: **Hospital volume and surgical mortality in the United States**. *N Engl J Med* 2002, **346**(15):1128-1137.
5. Morche J, Renner D, Pietsch B, Kaiser L, Bronneke J, Gruber S, Matthias K: **International comparison of minimum volume standards for hospitals**. *Health Policy* 2018, **122**(11):1165-1176.
6. Ho V: **Learning by Doing**. In: *Encyclopedia of Health Economics.* Edited by Culyer AJ. San Diego: Elsevier; 2014: 141-145.
7. Arrow KJ: **The Economic Implications of Learning by Doing**. *The Review of Economic Studies* 1962, **29**(3):155-173.
8. Hall BHRN: **Handbook of the economics of innovation**. Amsterdam; Boston: North Holland; 2010.
9. Huesch MD: **Learning by doing, scale effects, or neither? Cardiac surgeons after residency**. *Health Serv Res* 2009, **44**(6):1960-1982.
10. Gaynor M, Seider H, Vogt WB: **The Volume-Outcome Effect, Scale Economies, and Learning-by-Doing**. *American Economic Review* 2005, **95**(2):243-247.
11. Sfekas A: **Learning, forgetting, and hospital quality: an empirical analysis of cardiac procedures in Maryland and Arizona**. *Health Econ* 2009, **18**(6):697-711.
12. Taylor RA: **Principles of the Economic Behavior of Hospitals**. *Nebraska Journal of Economics and Business* 1972, **11**(2):45-53.
13. McGuire A: **The theory of the hospital: a review of the models**. *Soc Sci Med* 1985, **20**(11):1177-1184.
14. Pauly M, Redisch M: **The Not-For-Profit Hospital as a Physicians' Cooperative**. *The American Economic Review* 1973, **63**(1):87-99.
15. Newhouse JP: **Toward a Theory of Nonprofit Institutions: An Economic Model of a Hospital**. *The American Economic Review* 1970, **60**(1):64-74.
16. Halpern NA, Pastores SM: **Critical Care Medicine Beds, Use, Occupancy, and Costs in the United States: A Methodological Review**. *Crit Care Med* 2015, **43**(11):2452-2459.
17. Wunsch H: **The outbreak that invented intensive care**. *Nature* 2020.
18. Everberg G, Ibsen B: **[Transport of polio patients with respiratory symptoms]**. *Ugeskr Laeger* 1954, **116**(29):1077-1078.
19. Lassen HC: **A preliminary report on the 1952 epidemic of poliomyelitis in Copenhagen with special reference to the treatment of acute respiratory insufficiency**. *Lancet* 1953, **1**(6749):37-41.
20. Takala J: **A History of Intensive Care Medicine**. In: *The Wondrous Story of Anesthesia.* Edited by Eger Ii EI, Saidman LJ, Westhorpe RN. New York, NY: Springer New York; 2014: 785-798.

21. Weil MH, Tang W: **From intensive care to critical care medicine: a historical perspective**. *Am J Respir Crit Care Med* 2011, **183**(11):1451-1453.
22. Rothen HU, Stricker K, Einfalt J, Bauer P, Metnitz PG, Moreno RP, Takala J: **Variability in outcome and resource use in intensive care units**. *Intensive Care Med* 2007, **33**(8):1329-1336.
23. Marshall JC, Bosco L, Adhikari NK, Connolly B, Diaz JV, Dorman T, Fowler RA, Meyfroidt G, Nakagawa S, Pelosi P *et al*: **What is an intensive care unit? A report of the task force of the World Federation of Societies of Intensive and Critical Care Medicine**. *J Crit Care* 2017, **37**:270-276.
24. Wilcox ME, Chong CA, Niven DJ, Rubenfeld GD, Rowan KM, Wunsch H, Fan E: **Do intensivist staffing patterns influence hospital mortality following ICU admission? A systematic review and meta-analyses**. *Crit Care Med* 2013, **41**(10):2253-2274.
25. Angus DC, Shorr AF, White A, Dremsizov TT, Schmitz RJ, Kelley MA, Committee on Manpower for P, Critical Care S: **Critical care delivery in the United States: distribution of services and compliance with Leapfrog recommendations**. *Crit Care Med* 2006, **34**(4):1016-1024.
26. Murthy S, Wunsch H: **Clinical review: International comparisons in critical care - lessons learned**. *Crit Care* 2012, **16**(2):218.
27. Martin JM, Hart GK, Hicks P: **A unique snapshot of intensive care resources in Australia and New Zealand**. *Anaesth Intensive Care* 2010, **38**(1):149-158.
28. Prin M, Wunsch H: **International comparisons of intensive care: informing outcomes and improving standards**. *Curr Opin Crit Care* 2012, **18**(6):700-706.
29. Marshall JC, Bosco L, Adhikari NK, Connolly B, Diaz JV, Dorman T, Fowler RA, Meyfroidt G, Nakagawa S, Pelosi P *et al*: **What is an intensive care unit? A report of the task force of the World Federation of Societies of Intensive and Critical Care Medicine**. *Journal of Critical Care* 2017, **37**:270-276.
30. **Levels of care** [https://www.ics.ac.uk/Society/Guidance/PDFs/Levels_of_Care_25_3_21]
31. **Guidelines for intensive care unit admission, discharge, and triage**. *Critical Care Medicine* 1999, **27**(3):633-638.
32. Ward NS, Afessa B, Kleinpell R, Tisherman S, Ries M, Howell M, Halpern N, Kahn J, Members of Society of Critical Care Medicine Taskforce on ICUS: **Intensivist/patient ratios in closed ICUs: a statement from the Society of Critical Care Medicine Taskforce on ICU Staffing**. *Crit Care Med* 2013, **41**(2):638-645.
33. Gershengorn HB, Harrison DA, Garland A, Wilcox ME, Rowan KM, Wunsch H: **Association of Intensive Care Unit Patient-to-Intensivist Ratios With Hospital Mortality**. *JAMA Intern Med* 2017, **177**(3):388-396.
34. Wunsch H, Angus DC, Harrison DA, Collange O, Fowler R, Hoste EA, de Keizer NF, Kersten A, Linde-Zwirble WT, Sandiumenge A *et al*: **Variation in critical care services across North America and Western Europe**. *Crit Care Med* 2008, **36**(10):2787-2793, e2781-2789.
35. Guidet B, van der Voort PHJ, Csomos A: **Intensive care in 2050: healthcare expenditure**. *Intensive Care Medicine* 2017, **43**(8):1141-1143.
36. Gooch RA, Kahn JM: **ICU bed supply, utilization, and health care spending: an example of demand elasticity**. *JAMA* 2014, **311**(6):567-568.

37. Takala J, Moser A, Raj R, Pettila V, Irincheeva I, Selander T, Kiiski O, Varpula T, Reinikainen M, Jakob SM: **Variation in severity-adjusted resource use and outcome in intensive care units**. *Intensive Care Med* 2022, **48**(1):67-77.
38. Knaus WA, Wagner DP, Zimmerman JE, Draper EA: **Variations in mortality and length of stay in intensive care units**. *Ann Intern Med* 1993, **118**(10):753-761.
39. Churpek MM, Gupta S, Spicer AB, Parker WF, Fahrenbach J, Brenner SK, Leaf DE, Investigators S-C: **Hospital-Level Variation in Death for Critically Ill Patients with COVID-19**. *Am J Respir Crit Care Med* 2021, **204**(403-411).
40. Gillies MA, Power GS, Harrison DA, Fleming A, Cook B, Walsh TS, Pearse RM, Rowan KM: **Regional variation in critical care provision and outcome after high-risk surgery**. *Intensive Care Med* 2015, **41**(10):1809-1816.
41. Kadri SS, Sun J, Lawandi A, Strich JR, Busch LM, Keller M, Babiker A, Yek C, Malik S, Krack J *et al*: **Association Between Caseload Surge and COVID-19 Survival in 558 U.S. Hospitals, March to August 2020**. *Ann Intern Med* 2021, **174**(9):1240-1251.
42. Rapoport J, Teres D, Barnett R, Jacobs P, Shustack A, Lemeshow S, Norris C, Hamilton S: **A comparison of intensive care unit utilization in Alberta and western Massachusetts**. *Crit Care Med* 1995, **23**(8):1336-1346.
43. Bekelman JE, Halpern SD, Blankart CR, Bynum JP, Cohen J, Fowler R, Kaasa S, Kwietniewski L, Melberg HO, Onwuteaka-Philipsen B *et al*: **Comparison of Site of Death, Health Care Utilization, and Hospital Expenditures for Patients Dying With Cancer in 7 Developed Countries**. *JAMA* 2016, **315**(3):272-283.
44. Iwashyna TJ, Cooke CR, Wunsch H, Kahn JM: **Population burden of long-term survivorship after severe sepsis in older Americans**. *J Am Geriatr Soc* 2012, **60**(6):1070-1077.
45. Hatfield KM, Dantes RB, Baggs J, Sapiano MRP, Fiore AE, Jernigan JA, Epstein L: **Assessing Variability in Hospital-Level Mortality Among U.S. Medicare Beneficiaries With Hospitalizations for Severe Sepsis and Septic Shock**. *Crit Care Med* 2018, **46**(11):1753-1760.
46. Bauer M, Gerlach H, Vogelmann T, Preissing F, Stiefel J, Adam D: **Mortality in sepsis and septic shock in Europe, North America and Australia between 2009 and 2019-results from a systematic review and meta-analysis**. *Crit Care* 2020, **24**(1):239.
47. Seymour CW, Gesten F, Prescott HC, Friedrich ME, Iwashyna TJ, Phillips GS, Lemeshow S, Osborn T, Terry KM, Levy MM: **Time to Treatment and Mortality during Mandated Emergency Care for Sepsis**. *N Engl J Med* 2017, **376**(23):2235-2244.
48. Ferrer R, Artigas A, Levy MM, Blanco J, Gonzalez-Diaz G, Garnacho-Montero J, Ibanez J, Palencia E, Quintana M, de la Torre-Prados MV *et al*: **Improvement in process of care and outcome after a multicenter severe sepsis educational program in Spain**. *JAMA* 2008, **299**(19):2294-2303.
49. Rhee C, Yu T, Wang R, Kadri SS, Fram D, Chen HC, Klompas M, Program CDCPE: **Association Between Implementation of the Severe Sepsis and Septic Shock Early Management Bundle Performance Measure and Outcomes in Patients With Suspected Sepsis in US Hospitals**. *JAMA Netw Open* 2021, **4**(12):e2138596.
50. Park RE, Brook RH, Kosecoff J, Keesey J, Rubenstein L, Keeler E, Kahn KL, Rogers WH, Chassin MR: **Explaining variations in hospital death rates. Randomness, severity of illness, quality of care**. *JAMA* 1990, **264**(4):484-490.

51. ColeBrook LP, A.W. : **Treatment of 106 cases of peurperal fever by sulphanilamide(streptocide).** *Lancet* 1937, **230**(5961):1237-1242.
52. OECD: **Beyond containment. Health systems responses to COVID-19 in the OECD**. In: *Tackling coronavirus (COVID-19): Contributing to a Global Effort.* 2020.
53. Halpern NA, Pastores SM: **Critical care medicine in the United States 2000-2005: an analysis of bed numbers, occupancy rates, payer mix, and costs**. *Crit Care Med* 2010, **38**(1):65-71.
54. Chen LM, Render M, Sales A, Kennedy EH, Wiitala W, Hofer TP: **Intensive care unit admitting patterns in the Veterans Affairs health care system**. *Arch Intern Med* 2012, **172**(16):1220-1226.
55. Rhodes A, Ferdinande P, Flaatten H, Guidet B, Metnitz PG, Moreno RP: **The variability of critical care bed numbers in Europe**. *Intensive Care Med* 2012, **38**(10):1647-1653.
56. Wunsch H, Angus DC, Harrison DA, Linde-Zwirble WT, Rowan KM: **Comparison of medical admissions to intensive care units in the United States and United Kingdom**. *Am J Respir Crit Care Med* 2011, **183**(12):1666-1673.
57. Carr BG, Addyson DK, Kahn JM: **Variation in critical care beds per capita in the United States: implications for pandemic and disaster planning**. *JAMA* 2010, **303**(14):1371-1372.
58. **Population and household estimates, England and Wales: Census 2021** [https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/populationandhouseholdestimatesenglandandwales/census2021]
59. **Critical Care Bed Capacity and Urgent Operations Cancelled** [https://www.england.nhs.uk/statistics/statistical-work-areas/critical-care-capacity/]
60. Bravata DM, Perkins AJ, Myers LJ, Arling G, Zhang Y, Zillich AJ, Reese L, Dysangco A, Agarwal R, Myers J *et al*: **Association of Intensive Care Unit Patient Load and Demand With Mortality Rates in US Department of Veterans Affairs Hospitals During the COVID-19 Pandemic**. *JAMA Netw Open* 2021, **4**(1):e2034266.
61. Wilcox ME, Rowan KM, Harrison DA, Doidge JC: **Does Unprecedented ICU Capacity Strain, As Experienced During the COVID-19 Pandemic, Impact Patient Outcome?** *Crit Care Med* 2022, **50**(6):e548-e556.
62. Culyer AJNJP: **Handbook of health economics. Volume 1A Volume 1A**. 2000.
63. Guidet B, Leblanc G, Simon T, Woimant M, Quenot JP, Ganansia O, Maignan M, Yordanov Y, Delerme S, Doumenc B *et al*: **Effect of Systematic Intensive Care Unit Triage on Long-term Mortality Among Critically Ill Elderly Patients in France: A Randomized Clinical Trial**. *JAMA* 2017, **318**(15):1450-1459.
64. Barnato AE, Herndon MB, Anthony DL, Gallagher PM, Skinner JS, Bynum JP, Fisher ES: **Are regional variations in end-of-life care intensity explained by patient preferences?: A Study of the US Medicare Population**. *Med Care* 2007, **45**(5):386-393.
65. Irwin KE, Greer JA, Khatib J, Temel JS, Pirl WF: **Early palliative care and metastatic non-small cell lung cancer: potential mechanisms of prolonged survival**. *Chron Respir Dis* 2013, **10**(1):35-47.
66. Rhee C, Dantes R, Epstein L, Murphy DJ, Seymour CW, Iwashyna TJ, Kadri SS, Angus DC, Danner RL, Fiore AE *et al*: **Incidence and Trends of Sepsis in US Hospitals Using Clinical vs Claims Data, 2009-2014**. *Jama* 2017, **318**(13):1241-1249.

67. Kaukonen K-M, Bailey M, Suzuki S, Pilcher D, Bellomo R: **Mortality related to severe sepsis and septic shock among critically ill patients in Australia and New Zealand, 2000-2012**. *Jama* 2014, **311**(13):1308-1316.
68. Fleischmann-Struzek C, Mikolajetz A, Schwarzkopf D, Cohen J, Hartog CS, Pletz M, Gastmeier P, Reinhart K: **Challenges in assessing the burden of sepsis and understanding the inequalities of sepsis outcomes between National Health Systems: secular trends in sepsis and infection incidence and mortality in Germany**. *Intensive care medicine* 2018, **44**(11):1826-1835.
69. Balsa AI, McGuire TG: **Prejudice, clinical uncertainty and stereotyping as sources of health disparities**. *J Health Econ* 2003, **22**(1):89-116.
70. Crilly T, Le Grand JJ: **The motivation and behaviour of hospital Trusts**. *Soc Sci Med* 2004, **58**(10):1809-1823.
71. Black N, Payne M: **Directory of clinical databases: improving and promoting their use**. *Qual Saf Health Care* 2003, **12**(5):348-352.
72. Harrison DA, Brady AR, Rowan K: **Case mix, outcome and length of stay for admissions to adult, general critical care units in England, Wales and Northern Ireland: the Intensive Care National Audit & Research Centre Case Mix Programme Database**. *Crit Care* 2004, **8**(2):R99-111.
73. Young JD, Goldfrad C, Rowan K: **Development and testing of a hierarchical method to code the reason for admission to intensive care units: the ICNARC Coding Method. Intensive Care National Audit & Research Centre**. *Br J Anaesth* 2001, **87**(4):543-548.
74. Seymour CW, Liu VX, Iwashyna TJ, Brunkhorst FM, Rea TD, Scherag A, Rubenfeld G, Kahn JM, Shankar-Hari M, Singer M *et al*: **Assessment of Clinical Criteria for Sepsis: For the Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3)**. *JAMA* 2016, **315**(8):762-774.
75. Ferrando-Vivas P, Jones A, Rowan KM, Harrison DA: **Development and validation of the new ICNARC model for prediction of acute hospital mortality in adult critical care**. *J Crit Care* 2017, **38**:335-339.
76. Zimmerman JE, Kramer AA, McNair DS, Malila FM, Shaffer VL: **Intensive care unit length of stay: Benchmarking based on Acute Physiology and Chronic Health Evaluation (APACHE) IV**. *Crit Care Med* 2006, **34**(10):2517-2529.
77. de Keizer NF, Bonsel GJ, Goldfad C, Rowan KM: **The added value that increasing levels of diagnostic information provide in prognostic models to estimate hospital mortality for adult intensive care patients**. *Intensive Care Med* 2000, **26**(5):577-584.
78. Arts DG, Bosman RJ, de Jonge E, Joore JC, de Keizer NF: **Training in data definitions improves quality of intensive care data**. *Crit Care* 2003, **7**(2):179-184.
79. Chen LM, Martin CM, Morrison TL, Sibbald WJ: **Interobserver variability in data collection of the APACHE II score in teaching and community hospitals**. *Crit Care Med* 1999, **27**(9):1999-2004.
80. Simkins M, Iqbal A, Gronemeyer A, Konzen L, White J, Koenig M, Palmer C, Kerby P, Buckman S, Despotovic V *et al*: **Inter-Rater Reliability and Impact of Disagreements on Acute Physiology and Chronic Health Evaluation IV Mortality Predictions**. *Crit Care Explor* 2019, **1**(10):e0059.
81. Wunsch H, Brady AR, Rowan K: **Impact of exclusion criteria on case mix, outcome, and length of stay for the severity of disease scoring methods in common use in critical care**. *J Crit Care* 2004, **19**(2):67-74.

82. Harrison DA, Brady AR, Parry GJ, Carpenter JR, Rowan K: **Recalibration of risk prediction models in a large multicenter cohort of admissions to adult, general critical care units in the United Kingdom**. *Crit Care Med* 2006, **34**(5):1378-1388.
83. Harrison DA, Parry GJ, Rowan K: **Risk prediction models: recalibration**. *Crit Care Med* 2007, **35**(1):326-327.
84. Harrison DA, Ferrando-Vivas P, Shahin J, Rowan KM. In: *Ensuring comparisons of health-care providers are fair: development and validation of risk prediction models for critically ill patients.* Southampton (UK); 2015.
85. Luft HS, Bunker JP, Enthoven AC: **Should operations be regionalized? The empirical relation between surgical volume and mortality. 1979**. *Clin Orthop Relat Res* 2007, **457**:3-9.
86. Halm EA, Lee C, Chassin MR: **Is volume related to outcome in health care? A systematic review and methodologic critique of the literature**. *Ann Intern Med* 2002, **137**(6):511-520.
87. Mesman R, Faber MJ, Berden B, Westert GP: **Evaluation of minimum volume standards for surgery in the Netherlands (2003-2017): A successful policy?** *Health Policy* 2017, **121**(12):1263-1273.
88. Bhattarai N, McMeekin P, Price C, Vale L: **Economic evaluations on centralisation of specialised healthcare services: a systematic review of methods**. *BMJ Open* 2016, **6**(5):e011214.
89. Metcalfe D, Bouamra O, Parsons NR, Aletrari MO, Lecky FE, Costa ML: **Effect of regional trauma centralization on volume, injury severity and outcomes of injured patients admitted to trauma centres**. *Br J Surg* 2014, **101**(8):959-964.
90. Morris S, Hunter RM, Ramsay AI, Boaden R, McKevitt C, Perry C, Pursani N, Rudd AG, Schwamm LH, Turner SJ *et al*: **Impact of centralising acute stroke services in English metropolitan areas on mortality and length of hospital stay: difference-in-differences analysis**. *BMJ* 2014, **349**:g4757.
91. Fitzsimons KJ, Mukarram S, Copley LP, Deacon SA, van der Meulen JH: **Centralisation of services for children with cleft lip or palate in England: a study of hospital episode statistics**. *BMC Health Serv Res* 2012, **12**:148.
92. Christiansen T, Vrangbaek K: **Hospital centralization and performance in Denmark-Ten years on**. *Health Policy* 2018, **122**(4):321-328.
93. Magnussen J, Hagen TP, Kaarboe OM: **Centralized or decentralized? A case study of Norwegian hospital reform**. *Soc Sci Med* 2007, **64**(10):2129-2137.
94. Nimptsch U, Mansky T: **Hospital volume and mortality for 25 types of inpatient treatment in German hospitals: observational study using complete national data from 2009 to 2014**. *BMJ Open* 2017, **7**(9):e016184.
95. Lindenauer PK, Behal R, Murray CK, Nsa W, Houck PM, Bratzler DW: **Volume, quality of care, and outcome in pneumonia**. *Ann Intern Med* 2006, **144**(4):262-269.
96. Ross JS, Normand SL, Wang Y, Ko DT, Chen J, Drye EE, Keenan PS, Lichtman JH, Bueno H, Schreiner GC *et al*: **Hospital volume and 30-day mortality for three common medical conditions**. *N Engl J Med* 2010, **362**(12):1110-1118.
97. Haut ER, Pronovost PJ, Schneider EB: **Limitations of administrative databases**. *JAMA* 2012, **307**(24):2589; author reply 2589-2590.
98. Kumbhani DJ, Fonarow GC, Heidenreich PA, Schulte PJ, Lu D, Hernandez A, Yancy C, Bhatt DL: **Association Between Hospital Volume, Processes of Care, and Outcomes**

in Patients Admitted With Heart Failure: Insights From Get With The Guidelines-Heart Failure. *Circulation* 2018, **137**(16):1661-1670.
99. Joynt KE, Orav EJ, Jha AK: **The association between hospital volume and processes, outcomes, and costs of care for congestive heart failure**. *Ann Intern Med* 2011, **154**(2):94-102.
100. Pronovost PJ, Angus DC, Dorman T, Robinson KA, Dremsizov TT, Young TL: **Physician staffing patterns and clinical outcomes in critically ill patients: a systematic review**. *JAMA* 2002, **288**(17):2151-2162.
101. Angus DC, Black N: **Improving care of the critically ill: institutional and health-care system approaches**. *Lancet* 2004, **363**(9417):1314-1320.
102. Nguyen YL, Wallace DJ, Yordanov Y, Trinquart L, Blomkvist J, Angus DC, Kahn JM, Ravaud P, Guidet B: **The Volume-Outcome Relationship in Critical Care: A Systematic Review and Meta-analysis**. *Chest* 2015, **148**(1):79-92.
103. Viswanathan M, Patnode CD, Berkman ND, Bass EB, Chang S, Hartling L, Murad MH, Treadwell JR, Kane RL: **Recommendations for assessing the risk of bias in systematic reviews of health-care interventions**. *J Clin Epidemiol* 2018, **97**:26-34.
104. Galbraith RF: **A note on graphical presentation of estimated odds ratios from several clinical trials**. *Stat Med* 1988, **7**(8):889-894.
105. Higgins JP, Thompson SG, Deeks JJ, Altman DG: **Measuring inconsistency in meta-analyses**. *BMJ* 2003, **327**(7414):557-560.
106. Begg CB, Mazumdar M: **Operating characteristics of a rank correlation test for publication bias**. *Biometrics* 1994, **50**(4):1088-1101.
107. Shi L, Lin L: **The trim-and-fill method for publication bias: practical guidelines and recommendations based on a large database of meta-analyses**. *Medicine (Baltimore)* 2019, **98**(23):e15987.
108. Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, Bellomo R, Bernard GR, Chiche JD, Coopersmith CM *et al*: **The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3)**. *JAMA* 2016, **315**(8):801-810.
109. Rudd KE, Johnson SC, Agesa KM, Shackelford KA, Tsoi D, Kievlan DR, Colombara DV, Ikuta KS, Kissoon N, Finfer S *et al*: **Global, regional, and national sepsis incidence and mortality, 1990-2017: analysis for the Global Burden of Disease Study**. *Lancet* 2020, **395**(10219):200-211.
110. **Service delivery and safety: Improving the prevention, diagnosis and clinical management of sepsis** [https://www.who.int/servicedeliverysafety/areas/sepsis/en/]
111. Ikoma N, Kim B, Elting LS, Shih YT, Badgwell BD, Mansfield P: **Trends in Volume-Outcome Relationship in Gastrectomies in Texas**. *Ann Surg Oncol* 2019, **26**(9):2694-2702.
112. Lee KC, Sethuraman K, Yong J: **On the Hospital Volume and Outcome Relationship: Does Specialization Matter More Than Volume?** *Health Serv Res* 2015, **50**(6):2019-2036.
113. O'Donnell TFX, Boitano LT, Deery SE, Lancaster RT, Siracuse JJ, Schermerhorn ML, Scali ST, Patel VI: **Hospital Volume Matters: The Volume-Outcome Relationship in Open Juxtarenal AAA Repair**. *Ann Surg* 2020, **271**(1):184-190.

114. Sewalt CA, Wiegers EJA, Venema E, Lecky FE, Schuit SCE, Den Hartog D, Lingsma HF: **The volume-outcome relationship in severely injured patients: A systematic review and meta-analysis**. *J Trauma Acute Care Surg* 2018, **85**(4):810-819.
115. Luft HS, Bunker JP, Enthoven AC: **Should operations be regionalized? The empirical relation between surgical volume and mortality**. *N Engl J Med* 1979, **301**(25):1364-1369.
116. Alsfasser G, Leicht H, Gunster C, Rau BM, Schillinger G, Klar E: **Volume-outcome relationship in pancreatic surgery**. *Br J Surg* 2016, **103**(1):136-143.
117. Wright JD: **The Volume-Outcome Paradigm for Gynecologic Surgery: Clinical and Policy Implications**. *Clin Obstet Gynecol* 2020.
118. Gu WJ, Wu XD, Zhou Q, Zhang J, Wang F, Ma ZL, Gu XP: **Relationship between Annualized Case Volume and Mortality in Sepsis: A Dose-Response Meta-analysis**. *Anesthesiology* 2016, **125**(1):168-179.
119. Walkey AJ, Wiener RS: **Hospital case volume and outcomes among patients hospitalized with severe sepsis**. *Am J Respir Crit Care Med* 2014, **189**(5):548-555.
120. Shahin J, Harrison DA, Rowan KM: **Relation between volume and outcome for patients with severe sepsis in United Kingdom: retrospective cohort study**. *BMJ* 2012, **344**:e3394.
121. Peelen L, de Keizer NF, Peek N, Scheffer GJ, van der Voort PH, de Jonge E: **The influence of volume and intensive care unit organization on hospital mortality in patients admitted with severe sepsis: a retrospective multicentre cohort study**. *Crit Care* 2007, **11**(2):R40.
122. Engel C, Brunkhorst FM, Bone HG, Brunkhorst R, Gerlach H, Grond S, Gruendling M, Huhle G, Jaschinski U, John S *et al*: **Epidemiology of sepsis in Germany: results from a national prospective multicenter study**. *Intensive Care Med* 2007, **33**(4):606-618.
123. Reinikainen M, Karlsson S, Varpula T, Parviainen I, Ruokonen E, Varpula M, Ala-Kokko T, Pettila V: **Are small hospitals with small intensive care units able to treat patients with severe sepsis?** *Intensive Care Med* 2010, **36**(4):673-679.
124. Needham DM, Bronskill SE, Rothwell DM, Sibbald WJ, Pronovost PJ, Laupacis A, Stukel TA: **Hospital volume and mortality for mechanical ventilation of medical and surgical patients: a population-based analysis using administrative data**. *Crit Care Med* 2006, **34**(9):2349-2354.
125. Hawks L, Himmelstein DU, Woolhandler S, Bor DH, Gaffney A, McCormick D: **Trends in Unmet Need for Physician and Preventive Services in the United States, 1998-2017**. *JAMA Intern Med* 2020, **180**(3):439-448.
126. Dudley RA, Johansen KL, Brand R, Rennie DJ, Milstein A: **Selective referral to high-volume hospitals: estimating potentially avoidable deaths**. *JAMA* 2000, **283**(9):1159-1166.
127. Joseph B, Morton JM, Hernandez-Boussard T, Rubinfeld I, Faraj C, Velanovich V: **Relationship between hospital volume, system clinical resources, and mortality in pancreatic resection**. *J Am Coll Surg* 2009, **208**(4):520-527.
128. Macomber CW, Shaw JJ, Santry H, Saidi RF, Jabbour N, Tseng JF, Bozorgzadeh A, Shah SA: **Centre volume and resource consumption in liver transplantation**. *HPB (Oxford)* 2012, **14**(8):554-559.
129. Nguyen YL, Milbrandt EB, Weissfeld LA, Kahn JM, Chiche JD, Aegerter P, Clermont G, Kellum JA, Guidet B, Angus DC: **Intensive care unit renal support therapy volume is not associated with patient outcome**. *Crit Care Med* 2011, **39**(11):2470-2477.

130. Vincent JL, Marshall JC, Namendys-Silva SA, Francois B, Martin-Loeches I, Lipman J, Reinhart K, Antonelli M, Pickkers P, Njimi H *et al*: **Assessment of the worldwide burden of critical illness: the intensive care over nations (ICON) audit**. *Lancet Respir Med* 2014, **2**(5):380-386.
131. Morris S, Ramsay AIG, Boaden RJ, Hunter RM, McKevitt C, Paley L, Perry C, Rudd AG, Turner SJ, Tyrrell PJ *et al*: **Impact and sustainability of centralising acute stroke services in English metropolitan areas: retrospective analysis of hospital episode statistics and stroke national audit data**. *BMJ* 2019, **364**:l1.
132. Harrison DA, Lone NI, Haddow C, MacGillivray M, Khan A, Cook B, Rowan KM: **External validation of the Intensive Care National Audit & Research Centre (ICNARC) risk prediction model in critical care units in Scotland**. *BMC Anesthesiol* 2014, **14**:116.
133. Harrison DA, Parry GJ, Carpenter JR, Short A, Rowan K: **A new risk prediction model for critical care: the Intensive Care National Audit & Research Centre (ICNARC) model**. *Crit Care Med* 2007, **35**(4):1091-1098.
134. Harrison DA, Welch CA, Eddleston JM: **The epidemiology of severe sepsis in England, Wales and Northern Ireland, 1996 to 2004: secondary analysis of a high quality clinical database, the ICNARC Case Mix Programme Database**. *Crit Care* 2006, **10**(2):R42.
135. Vandenbroucke JP, von Elm E, Altman DG, Gotzsche PC, Mulrow CD, Pocock SJ, Poole C, Schlesselman JJ, Egger M, Initiative S: **Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration**. *Epidemiology* 2007, **18**(6):805-835.
136. Royston P: **Model Selection for Univariable Fractional Polynomials**. *The Stata Journal* 2017, **17**(3):619-629.
137. Lusa L, Ahlin C: **Restricted cubic splines for modelling periodic data**. *PLoS One* 2020, **15**(10):e0241364.
138. Rabe-Hesketh S, Skrondal A, Skrondal A: **Multilevel and longitudinal modeling using stata**, 2nd edn. College Station, Tex.: Stata Press Publication; 2008.
139. Austin PC, Wagner P, Merlo J: **The median hazard ratio: a useful measure of variance and general contextual effects in multilevel survival analysis**. *Stat Med* 2017, **36**(6):928-938.
140. Royston P: **Model selection for univariable fractional polynomials**. *Stata J* 2017, **17**(3):619-629.
141. Binder H, Sauerbrei W: **Adding local components to global functions for continuous covariates in multivariable regression modeling**. *Stat Med* 2010, **29**(7-8):808-817.
142. Ding P, VanderWeele TJ: **Sensitivity Analysis Without Assumptions**. *Epidemiology* 2016, **27**(3):368-377.
143. VanderWeele TJ, Ding P: **Sensitivity Analysis in Observational Research: Introducing the E-Value**. *Ann Intern Med* 2017, **167**(4):268-274.
144. Schunck R, Perales F: **Within- and Between-cluster Effects in Generalized Linear Mixed Models: A Discussion of Approaches and the Xthybrid command**. *The Stata Journal* 2017, **17**(1):89-115.
145. Panageas KS, Schrag D, Riedel E, Bach PB, Begg CB: **The effect of clustering of outcomes on the association of procedure volume and surgical outcomes**. *Ann Intern Med* 2003, **139**(8):658-665.

146. Abbenbroek B, Duffield CM, Elliott D: **The intensive care unit volume-mortality relationship, is bigger better? An integrative literature review**. *Aust Crit Care* 2014, **27**(4):157-164; quiz 165.
147. Thompson P: **Chapter 10 - Learning by Doing**. In: *Handbook of the Economics of Innovation.* Edited by Hall BH, Rosenberg N, vol. 1: North-Holland; 2010: 429-476.
148. Alchian A: **Reliability of Progress Curves in Airframe Production**. *Econometrica* 1963, **31**(4):679-693.
149. Argote L, Miron-Spektor E: **Organizational Learning: From Experience to Knowledge**. *Organization Science* 2011, **22**(5):1123-1137.
150. Finks JF, Osborne NH, Birkmeyer JD: **Trends in hospital volume and operative mortality for high-risk surgery**. *N Engl J Med* 2011, **364**(22):2128-2137.
151. Maharaj R, McGuire A, Street A: **Association of Annual Intensive Care Unit Sepsis Caseload With Hospital Mortality From Sepsis in the United Kingdom, 2010-2016**. *JAMA Netw Open* 2021, **4**(6):e2115305.
152. Mesman R, Westert GP, Berden BJ, Faber MJ: **Why do high-volume hospitals achieve better outcomes? A systematic review about intermediate factors in volume-outcome relationships**. *Health Policy* 2015, **119**(8):1055-1067.
153. Nicholas LH, Stith SS: **Are Transplant Centers That Meet Insurer Minimum Volume Requirements Better Quality?** *Med Care Res Rev* 2021, **78**(5):502-510.
154. Leung S, Pastores SM, Oropello JM, Lilly CM, Galvagno SM, Jr., Badjatia N, Jacobi J, Herr DL, Oliveira JD, Academic Leaders in Critical Care Medicine Task Force of the Society of Critical Care M: **Regionalization of Critical Care in the United States: Current State and Proposed Framework From the Academic Leaders in Critical Care Medicine Task Force of the Society of Critical Care Medicine**. *Crit Care Med* 2022, **50**(1):37-49.
155. Gaffney AW: **ICU Equity and Regionalization in the COVID-19 Era**. *Ann Am Thorac Soc* 2022.
156. Cooper Z, Gibbons S, Jones S, McGuire A: **Does Hospital Competition Save Lives? Evidence from the English NHS Patient Choice Reforms**. *Econ J (London)* 2011, **121**(554):F228-F260.
157. Gaynor M, Seider H, Vogt WB: **The Volume-Outcome Effect, Scale Economies, and Learning-by-Doing**. *The American Economic Review* 2005, **95**:243-247.
158. Van Gestel R, Muller T, Bosmans J: **Does my high blood pressure improve your survival? Overall and subgroup learning curves in health**. *Health Econ* 2017, **26**(9):1094-1109.
159. Ho V: **Learning and the evolution of medical technologies: the diffusion of coronary angioplasty**. *J Health Econ* 2002, **21**(5):873-885.
160. Ho SS, Chan YL, Yeung DK, Metreweli C: **Blood flow volume quantification of cerebral ischemia: comparison of three noninvasive imaging techniques of carotid and vertebral arteries**. *AJR Am J Roentgenol* 2002, **178**(3):551-556.
161. Avdic D, Lundborg P, Vikstrom J: **Estimating returns to hospital volume: Evidence from advanced cancer surgery**. *J Health Econ* 2019, **63**:81-99.
162. World Health A: **Improving the prevention, diagnosis and clinical management of sepsis**. In. Geneva: World Health Organization; 2017.
163. Walton NT, Mohr NM: **Concept review of regionalized systems of acute care: Is regionalization the next frontier in sepsis care?** *J Am Coll Emerg Physicians Open* 2022, **3**(1):e12631.

164. Vandenbroucke JP, von Elm E, Altman DG, Gotzsche PC, Mulrow CD, Pocock SJ, Poole C, Schlesselman JJ, Egger M, Initiative S: **Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration**. *Int J Surg* 2014, **12**(12):1500-1524.
165. Benkard CL: **Learning and Forgetting: The Dynamics of Aircraft Production**. *American Economic Review* 2000, **90**(4):1034-1054.
166. **Causality, learning and forgetting in surgery.** [http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.183.4869.]
167. Ervin JN, Kahn JM, Cohen TR, Weingart LR: **Teamwork in the intensive care unit**. *Am Psychol* 2018, **73**(4):468-477.
168. Gowrisankaran G, Olin JM: **Causality , Learning and Forgetting in Surgery**. In*: 2005*; 2005.
169. Facchini G: **Forgetting-by-not-doing: The case of surgeons and cesarean sections**. *Health Econ* 2022, **31**(3):481-495.
170. Shuhaiber JH, Goldsmith K, Nashef SA: **Impact of cardiothoracic resident turnover on mortality after cardiac surgery: a dynamic human factor**. *Ann Thorac Surg* 2008, **86**(1):123-130; discussion 130-121.
171. Bozic KJ, Maselli J, Pekow PS, Lindenauer PK, Vail TP, Auerbach AD: **The influence of procedure volumes and standardization of care on quality and efficiency in total joint replacement surgery**. *J Bone Joint Surg Am* 2010, **92**(16):2643-2652.
172. Huguet M: **Centralization of care in high volume hospitals and inequalities in access to care**. *Soc Sci Med* 2020, **260**:113177.
173. Lott JP, Iwashyna TJ, Christie JD, Asch DA, Kramer AA, Kahn JM: **Critical illness outcomes in specialty versus general intensive care units**. *Am J Respir Crit Care Med* 2009, **179**(8):676-683.
174. Kramer AH, Zygun DA: **Do neurocritical care units save lives? Measuring the impact of specialized ICUs**. *Neurocrit Care* 2011, **14**(3):329-333.
175. Diringer MN, Edwards DF: **Admission to a neurologic/neurosurgical intensive care unit is associated with reduced mortality rate after intracerebral hemorrhage**. *Crit Care Med* 2001, **29**(3):635-640.
176. Damian MS, Ben-Shlomo Y, Howard R, Harrison DA: **Admission patterns and survival from status epilepticus in critical care in the UK: an analysis of the Intensive Care National Audit and Research Centre Case Mix Programme database**. *Eur J Neurol* 2020, **27**(3):557-564.
177. McCredie VA, Alali AS, Scales DC, Rubenfeld GD, Cuthbertson BH, Nathens AB: **Impact of ICU Structure and Processes of Care on Outcomes After Severe Traumatic Brain Injury: A Multicenter Cohort Study**. *Crit Care Med* 2018, **46**(7):1139-1149.
178. Damian MS, Ben-Shlomo Y, Howard R, Bellotti T, Harrison D, Griggs K, Rowan K: **The effect of secular trends and specialist neurocritical care on mortality for patients with intracerebral haemorrhage, myasthenia gravis and Guillain-Barre syndrome admitted to critical care : an analysis of the Intensive Care National Audit & Research Centre (ICNARC) national United Kingdom database**. *Intensive Care Med* 2013, **39**(8):1405-1412.
179. de la Paz A, Orgel R, Hartsell SE, Pauley E, Katz JN: **Getting cardiogenic shock patients to the right place-How initial intensive care unit triage decisions impact processes of care and outcomes**. *Am Heart J* 2020, **230**:66-70.

180. Elf K, Nilsson P, Enblad P: **Outcome after traumatic brain injury improved by an organized secondary insult program and standardized neurointensive care**. *Crit Care Med* 2002, **30**(9):2129-2134.
181. Cram P, Rosenthal GE, Vaughan-Sarrazin MS: **Cardiac revascularization in specialty and general hospitals**. *N Engl J Med* 2005, **352**(14):1454-1462.
182. Littek W: **Labor, Division of**. In: *International Encyclopedia of the Social & Behavioral Sciences.* Edited by Smelser NJ, Baltes PB. Oxford: Pergamon; 2001: 8220-8226.
183. Smith A: **Wealth of Nations. Books 1-3 Books 1-3**; 2009.
184. Ricardo DMJR: **The works of David Ricardo**. London: J. Murray; 1888.
185. Eastaugh SR: **Hospital costs and specialization: benefits of limiting the number of product lines**. *J Health Care Finance* 2009, **36**(2):24-34.
186. Zwanziger J, Melnick GA, Simonson L: **Differentiation and specialization in the California hospital industry 1983 to 1988**. *Med Care* 1996, **34**(4):361-372.
187. Lee KS, Chun KH, Lee JS: **Reforming the hospital service structure to improve efficiency: urban hospital specialization**. *Health Policy* 2008, **87**(1):41-49.
188. **Health at a glance. Europe 2020**
189. Wunsch H, Gershengorn H, Scales DC: **Economics of ICU organization and management**. *Critical care clinics* 2012, **28**(1):25-v.
190. Evans L, Best C: **Managing malnutrition in patients with dementia**. *Nurs Stand* 2015, **29**(28):50-57.
191. Barro JR, Huckman RS, Kessler DP: **The effects of cardiac specialty hospitals on the cost and quality of medical care**. *J Health Econ* 2006, **25**(4):702-721.
192. Grasic K, Mason AR, Street A: **Paying for the quantity and quality of hospital care: the foundations and evolution of payment policy in England**. *Health Econ Rev* 2015, **5**(1):50.
193. Gaughan J, Gutacker N, Grasic K, Kreif N, Siciliani L, Street A: **Paying for efficiency: Incentivising same-day discharges in the English NHS**. *J Health Econ* 2019, **68**:102226.
194. Schneider JE, Miller TR, Ohsfeldt RL, Morrisey MA, Zelner BA, Pengxiang L: **The economics of specialty hospitals**. *Med Care Res Rev* 2008, **65**(5):531-553; discussion 554-563.
195. Barro JR, Huckman RS, Kessler DP: **The Effects of Cardiac Specialty Hospitals on the Cost and Quality of Medical Care**. *Health Economics eJournal* 2006.
196. KC DS, Terwiesch C: **The Effects of Focus on Performance: Evidence from California Hospitals**. *Management Science* 2011, **57**(11):1897-1912.
197. Huckman RS, Zinner DE: **Does Focus Improve Operational Performance? Lessons from the Management of Clinical Trials**. *Strategic Management Journal* 2008, **29**(2):173-193.
198. W. S: **The focused factory**. *Harvard Business Review* 1974, **52**:113-120.
199. Kuntz L, Scholtes S, Sülz S: **Separate and Concentrate: Accounting for Patient Complexity in General Hospitals**. *Management Science* 2019, **65**(6):2482-2501.
200. Hagen TP, Vaughan-Sarrazin MS, Cram P: **Relation between hospital orthopaedic specialisation and outcomes in patients aged 65 and older: retrospective analysis of US Medicare data**. *BMJ* 2010, **340**:c165.
201. Schneider JE, Miller TR, Ohsfeldt RL, Morrisey MA, Zelner BA, Li P: **The Economics of Specialty Hospitals**. *Medical Care Research and Review* 2008, **65**(5):531-553.

202. Regge MD, Trybou J, Gemmel P, Duyck P, Annemans L: **Physician-owned specialized facilities: focused factories or destructive competition?: a systematic review**. In*: 2013*; 2013.
203. Kramer AH, Zygun DA: **Neurocritical care: why does it make a difference?** *Current Opinion in Critical Care* 2014, **20**(2):174-181.
204. Goldfarb M, van Diepen S, Liszkowski M, Jentzer JC, Pedraza I, Cercek B: **Noncardiovascular Disease and Critical Care Delivery in a Contemporary Cardiac and Medical Intensive Care Unit**. *Journal of Intensive Care Medicine* 2019, **34**(7):537-543.
205. Cooper Z, Gibbons S, Skellern M: **Does competition from private surgical centres improve public hospitals' performance? Evidence from the English National Health Service**. *Journal of Public Economics* 2018, **166**:63-80.
206. Regge MD, Trybou J, Gemmel P, Duyck P, Annemans L: **Physician-Owned Specialized Facilities: Focused Factories or Destructive Competition?** *Academy of Management Proceedings* 2013, **2013**(1):13667.
207. Swanson A: **Physician investment in hospitals: Specialization, selection, and quality in cardiac care**. *J Health Econ* 2021, **80**:102519.
208. Gabel JR, Fahlman C, Kang R, Wozniak G, Kletke P, Hay JW: **Where do I send thee? Does physician-ownership affect referral patterns to ambulatory surgery centers?** *Health Aff (Millwood)* 2008, **27**(3):w165-174.
209. Clark JR, Huckman RS: **Broadening Focus: Spillovers, Complementarities, and Specialization in the Hospital Industry**. *Management Science* 2012, **58**(4):708-722.
210. Conrad Kobel ET: **Hospital specialisation within a DRG-Framework: The Austrian Case**. In: *Faculty of Economics and Statistics, University of Innsbruck.* Edited by Papers W; 2013.
211. Lindlbauer I, Schreyogg J: **The relationship between hospital specialization and hospital efficiency: do different measures of specialization lead to different results?** *Health Care Manag Sci* 2014, **17**(4):365-378.
212. Lindlbauer I, Schreyögg J: **The relationship between hospital specialization and hospital efficiency: do different measures of specialization lead to different results?** *Health Care Management Science* 2014, **17**(4):365-378.
213. Ferrando-Vivas P, Jones A, Rowan KM, Harrison DA: **Development and validation of the new ICNARC model for prediction of acute hospital mortality in adult critical care**. *Journal of Critical Care* 2017, **38**:335-339.
214. Mielke D, Malinova V, Moerer O, Suntheim P, Voit M, Rohde V: **Does the subspecialty of an intensive care unit (ICU) has an impact on outcome in patients suffering from aneurysmal subarachnoid hemorrhage?** *Neurosurgical Review* 2019, **42**(1):147-153.
215. Buzzelli M, Su J: **Multi-level modelling in health research: a caution and rejoinder on temporally mismatched data**. *Soc Sci Med* 2006, **62**(5):1215-1218.
216. Bell A, Fairbrother M, Jones K: **Fixed and random effects models: making an informed choice**. *Quality & Quantity* 2019, **53**(2):1051-1074.
217. Skinner W: **The Focused Factory**. *Harvard Business Review* 1974, **May-June**:114-121.
218. McCrum ML, Lipsitz SR, Berry WR, Jha AK, Gawande AA: **Beyond volume: does hospital complexity matter?: an analysis of inpatient surgical mortality in the United States**. *Med Care* 2014, **52**(3):235-242.

219. Cram P, Vaughan-Sarrazin MS, Wolf B, Katz JN, Rosenthal GE: **A comparison of total hip and knee replacement in specialty and general hospitals**. *J Bone Joint Surg Am* 2007, **89**(8):1675-1684.
220. Arabi YM, Azoulay E, Al-Dorzi HM, Phua J, Salluh J, Binnie A, Hodgson C, Angus DC, Cecconi M, Du B *et al*: **How the COVID-19 pandemic will change the future of critical care**. *Intensive Care Med* 2021, **47**(3):282-291.
221. de Cruppe W, Malik M, Geraedts M: **Minimum volume standards in German hospitals: do they get along with procedure centralization? A retrospective longitudinal data analysis**. *BMC Health Serv Res* 2015, **15**:279.
222. Subramanian MP, Yang Z, Chang SH, Willis D, Zhang J, Semenkovich TR, Heiden BT, Kozower BD, Kreisel D, Meyers BF *et al*: **Minimum Volume Standards for Surgical Care of Early-Stage Lung Cancer: A Cost-Effectiveness Analysis**. *Ann Thorac Surg* 2022.
223. Stanak M, Strohmaier C: **Minimum volume standards in day surgery: a systematic review**. *BMC Health Serv Res* 2020, **20**(1):886.
224. MacKenzie EJ, Rivara FP, Jurkovich GJ, Nathens AB, Frey KP, Egleston BL, Salkever DS, Scharfstein DO: **A national evaluation of the effect of trauma-center care on mortality**. *N Engl J Med* 2006, **354**(4):366-378.
225. Birkmeyer JD, Finlayson EV, Birkmeyer CM: **Volume standards for high-risk surgical procedures: potential benefits of the Leapfrog initiative**. *Surgery* 2001, **130**(3):415-422.
226. Staebler S: **Regionalized systems of perinatal care: health policy considerations**. *Adv Neonatal Care* 2011, **11**(1):37-42.
227. Davie JH, R.M.; Mountford, J.; Morris, S. : **London's Hyperacute Stroke Units Improve Outcomes and Lower Costs**. *Harvard Business Review* 2013.
228. Barnato AE, Kahn JM, Rubenfeld GD, McCauley K, Fontaine D, Frassica JJ, Hubmayr R, Jacobi J, Brower RG, Chalfin D *et al*: **Prioritizing the organization and management of intensive care services in the United States: the PrOMIS Conference**. *Crit Care Med* 2007, **35**(4):1003-1011.
229. Goldstein H: **Hierarchical Data Modeling in the Social Sciences**. *Journal of Educational and Behavioral Statistics* 1995, **20**(2):201-204.
230. Goldstein H, Spiegelhalter DJ: **League Tables and Their Limitations: Statistical Issues in Comparisons of Institutional Performance**. *Journal of the Royal Statistical Society Series A (Statistics in Society)* 1996, **159**(3):385-443.
231. Orsini N, Greenland S: **A procedure to tabulate and plot results after flexible modeling of a quantitative covariate**. *Stata Journal* 2011, **11**(1):1-29.
232. Smith OM, Chant C, Burns KEA, Kaur M, Ashraf S, DosSantos CC, Hwang SW, Friedrich JO: **Characteristics, clinical course, and outcomes of homeless and non-homeless patients admitted to ICU: A retrospective cohort study**. *PLoS One* 2017, **12**(6):e0179207.
233. Ginde AA, Moss M, Shapiro NI, Schwartz RS: **Impact of older age and nursing home residence on clinical outcomes of US emergency department visits for severe sepsis**. *J Crit Care* 2013, **28**(5):606-611.
234. **English indices of deprivation** [https://www.gov.uk/government/collections/english-indices-of-deprivation]
235. Knaus WA, Draper EA, Wagner DP, Zimmerman JE: **APACHE II: a severity of disease classification system**. *Crit Care Med* 1985, **13**(10):818-829.

236. Harrison DA, Rowan KM: **Outcome prediction in critical care: the ICNARC model**. *Curr Opin Crit Care* 2008, **14**(5):506-512.
237. Levy MM, Artigas A, Phillips GS, Rhodes A, Beale R, Osborn T, Vincent JL, Townsend S, Lemeshow S, Dellinger RP: **Outcomes of the Surviving Sepsis Campaign in intensive care units in the USA and Europe: a prospective cohort study**. *Lancet Infect Dis* 2012, **12**(12):919-924.
238. Hastie T, Friedman J, Tisbshirani R: **The Elements of statistical learning : data mining, inference, and prediction**. New York: Springer; 2018.
239. Gelfand AE, Dey DK: **Bayesian Model Choice: Asymptotics and Exact Calculations**. *Journal of the Royal Statistical Society Series B (Methodological)* 1994, **56**(3):501-514.
240. Cameron AC, Triverdi PK: **Microeconometrics using Stata**. College Station, Tex: Stata Press; 2010.
241. Allison PD: **Fixed Effects Regression Models**. Thousand Oaks, Calif.: Sage Publications; 2009.

Appendix one: The Third International Consensus Definitions for Sepsis and Septic Shock[94]

Third International Consensus Definition of Sepsis[94]
Sepsis is defined as life-threatening organ dysfunction caused by a dysregulated host response to infection
Organ dysfunction can be defined as an acute change in SOFA (Sequential Organ Failure Assessment) score of ≥ 2 attributable to infection.
• The baseline SOFA score can be assumed to be zero in patients not known to have pre-existing organ dysfunction.
• A SOFA score≥ 2 reflects an overall mortality risk of 10 % in a general hospital population with suspected infection.
• Sepsis is a life-threatening condition that arises when the body's response to an infection injures its own tissues and organs
• Septic shock is a subset of sepsis in which underlying circulatory and cellular/metabolic abnormalities are profound enough to substantially increase mortality.
• Patients with septic shock can be identified with sepsis with persisting hypotension requiring vasopressors to maintain a mean arterial pressure ≥65mmHg and having a serum lactate level >2 mmol/L (18mg/dL) despite adequate volume resuscitation. Mortality in this sub-group >40%

Appendix two: Association of Intensive Care Unit Annual Sepsis Caseload with Patient Mortality from Sepsis in the United Kingdom, 2010-2016

Table of contents

d. eTable. 2 ICU characteristics between 2010 and 2016

4. Additional results
    a. eFigure 3. Results of the regression analysis with volume as quartiles
    b. eTable 3. Testing the statistical significance between models of increasing complexity with volume defined as a restricted cubic spline
    c. eFigure 4. Variation within ICUs across years empirical Bayes predication
    d. eFigure 5. Subgroups analysis: Adjusted probability of mortality.
5. Sensitivity analyses and robustness checks
    a. Fractional polynomial and selection procedure
        i. eTable 4. Comparisons of model fit for fractional polynomial
        ii. eFigure 6. Functional form of fractional polynomial
        iii. eTable 5. A comparison of the information criteria for the linear, fractional polynomial and restricted cubic spline models.
    b. E-values
        i. Rationale
        ii. Figure 7. Value of the joint minimum strength of association that an unmeasured confounder must have with both an increase in ICU sepsis volume and acute hospital mortality to explain away the volume outcome relationship
    c. Checking for exogeneity of ICU volume
        i. Rationale
        ii. eTable 6. Within and between cluster effects of ICU volume test for exogeneity

d. Checking robustness of the primary analysis to extremes in annual sepsis volume.

i. eTable7 describing quartiles after excluding extremes of annual sepsis volume with probability of acute hospital mortality.

ii. eFigure 8. The predicted mortality and marginal effects for the analysis excluding extremes of annual sepsis volume

The supplementary appendix is set out as follows: We first introduce the hierarchical structure of the data and the multilevel model. We then explore the specification of ICU sepsis volume as quartiles and then as a restricted cubic spline. We perform a sensitivity analysis of specifying volume as a fractional polynomial which is described later in the appendix. We explain the motivation for using a random effects model compared with a fixed effects model and describe the use of the associated empirical Bayes method to predict the ICU-level effect using the estimated random effects. The control variables used in the analysis are then described. We conduct a sensitivity analysis using E-values to explore the potential effects of any missing control variables which is described later in the appendix. We provide additional descriptive data and then return to the results of the empirical analysis and sensitivity analyses. To check the robustness of the random effects model we undertook a regression-based alternative to the Hausman test for endogeneity. We provide a more detailed rationale for the use of the random effects model and conclude with the results of the robustness check.

1. Statistical Methods

a) Multilevel models

i. Rationale

The data used in this study feature patients nested within years nested within ICUs, thereby forming a natural hierarchical structure that is suited to a multilevel modelling technique that models each level of the structure simultaneously. Patients treated in the same year, in the same ICU are more likely to have similarities than patients across different years and different ICUs. Multilevel models have been widely used in education for examining the performance across schools[267]. The use of multilevel modelling in health services research has also grown with the increasing availability of patient level data. A landmark paper by Goldstein and Spiegelhalter argued in favour of using the empirical Bayes analysis, which can be related to multilevel modelling in terms of sequential analysis of variation, to make institutional comparisons in terms of utilisation and mortality across health care settings[268].

ii. Empirical model

ii.a ICU volume as quartiles

As it is used commonly in the literature, we first consider volume in terms of quartiles. To account for the multi-level structure of the data, a 3-level random intercept model is specified as follows:

$$logit = \left\{Pr\left(y_{ijt} = 1 | x_{ijt}, u_{jt}^{(2)}, u_{j}^{(3)}\right)\right\}$$

$$= \beta_0 + \beta_1 ICU\ VOLUME\_Q2_{\ ijt} + \beta_2 ICU\ VOLUME\ \_Q3_{ijt}$$

$$+ \beta_4 ICU\ VOLUME\ \_Q4_{ijt} + \ \beta' \Omega'_{\ ijt} + \beta' \Phi_j + u_{jt}^{(2)} + \ u_j^{(3)}$$

In this model y_{ijt} is the outcome variable, within acute hospital mortality and ICU volume is categorised into quartiles using quartile 1 ($ICU\ VOLUME\_Q1$) as the reference category. Here Ω'_{ijt} is a vector of all patient- level covariates, Φ'_j is a vector for time invariant ICU-level covariates, $u_{jt}^{(2)} \sim N(0, \psi^{(2)}$ is the random intercept for year within the ICU and $u_j^{(3)} \sim N(0, \psi^{(3)})$ is the random intercept varying over ICUs. The random intercept $u_{jt}^{(2)}$ and $u_j^{(3)}$ are assumed to be independent of any covariates (exogenous).

ii.b ICU volume as restricted cubic splines

Categorisation of ICU volume into quartiles assumes that all values within the quartile have the same relationship with mortality. An alternative modelling strategy is to allow for a non-linear effect of ICU volume on mortality. To do this we fitted a restricted cubic splines of sepsis volume and assessed the model with a likelihood ratio test and information criteria. Cubic splines are defined as piecewise-polynomial line segments across the distribution of the ICU volume variable, x [269]. The splines are polynomials within intervals of the volume variable, that connect each segment of the distribution. A linear spline is a set of line segments that divides the ICU volume at intervals a, b and c referred to as knots (k). Cubic polynomials allow for fitting of non-linear curves. The cubic splines are made to join at the knots by restricting the first and second derivative of the function to agree at the knots (i.e., there should be no gap in the spline curve). The restricted cubic spline has the further restriction of being linear before the first knot and after the last knot. The regression coefficients determine the shape of the curve and are considered shape parameters. Standard statistical tests can determine if the coefficients are equal to zero i.e., whether or not there is an association between volume and mortality. It is usual to present the results of restricted cubic splines

graphically with confidence intervals [182]. There are several ways to compare model fit. Unlike fractional polynomials, there is no formal selection procedure for deciding on the best fitting restricted cubic spline model. We used information criteria (AIC and BIC) to compare non-nested models and likelihood ratio tests to compare nested models. All such tests lend support to our preferred specification A further specification of volume as a fractional polynomial is included as a sensitivity analysis.

iii. Fixed versus random effects

The choice between fixed and random effects specification should be based on the perspective of the investigator. Random effects assume the ICUs are a sample population randomly drawn from a common population, with the inference given to any omitted variables assumed to be related to the whole population. Fixed effects draw inference purely from the effects analysed within the sample. Our interest lies in making inference which supports generalisation to the underlying population of ICUs, and our preferred specification is a random effect one. We do test this choice through application of a Hausman test for misspecification: essentially testing whether the differences in individual effect can be attributed to *chance. The results* are included in eTable 6 and confirm the appropriateness of using a random effects model.

Time fixed effects would capture all unmeasured ICU or patient characteristics within a year. This assumes that there is between-ICU correlation within year.

Instead, the time random effects allow for correlation for the same ICU across years. A reason for this might be that the management of sepsis within each ICU is likely to be more similar over time than management across ICUs in the same time period. Hence, we believe that use of random effects to be a less restrictive and closer to reality than using fixed effects.

iv. Interpretation of the random effects

Random effects can be related to the shrinkage estimates in the empirical Bayes literture, as if there is small random intercept variance it can be thought of as an informative prior. Of course this could be counterbalanced by a large level-1 residual variance which would represent uniformative data or a small (ICU) cluster size which would reflect an uninformative cluster.

b) Control variables

To control for confounding factors the models take into account various patient-level and ICU level characteristics. We used E-values to explore the potential effect of missing control variables. This is described in more detail as a sensitivity analysis later in the appendix.

2. Descriptive statistics

a. Patient characteristics

We categorise age into quartiles and include dummy variables for gender. Ethnicity is categorised into White, Asian, Black, and Mixed/Other. Co-morbidities are a set of dummy

variables for the presence of very severe cardiovascular disease, severe respiratory disease, end stage kidney disease, severe liver disease, metastatic cancer and haematological malignancy. Level of dependency was categorised as being fully independent, requiring some assistance or being fully dependant on assistance. Usual residence prior to hospitalisation was categorised into home, a non-health related institution, a health-related institution such as a nursing home or hospice and no fixed address. Homelessness and residence in a health care institution prior to hospitalisation has been associated with higher mortality from sepsis even after adjustment for comorbidities and disease severity[270, 271]. Socioeconomic status was described using the Index of Multiple Deprivation (2011), categorized into quintiles. The IMD measure relative levels of deprivation in 32,844 small areas or neighbourhoods in England, by combining information about the residents´ income, employment, education, skills and training, health and disability, crime, barriers to housing and services, and environmental conditions. We categorise the index into quintiles, with IMD=1 indicating the most deprived neighbourhoods[272].

The type of admission was categorised as medical, elective surgery and emergency surgery. The severity of illness was quantified by the APACHE II and ICNARC scores. The APACHE II (Acute Physiology, Age, Chronic Health Evaluation) model is a severity of critical illness model developed in 1985 [273] . The score ranges from 0 to 71, with a higher score predicting a higher mortality. The ICNARC score was first published in 2007 and then recalibrated in 2014, 2015 and in 2018[274]. The score is from 0 to 100 based on weightings for deviations from normal in the twelve physiological parameters during the first 24 hours in the ICU as well as additional weights from age, indication for admission, surgical urgency, source of admission and cardiopulmonary resuscitation prior to admission[123, 178, 274]. The type of treatment

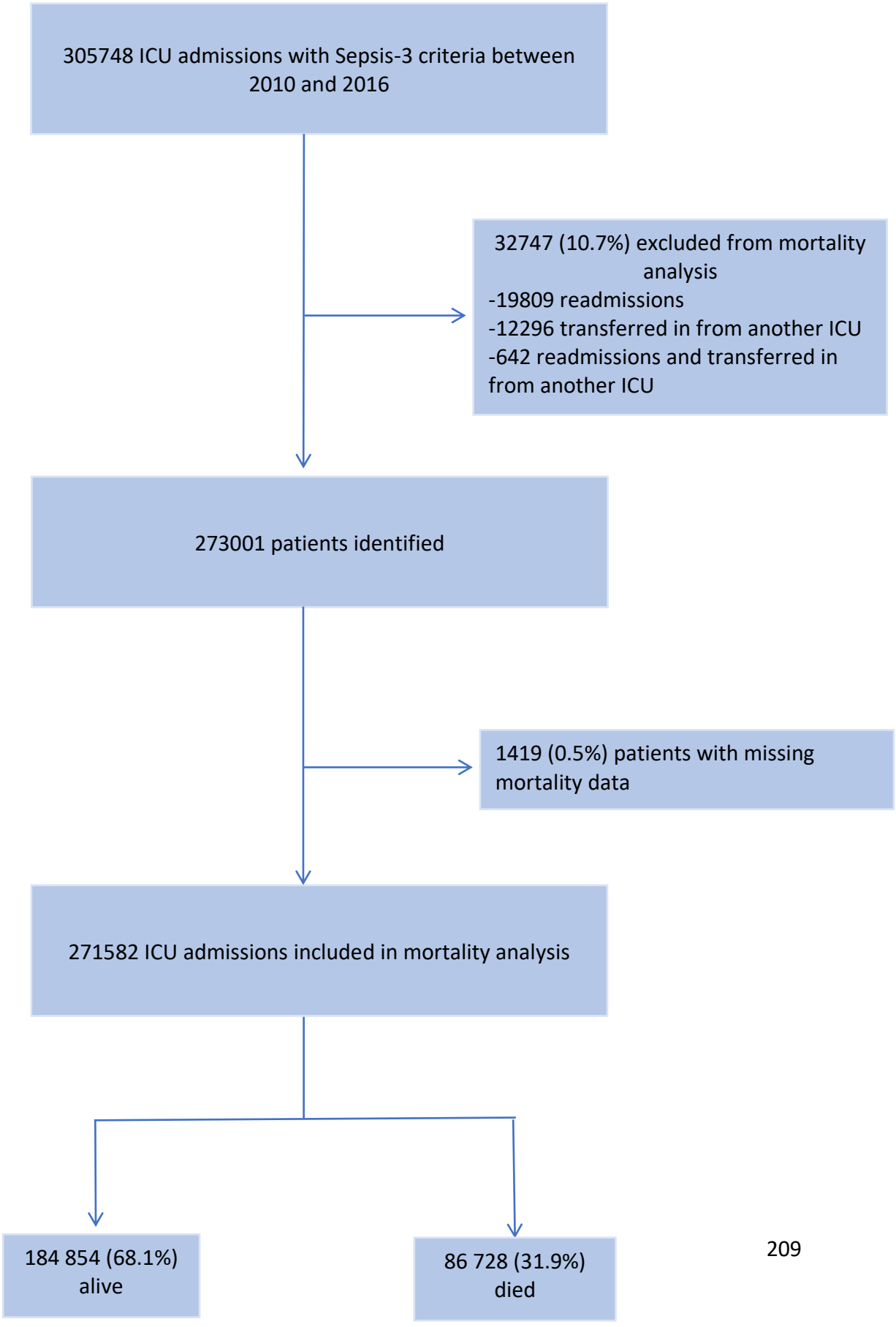
received in the ICU include renal replacement therapy, mechanical ventilation, and circulatory support. Septic shock is a high severity sub-category of sepsis with an expected unadjusted mortality of about 41%[275]. ICU length of stay is described in hours. Hospital length of stay is described in days and includes the days of hospitalisation prior to ICU admission.

b. ICU characteristics

ICU specific characteristics include the academic affiliation, number of ICU beds, annual bed occupancy, sepsis caseload and total caseload.

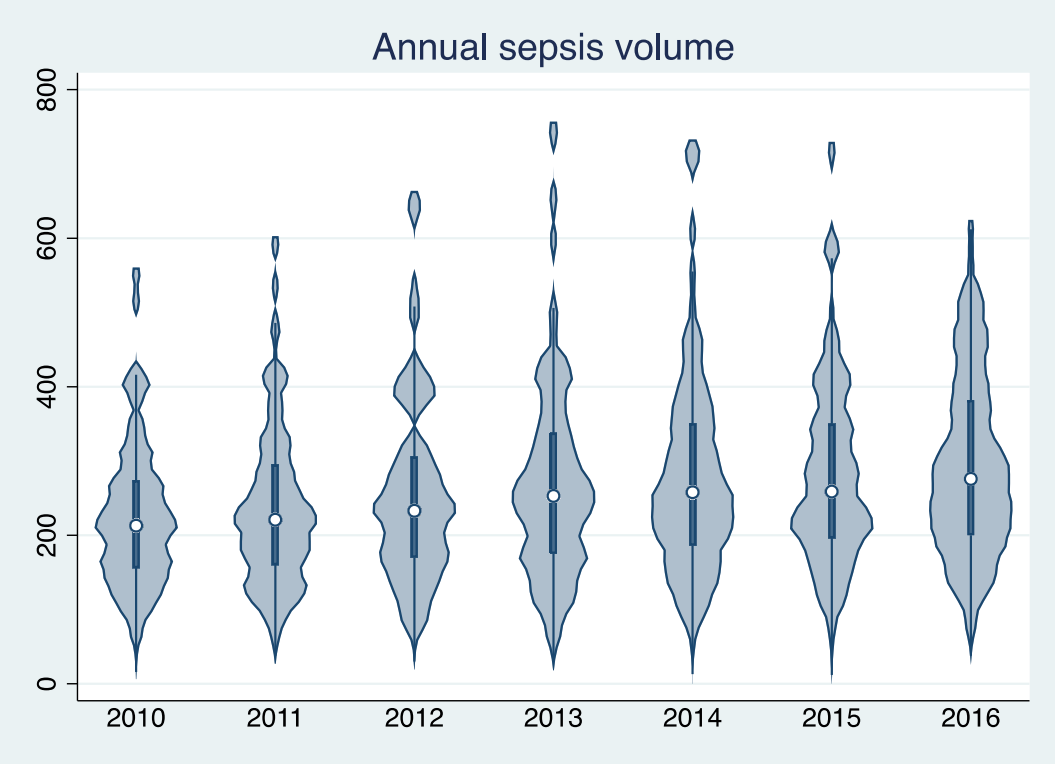
3. Additional data

a. eFigure 1. Study patient flow



b. eFigure 2. Violin plot of annual ICU volume and annual ICU sepsis volume between 2010 and 2016

Violin plots are a method for displaying the distribution of the data, combining a boxplot and a kernel density plot. The information contained is the median (white dot), the interquartile range (bar in the centre of the violin). The wider areas of the violin correspond to increased frequency at that y-value. The thin lines that extend from the bar are the lower and upper adjacent values, defined as first quartile -1.5 IQR and third quartile +1.5 IQR. The median total annual ICU volume increased from a caseload of 666 [IQR 459-911) in 2010 to a caseload of 827 [IQR 598-1288] in 2016. The median annual sepsis volume increased from 213 [IQR 156-274] in 2010 to 276 [IQR 200-382].



d. eTable 1. Characteristics of ICU admissions with sepsis between 2010 and 2016

Variable	Total Sepsis	Year							
		2010	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**P value**
	N=305738	N=36079	N=38643	N=41075	N=43100	N=46835	N=48685	N=51321	
Age in years									<0.001
<54	77082(25.2)	9540(26.4)	9950(25.8)	10000(24.3)	10411(24.1)	11464(24.5)	12237(25.1)	13480(26.2)	
54-66	77258(26.3)	9372(26.0)	9860(25.5)	10432(25.4)	10902(25.3)	11711(25.0)	12103(24.9)	12878(25.1)	
67-76	79820(26.1)	8967(24.9)	9952(25.8)	10562(25.7)	11284(26.1)	12455(26.6)	12920(26.4)	13680(26.7)	
>76	71578(23.4)	8200(22.7)	8881(23.0)	10081(24.5)	10503(24.3)	11205(23.9)	11425(23.5)	11283(22.0)	
Male gender	167895(54.9)	19638(54.4)	21225(54.9)	22435(54.6)	23541(54.6)	25699(55.3)	26922(55.3)	28435(55.4)	0.023
Ethnicity									<0.001
White	277787(90.9)	33153(92.0)	35534(92.1)	37721(92.0)	29429(91.5)	42474(90.7)	43718(89.8)	45758(89.2)	
Asian	10723(3.5)	1168(3.2)	1145(3.0)	1304(3.2)	1465(3.4)	1742(3.7)	1890(3.9)	2009(3.9)	
Black	6192(2.0)	663(1.8)	711(1.8)	773(1.9)	851(2.0)	1010(2.2)	1041(2.1)	1143(2.2)	
Mixed/other	10855(3.6)	1046(2.9)	1168(3.0)	1227(3.0)	1355(3.1)	1609(3.4)	2038(4.2)	2413(4.7)	
Comorbidities									
Cardiac	5375(1.8)	610(1.7)	616(1.6)	740(1.8)	708(1.7)	869(1.9)	894(1.8)	938(1.9)	0.012
Respiratory	13500(4.4)	1690(4.7)	1700(4.4)	1802(4.4)	1887(4.4)	2048(4.4)	2191(4.5)	2182(4.3)	0.092
ESRD*	5997(2.0)	669(1.9)	717(1.9)	728(1.8)	816(1.9)	925(2.0)	1049(2.2)	1093(2.1)	<0.001
Liver	7049(2.3)	796(2.2)	880(2.3)	945(2.3)	970(2.3)	1119(2.4)	1143(2.4)	1196(2.3)	0.652
Haematological malignancy	10828(3.6)	1194(3.2)	1416(4.0)	1426(3.5)	1474(3.43)	1713(3.7)	1738(3.6)	1867(3.7)	0.042
Metastatic cancer	7492(2.5)	729(2.0)	899(2.3)	935(2.3)	1035(2.4)	1226(2.6)	1281(2.6)	1387(2.7)	<0.001
Level of dependency prior to acute hospitalization									<0.001
Independent	208339(68.4)	25142(70.0)	26998(70.2)	28091(68.7)	29319(68.3)	31240(67.0)	32797(67.6)	34752(67.9)	
Some assistance	90536(29.7)	10147(28.2)	10848(28.2)	12065(29.5)	12843(29.9)	14564(31.2)	14748(30.4)	15321(30.0)	
Total dependence	5677(1.9)	620(1.7)	621(1.6)	748(1.8)	775(1.8)	839(1.8)	994(2.1)	1080(2.1)	
Usual residence prior to hospitalization									0.755
Home	296726(97.1)	35017(97.0)	37507(97.1)	39836(97.0)	41785(97.0)	45406(97.0)	47309(97.2)	49866(97.2)	
Work or non-health related institution	638(0.2)	75(0.2)	84(0.2)	90(0.2)	88(0.2)	94(0.2)	102(0.2)	105(0.2)	
Nursing home, hospice or health related institution	7327(2.4)	857(2.3)	914(2.4)	1002(2.4)	1085(2.5)	1169(2.5)	1126(2.3)	1174(2.3)	
No fixed address	1057(0.4)	134(0.4)	139(0.4)	149(0.4)	142(0.3)	166(0.4)	150(0.3)	177(0.3)	
IMD** quintile									0.075
I	77808(25.6)	9321(26.0)	9894(25.8)	10358(25.3)	11141(26.0)	12016(25.9)	12149(25.2)	12929(25.4)	
II	65377(21.5)	7608(21.2)	8357(21.8)	8833(21.6)	9146(21.3)	9922(21.3)	10540(21.8)	10971(21.6)	
III	59516(19.9)	7125(19.9)	7535(19.6)	8058(19.7)	8254(19.2)	9073(19.5)	9569(19.8)	9902(19.5)	
IV	53094(17.5)	6183(17.2)	6593(17.2)	7273(17.8)	7473(17.4)	8205(17.7)	8376(17.4)	8991(17.7)	
V	47809(15.8)	5663(15.8)	6025(15.7)	6360(15.9)	6831(15.9)	7255(15.6)	7620(15.8)	8055(15.8)	

APACHE II score, mean(95% CI)	18.4(18.3-18.4)	18.6(18.5-18.7)	18.6(18.5-18.7)	18.5(18.4-18.6)	18.4(18.3-18.5)	18.4(18.3-18.4)	18.2(18.2-18.3)	18.0(17.9-18.0)	<0.001
ICNARC score, mean(95% CI)	20.6(20.6-20.7)	21.4(21.2-21.4)	21.1(21.0-21.2)	20.9(20.8-21.0)	20.8(20.7-20.9)	20.5(20.4-20.5)	20.3(20.3-20.4)	20.0(19.9-20.1)	<0.001
ICNARC predicted probability of death, mean% (95% CI)	29.6(29.5-29.7)	30.9(30.6-31.2)	30.6(30.2-30.8)	30.4(30.1-30.6)	30.2(30.0-30.4)	29.4(29.1-29.7)	28.9(28.6-29.0)	27.8(27.5-28.0)	<0.001
Renal failure in the first 24 hours	25731(8.6)	3348(9.5)	3588(9.5)	3593(8.9)	3657(8.6)	3826(8.3)	3888(8.1)	3831(7.6)	<0.001
Mechanical ventilation	16600(54.3)	22140(61.3)	22688(58.7)	23091(5.2)	23534(54.6)	24440(52.2)	24644(50.6)	25469(48.6)	<0.001
Septic shock	58911(19.3)	7740(21.5)	8008(20.7)	8272(20.1)	8426(19.6)	8813(18.8)	9015(18.5)	8637(16.8)	<0.001
Non-surgical	227533(74.4)	26545(73.6)	28470(73.7)	30343(73.9	32099(74.5)	34890(74.5)	36359(74.7)	38827(75.7)	<0.001
ICU length of stay in hours, median IQR	93(43-198)	94(42-210)	92(42-204)	92(42-196)	92(42-1)95)	93(44-195)	93(43-196)	96(45-195)	<0.001
Hospital length of stay in days, median IQR	15(7-32)	26(8-34)	16(8-33)	16(7-32)	16(7-32)	15(8-31)	15(7-30)	15(7-29)	<0.001
Transferred in	12938(4.2)	1707(4.7)	1801(4.7)	1761(4.3)	1838(4.3)	1887(4.0)	1924(4.0)	2020(3.9)	<0.001
Readmission	20451(6.7)	2404(6.7)	2863(7.4)	2905(7.1)	2966(6.9)	3114(6.7)	3115(6.4)	3084(6.0)	<0.001
Unadjusted ICU mortality	62277(22.8)	8121(25.3)	8208(24.1)	8598(23.5)	8830(23.0)	9172(21.9)	9543(21.8)	9805(21.2)	<0.001
Unadjusted Hospital mortality	86728(31.9)	11326(35.6)	11413(33.7)	11945(32.9)	12325(32.2)	12831(30.7)	13331(30.6)	13557(29.5)	<0.001

*ESRD= end stage renal disease

**IMD= Index of Multiple Deprivation 2011

For categorical variables a Chi squared test was used. The null hypothesis was that there is no difference in the distribution of responses to the outcome across comparison groups. For continuous variables we use the ANOVA to analyse the differences in means between groups.

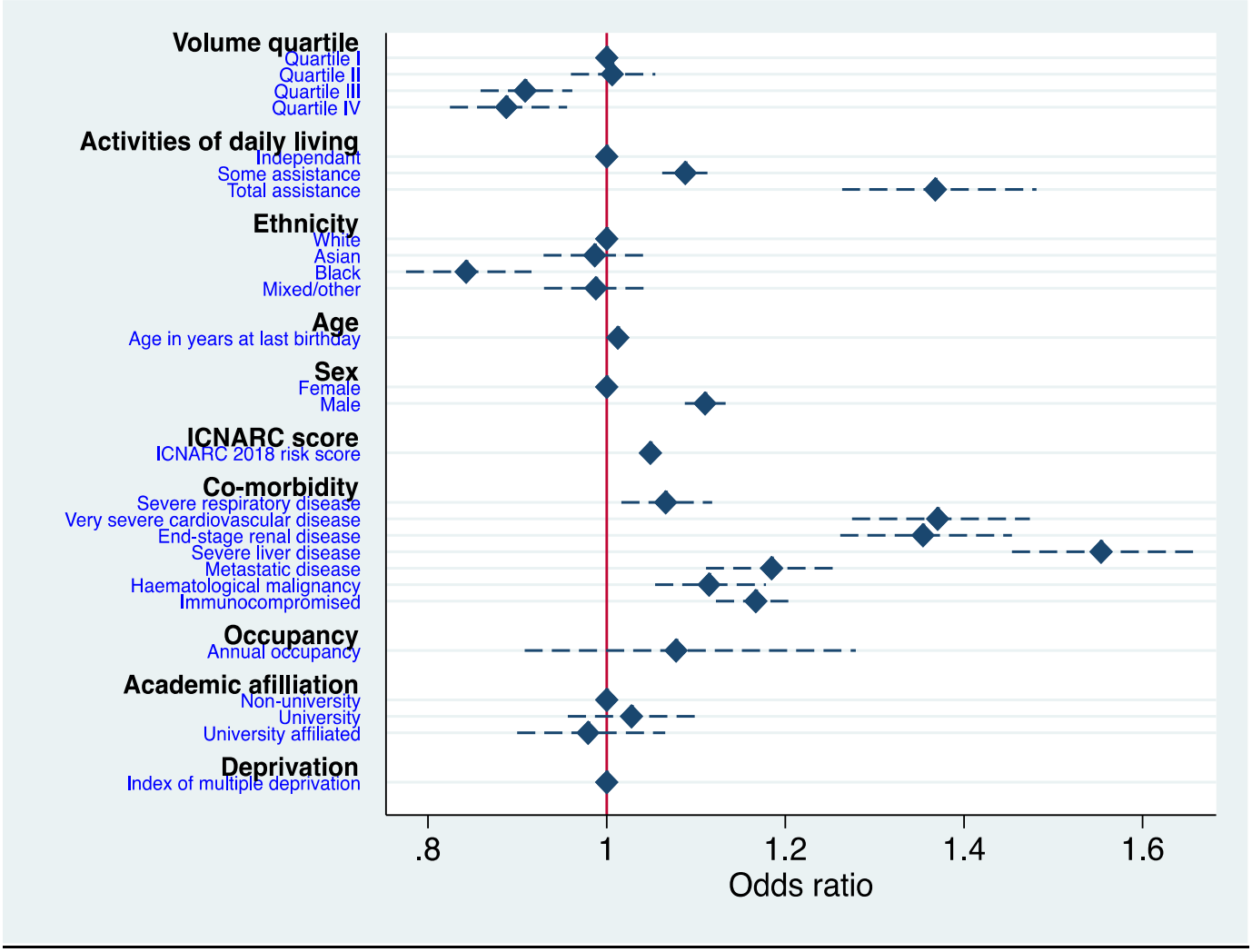
eTable 2 ICU characteristics between 2010 and 2016

Variable	**Number**								
		2010	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**p-value**
ICU beds, median (IQR)	13(9-18)	11(8-16)	12(9-17)	12(9-17)	13(9-19)	13(9-19)	14(10-12)	15(10-20)	<0.001
Occupancy %, median (IQR)	73.5(67.9-79.5)	71.0(65.1-76-2)	71.6(64.7-77.7)	72.6(67.3-79.1)	74.2(68.7-80.1)	73.4(68.2-79.7)	75.0(69.8-80.0)	75.8(80.5-80.8)	<0.001
Sepsis volume, median (IQR)	242(177-334)	213(156-274)	221(160-295)	233(170-306)	253(175-338)	258(186-351)	259(196-350)	276(200-382)	<0.001
Non-sepsis volume, median (IQR)	497(346-747)	432(295-607)	472(325-684)	493(346-722)	496(340-774)	519(368-819)	527(377-871)	555(382-819)	<0.001
Total volume, median (IQR)	742(533-1087)	666(459-911)	691(492-994)	729(519-1085)	732(519-1175)	774(577-1229)	785(596-1254)	827(598-1288)	<0.001

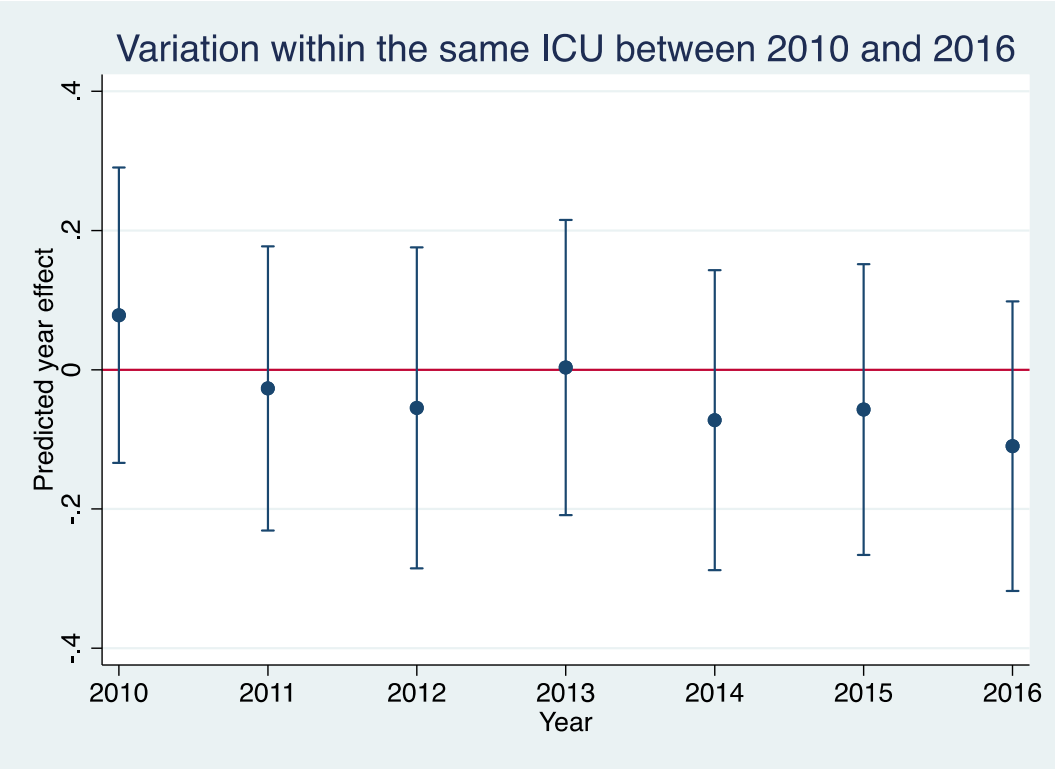
For categorical variables a Chi squared test was used. The null hypothesis was that there is no difference in the distribution of responses to the outcome across comparison groups. For continuous variables we use the ANOVA to analyse the differences in means between groups.

4. Additional results

a. eFigure 3. Odds ratio for mortality with volume as quartiles.



c._eFigure 4. The within-ICU variation across between 2010 and 2016.

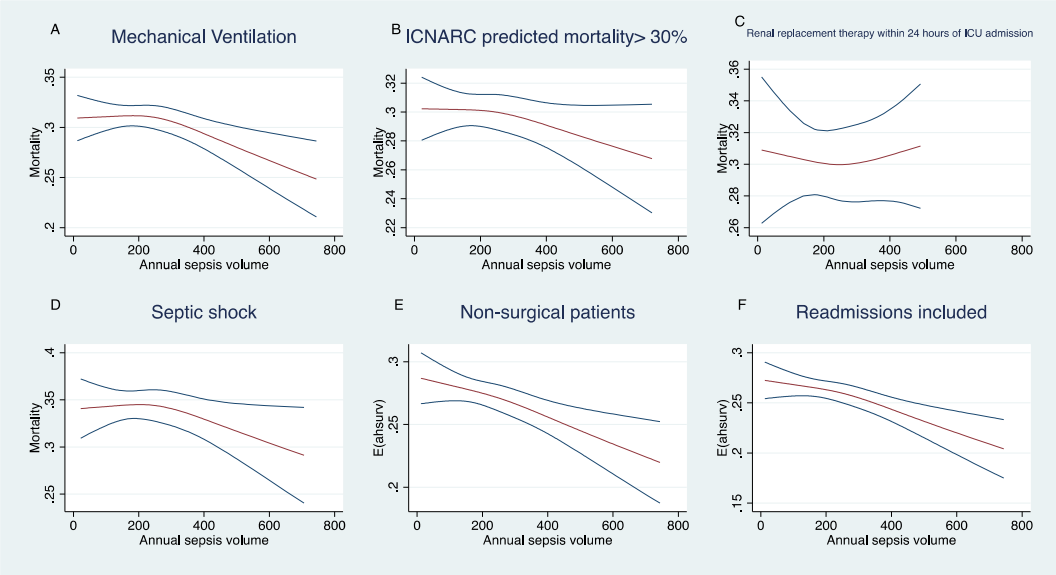


The figure represents variation in mortality not explained by measurable characteristics within the same ICU over the study period.

d. eTable 3. Testing the statistical significance between models of increasing complexity with volume defined as a restricted cubic spline.

Sensitivity Analysis	Chi2	df	p-value
Sepsis expressed cubic splines			
3 knots	16.67	2	0.0002
4 knots	18.65	3	0.0003
5 knots	18.71	4	0.0009
6 knots	20.83	5	0.0009

e. eFigure 5. Subgroup analysis: (A) Mechanical ventilation (B) ICNARC predicted mortality >30% (C) Renal replacement therapy within 24 hours of admission (D) Septic shock (E) Non-surgical patients (F) Readmissions included



There was no enhanced reduction in mortality from ICU volume in patients with more severe illness as characterised by the subgroup of patients with mechanical ventilation, ICNARC predicted mortality >30%, renal replacement therapy within 24 ours of admission and septic shock. The reduction of mortality was consistent in the subgroup of patients with sepsis not requiring surgical procedures. The inclusion of the outcome from subsequent ICU readmission episodes did not impact the observed volume-outcome relationship.

5. Sensitivity analysis

a. Fractional polynomial and selection procedure

i. eTable 4. Output from model fitting procedure for fractional polynomial showing deviance and powers of model

Model	Sepsis volume	df	Deviance	Deviance difference	P-value	Powers
Model 1	Omitted	0	232492.13	16968	<0.0001	
	Linear	1	232475.76	0.596	0.440	1
	M=1	2	232475.16	0.000	-	2
Model 2	Omitted	0	232492.13	17.735	<0.001	
	Linear	1	232475.76	2.531	0.282	1
	M=1	2	232475.16	1.934	0.164	2
	M=2	3	232473.23	0.00	-	3 3
Model 3	Omitted	0	232492.13	233.253	0.001	
	Linear	1	232475.76	6.899	0.230	1
	M=1	2	232475.16	6.069	0.179	2
	M=2	3	232473.23	4.369	0.226	3 3
	M=3	6	232468.88	0.00	-	-2 0 0

The general formulation of fractional polynomials is:

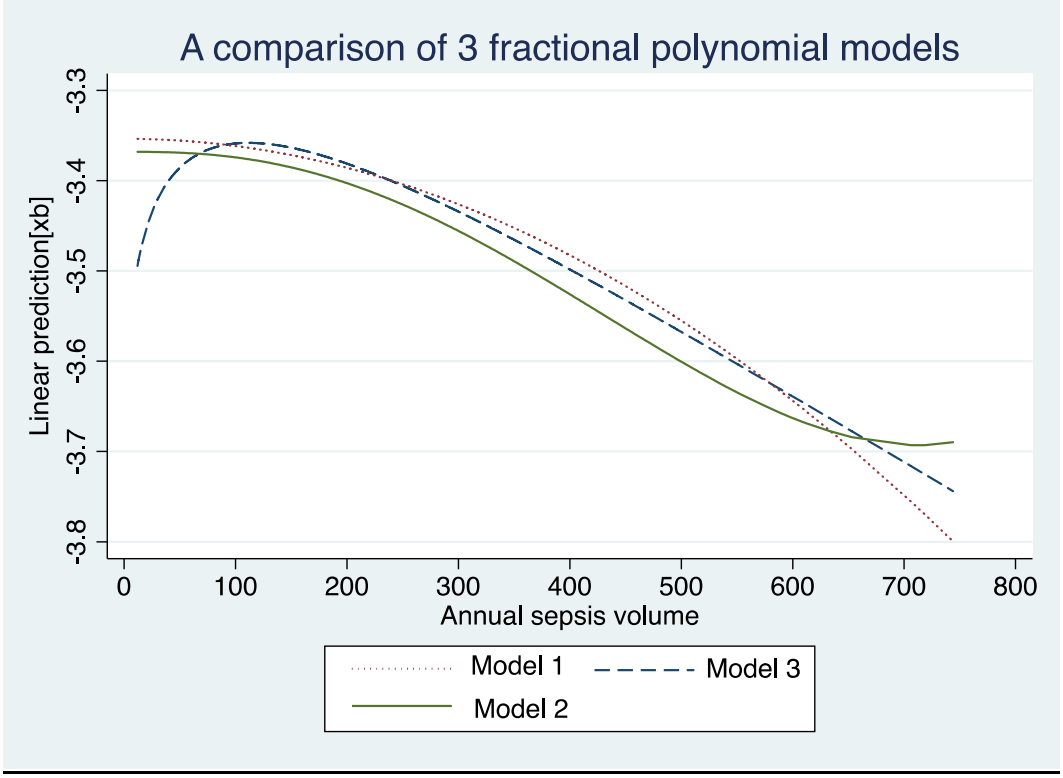
$$x^{(p_1,p_2,\ldots,p_m)'}\beta = \beta_0 + \beta_1 x^{(p_1)} + \beta_2 x^{(p_2)} + \cdots + \beta_m x^{(p_m)}$$

x^0 is interpreted as $\ln(x)$ and repeat powers are multiplied by $\ln(x)$, where $x$indicates ICU volume.

The models are constructed iteratively. In Model 1, the highest power is 2 thus the fractional polynomial with the dimension $M = 1$ would be $\beta_1 x^2$. The linear model included as part of the selection procedure would be $\beta_1 x$. In Model 2 the highest FP where the dimension $M = 2$ would be $\beta_1 x^3 + \beta_2 x^3 \ln(x)$ and included the preceding iterations of Model 1 for comparison. The highest power in Model 3 where the dimension $M = 3$ is $\beta_1 x^{-2} + \beta_2 \ln(x) + \beta_3 \ln(x) \ln(x)$.

The selection procedure compares the highest FP model with the model that omits the x variable. If significant, the selection procedure, then compares the FP model with the linear model. If this in turn is significant, the selection procedure then compares the most complex FP model with the next model down the list of iterations. A more complex model is chosen if that model fits the data better based on significance criteria.

ii. eFigure 6. Functional form of fractional polynomial models showing Model 1, Model 2 and Model 3.



The graph describes the functional form of the FP models where $Model\ 1 = \beta_1 x^2$, $Model\ 2 = \beta_1 x^3 + \beta_2 x^3 \ln(x)$ and $Model\ 3 = \beta_1 x^{-2} + \beta_2 \ln(x) + \beta_3 \ln(x) \ln(x)$. The models are in increasing complexity. The models of higher complexity show less stable results at the extremes of ICU volume [185].

iii. eTable 5. A comparison of the log likelihood and information criteria for the linear, fractional polynomial and restricted cubic spline models.

Model	df	Log likelihood	AIC	BIC
Categorical	25	-116234	232518.9	232780.4
Linear	23	-116237.9	232521.8	232762.4
Restricted cubic splines				
3 knots	24	-116237.1	232522.2	232773.3
4 knots	25	-116236.1	232522.1	232783.7
5 knots	26	-116236.0	232524.1	232796.1
6 knots	27	-116234.9	232523.9	232806.3
Fractional polynomial				
Model 1	23	-116237.6	232521.2	232761.8
Model 2	23	-116236.6	232519.2	232843.3
Model 3	25	-116234.4	232518.9	232780.4

The model that minimises both AIC and BIC is preferred but no model does this for this data. At higher sample sizes the AIC may select the model that is too complex and the BIC has a higher probability of selecting the true model[276]. As the sample size increases, the BIC offers more general consistency. Using likelihood alone as a model selection criterion in not advised as it will tend to select overly-paramaterized models[277]. The 3-knot model has the lowest BIC is the considered best fitting restricted cubic spline. In terms of FP models, Model 1 minimises the BIC. There are several competing models, one of which can be categorially declared as better than the rest.

b. E-values

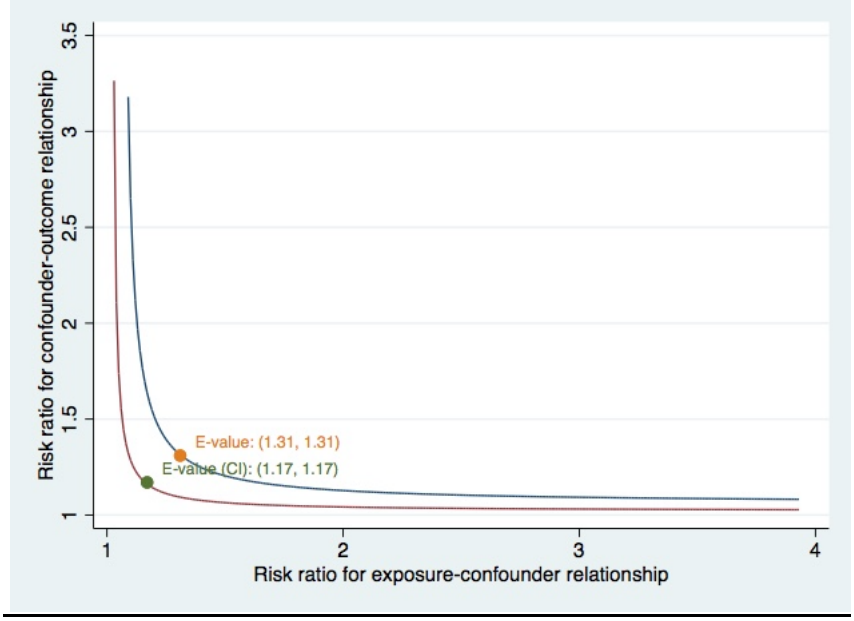
ii. Rationale

Regression analyses provide some control for measured confounding but there is potential for unmeasured confounding. Being unmeasured, it is a challenge to assess what impact this latter form confounding might have on the analyses[187]. Ding and VanderWeele have proposed a bounding factor approach to assess the sensitivity of the results to unmeasured confounders[187]. They define a bounding factor (E-value) which is the minimum strength of association on a risk-ratio scale that the unmeasured confounder would need to have a material effect on the results[188]. The E-value describes two parameters, the risk ratio for the confounder-outcome relationship and the risk ratio of the exposure confounder relationship.

There is no absolute threshold for an E-value and the reader must then assess whether the magnitude of the unmeasured confounder is plausible[188]. In this study, the OR for acute hospital mortality for quartile IV compared with quartile I of sepsis volume was 0.89 with a 95% CI of 0.82-0.96.The E-value for this point estimate is 1.31 and for the lower confidence limit is 1.17. This would mean that the observed OR could be explained by the presence of unmeasured confounding associated with both the exposure and the outcome by a risk ratio of 1.31, above and beyond measured confounders. Although a RR of 1.31 appears modest,

in the current study this seems unlikely because this would imply that the unmeasured covariates would have to be similar in magnitude to measured covariates like severe respiratory disease or haematologic malignancy. The Case Mix Program is a high-quality clinical database, and we suggest that an unmeasured confounder of this magnitude is unlikely.

ii. eFigure 7. Value of the joint minimum strength of association that an unmeasured confounder must have with both an increase in ICU sepsis volume and acute hospital mortality to explain away the volume outcome relationship, expressed as a risk ratio



The E-value describes 2 parameters, the risk ratio for the confounder-outcome relationship and the risk ratio of the exposure confounder relationship. The area above the two lines are the joint exposure confounder risk ratio and confounder outcome risk ratio that would be required to explain the observed treatment effect.

c. Checking for exogeneity ICU volume

i. Rationale

One of the assumptions of the multilevel random effects model is exogeneity- specifically in our model it would require that ICU volume is uncorrelated with the ICU-level random effect. [189]. The random effects model assumes that the within- and between ICU effects are the same and uses a weighted average of the within and between ICU effects in estimation. The p-value tests the significance of this difference. This is a regression-based alternative to the Hausman test for endogeneity of a regressor[278, 279].

Stated simply, the exogeneity assumption holds i.e., the random effects model of no correlation between the random effect and the level-1 covariates holds[279]. We describe these results as part of a robustness check to our model assumptions.

ii. eTable 6. Within and between cluster effects of ICU volume test for exogeneity.

Variable	**Coefficient**	**Standard Error**	**p-value**
Sepsis Volume			
Quartile I (W)	0.1474	0.0397	
Quartile II (W)	0.1541	0.3254	
Quartile III(W)	0.0371	0.0276	
Quartile IV(W)	*(omitted)*		
Quartile I (B)	0.0442	0.0607	
Quartile II (B)	0.0207	0.0634	
Quartile III(B)	0.0146	0.0678	
Quartile IV(B)	*(omitted)*		
Quartile I (B-W)	**0.0441**	**0.0607**	**0.467**
Quartile II (B-W)	**0.0207**	**0.0634**	**0.745**
Quartile III(B-W)	**0.0146**	**0.0678**	**0.830**
Quartile IV(B-W)	*(omitted)*		
Restricted Cubic Spline			
Spline-1 (W)	-0.0003	0.00003	
Spline-2(W)	-0.0006	0.00003	
Spline-1 (B)	-0.0003	0.0004	
Spline-2 (B)	0.0003	0.0005	
Spline-1 (B-W)	**$4.23e^{-6}$**	**0.0004**	**0.9930**
Spline-2(B-W)	**0.0009**	**0.0006**	**0.1551**

(W)= within cluster; (B)= between cluster; (B-W) = difference in between and within-cluster effects.

The lack of statistical significance, as highlighted by the highlighted p values, in the between- and within-cluster effects for ICU volume imply a lack of correlation in the ICU volume and the ICU random effect, in support of the assumption that ICU volume is exogenous.

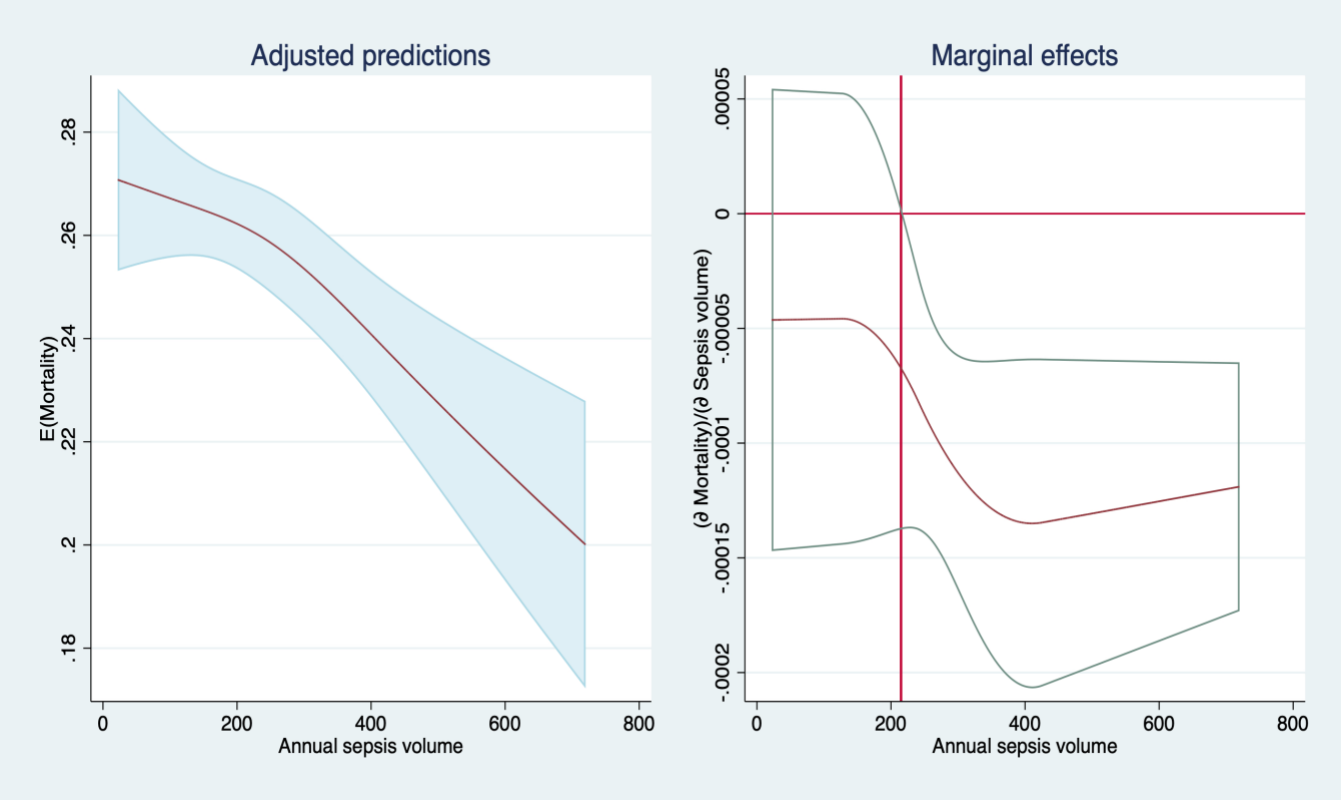
d. Checking the robustness of the results to extremes of annual sepsis volume.

In this analysis we excluded ICUs with <20 sepsis patients per year and ICUs with >740 sepsis patients per year. This resulted in excluding 801 patients from the analysis.

i. eTable7 describing quartiles after excluding extremes of annual sepsis volume with probability of acute hospital mortality.

Quartile	Range	Mortality	95% CI
I	22-177	31.2 %	30.6% - 31.8%
II	178-242	31.3%	30.7% -31.9%
III	243-331	29.8%	29.2-30.4%
IV	332-720	29.3%	28.5%-30.2%

ii. eFigure 8. The predicted mortality and marginal effects for the analysis excluding extremes of annual sepsis volume



This analysis is consistent with the primary analysis. We find an identical threshold for the favorable volume- outcome relationship of 215 patients. This suggests that the primary analysis was not biased by the extremes of ICU sepsis volume

Appendix three: The association between ICU specialisation and mortality in the UK

Table of Contents

1. The ICNARC Coding Method (ICM)

The ICM is a structured, numerical method for uniformly coding critical care patients' reasons for admission; assigning unique numerical codes to all conditions/disease processes which commonly necessitate patients' admission to critical care.

ICM Structure:

The resulting hierarchical structure is composed of the following five tiers:

Type: has the admission had surgery for the condition being coded?

System: which body system is involved?

Site: which anatomical site(s) within the body system are involved?

Process: what physiological or pathological process is involved?

Condition: what is the condition being coded?

Each tier of the ICM code is represented by a number, reflecting the chosen option for that tier. By working through the tiers, the ICM constructs a unique numerical code which, when complete, relates to a specific condition.

The code is defined and built as follows:

Type: has the admission had surgery for the condition being coded?

The ICM code begins by distinguishing whether a patient has had surgery for the condition being coded. This is done by selecting either 1, the patient had surgery for the condition, or 2, they did not.

System: which body system is involved?

The ICM codes the body system affected, and each body system has a unique code, e.g. the respiratory system is represented by the number 1.

Site: which anatomical site(s) within the body system are involved?

The site tier reflects the specific site within the selected body system, e.g. lungs are always coded using 4, within the respiratory system.

Process: what physiological or pathological process is involved?

The process tier reflects the physiological or pathological process which is affecting the specific site. Each process code is the same across every system, e.g. infection will always be coded as 27.

Condition: what is the condition being coded?

The final tier of the ICM identifies the specific condition, e.g. Bacterial pneumonia is coded as:

Type: Non-surgical - 2

System: Respiratory - 1

Site: Lungs – 4

Process: Infection – 27

Condition: Bacterial pneumonia – 1

The full code for bacterial pneumonia is therefore 2.1.4.27.1

2. Alternate measures of specialisation

2.1. The exponent (E)

The exponent (E) in share is calculated as the exponent of the S_{kjm}

$$E = e^{S_{kjm}}$$

This is an upward sloping function where $e^0 = 1$ when the $S_{kjm} = 0$ and $e^1 = 2.718$ when $S_{kjm} = 1$. This specification addresses the potential non-linear relationship between specialisation and mortality.

2.2. Relative difference (RD)

RD can be described as the difference between the share S_{kjm} and the reference share.

$$RD = \left[\frac{n_{kjm}}{N_{jm}}\right] - \left[\frac{n_{km}}{N_m}\right]$$

The RD will be zero when the distribution of patients with ICU j is identical to the reference distribution. Values greater than zero imply higher levels of specialisation than the average ICU and negative values indicate ICUs with less than average levels of specialisation.

3. Hierarchical model

Hierarchical models overcome both the atomistic fallacy of individual risk factor epidemiology and the ecological fallacy of aggregated data[244]. This approach aims to answer the question of: do higher level contexts such as specialisation impact on individual patients and in what magnitude?

To account for the hierarchical structure of the data, patients i (level 1) within the same month m (level2) are within the ICU j (level 3), a 3-level hierarchical logistic regression model is specified as follows:

$$logit = \left\{Pr\left(y_{ijm} = 1|X'_{ijm}, u_{jm}^{(2)}, u_j^{(3)}\right)\right\}$$

$$= \beta_0 + \beta_1 S_{kjm} + \beta_2 Vol_{jm} + \beta_3 ICNARC_{ijm} + \beta' \Omega'_{ijm} + \beta' \Phi_j + u_{jm}^{(2)} + u_j^{(3)}$$

The outcome variable, acute hospital mortality, y_{ijm} given a set of covariates, X'. The exposure is level of specialisation for diagnostic group k in the ICU S_{kjm} for specialisation. In the model , , X' includes monthly caseload volume Vol_{jym} , and patient level severity of illness as measured by the ICNARC$_{2018}$ score, $ICNARC_{ijm}$. Here Ω'_{ijm} is a vector of all other patient-level covariates, Φ'_j is a vector for other ICU-level covariates, $u_{jm}^{(2)} \sim N(0, \psi^{(2)})$ is the random intercept for the month with the ICU j and $u_j^{(3)} \sim N(0, \psi^{(3)})$ is the random intercept varying over ICUs. The random intercepts $u_{jm}^{(2)}$ and $u_j^{(3)}$ are assumed to be independent of any covariates (exogenous). The random intercept $u_j^{(3)}$ represents unexplained variation in mortality across ICUs, and the random intercept $u_{jm}^{(2)}$ represents unexplained variation between months in the same ICU. The hierarchical model accounts

for the clustering of the data within months and ICUs and provide a more accurate estimate of standard errors for clustered data than fixed effects models[245].

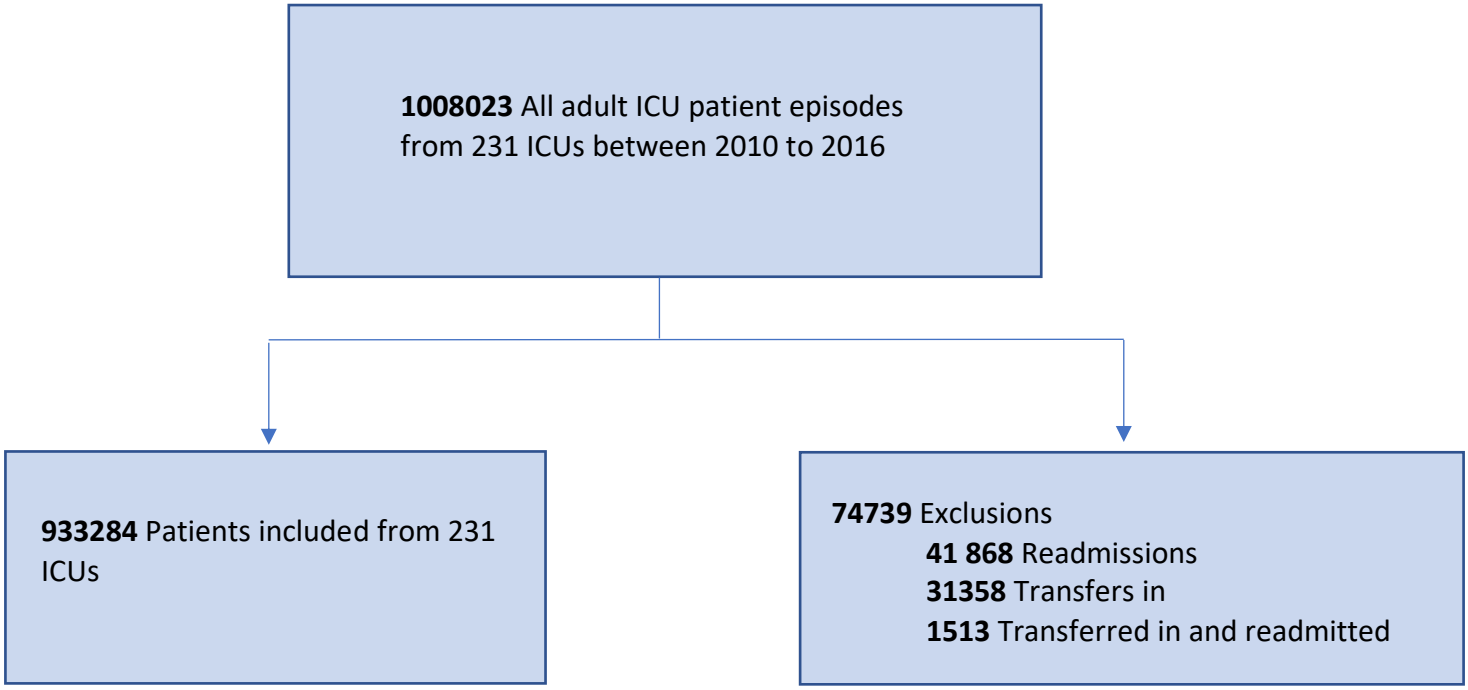
Hierarchical models have lower units of aggregation nested within higher units.

The hierarchical model framework can overcome a group level omitted variable. The addition of a group level average covariate $\overline{X_j}$ will absorb all correlations between X and the group level random effects. This approach is equivalent to a fixed effects approach.

Hierarchical models can also test for heterogeneity in coefficients across groups.

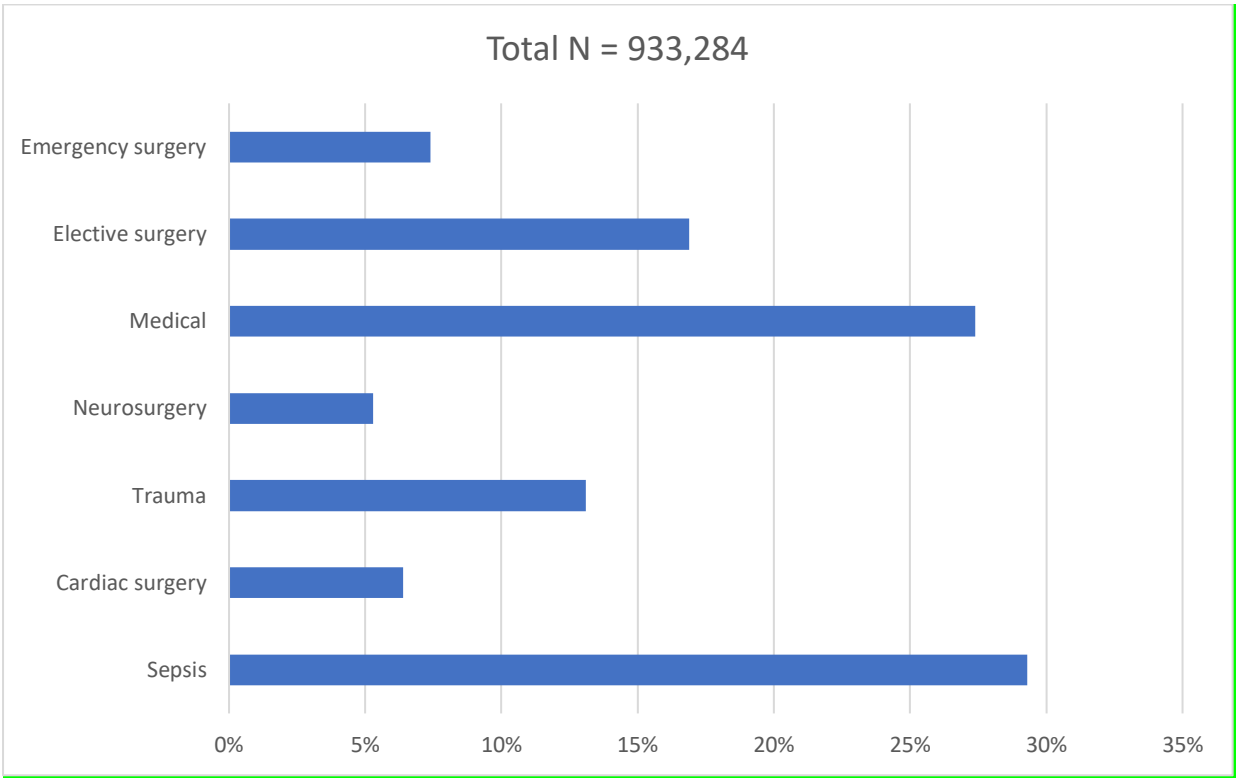
Hierarchical models allow a framework to account for clustering, omitted group-level variables and heterogeneity of coefficients between ICUs.

4. eFigure 1 Patient flows.



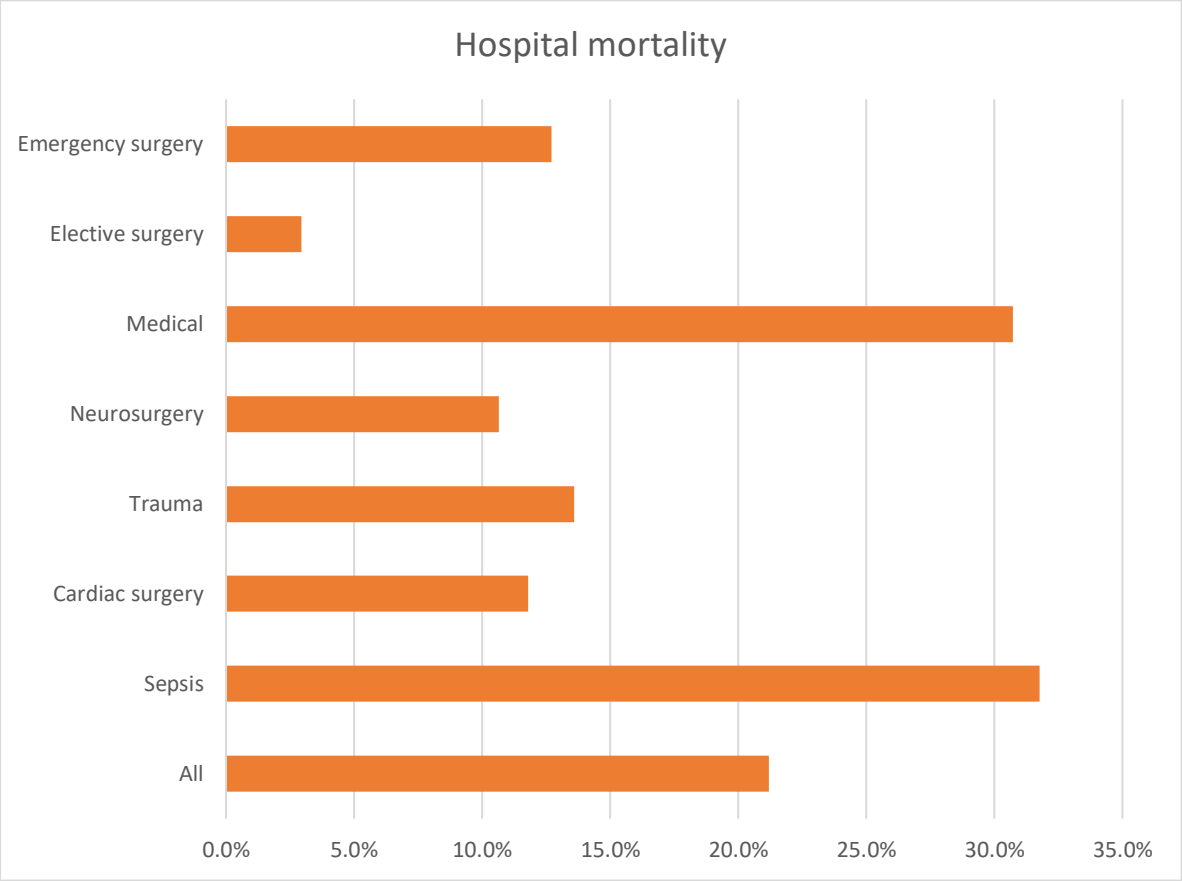
5. Descriptive data

5.1. eFigure 2. Distribution of specialist patients.



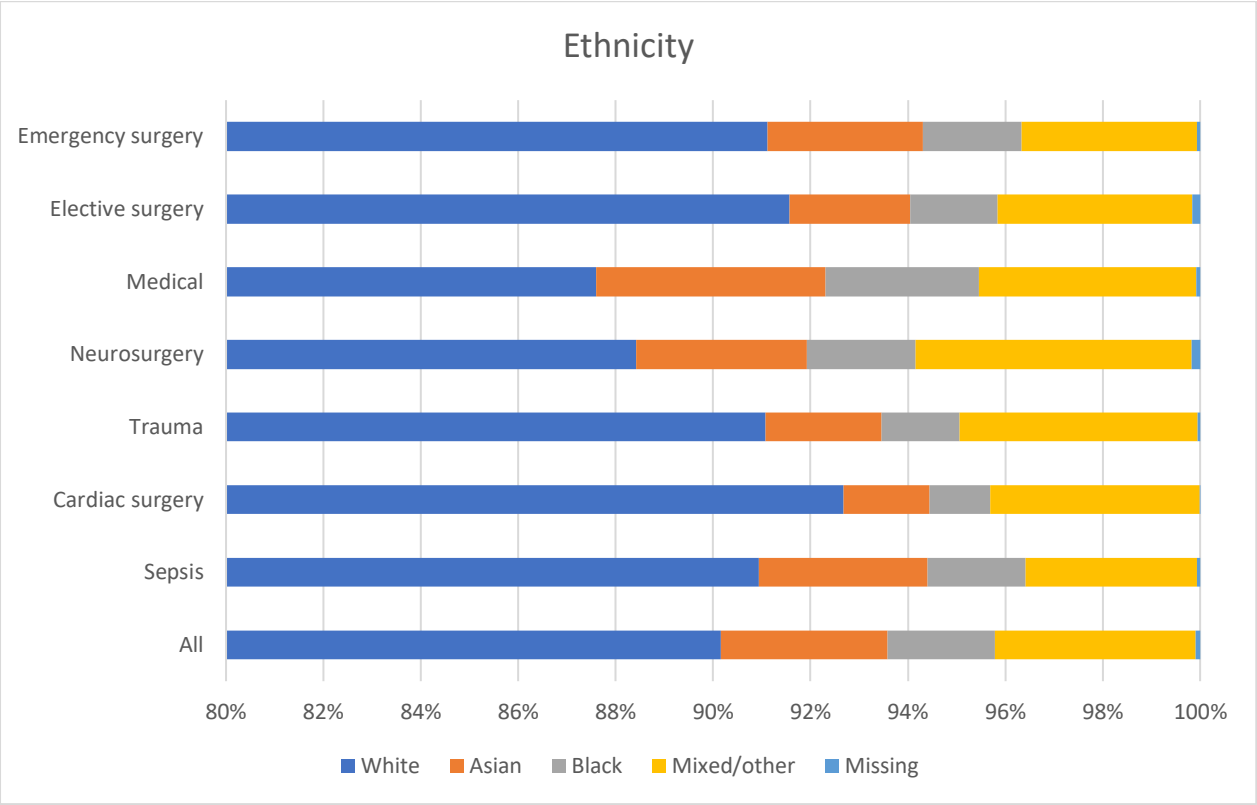
There was a total of 933284 patients. Sepsis was the most common diagnostic group and accounted or 29% of patients. Cardiac and neurosurgery patients were the smallest share of patients (6.4% and 5.3% respectively).

5.2. eFigure 3. Unadjusted mortality across specialties.



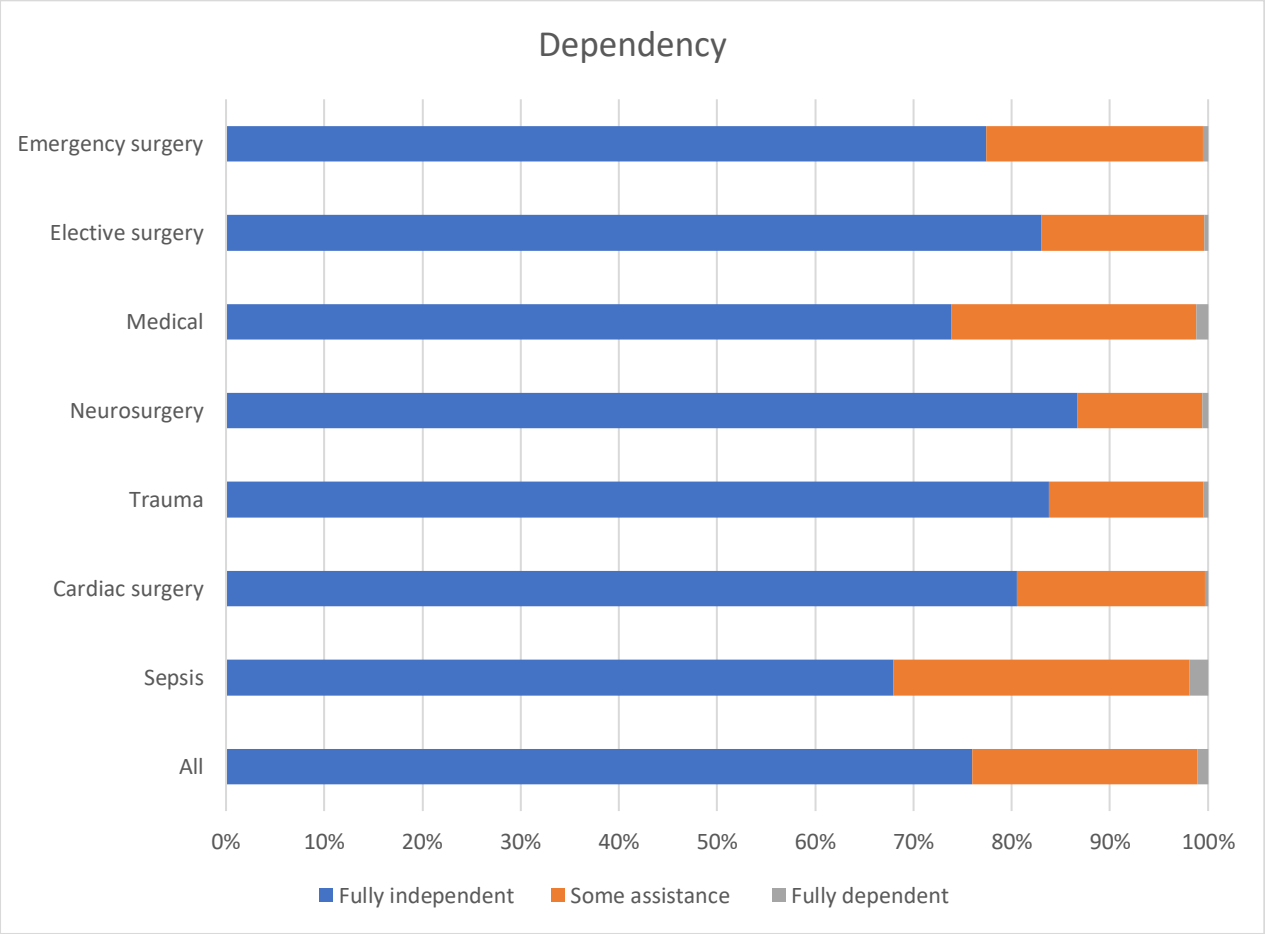
The unadjusted hospital mortality for all ICU patients is 21%. Mortality is higher in sepsis (31.8%) & medical (30.7%) and low in elective surgery (2.9%).

5.3. eFigure 4. Distribution of ethnicity across specialties.



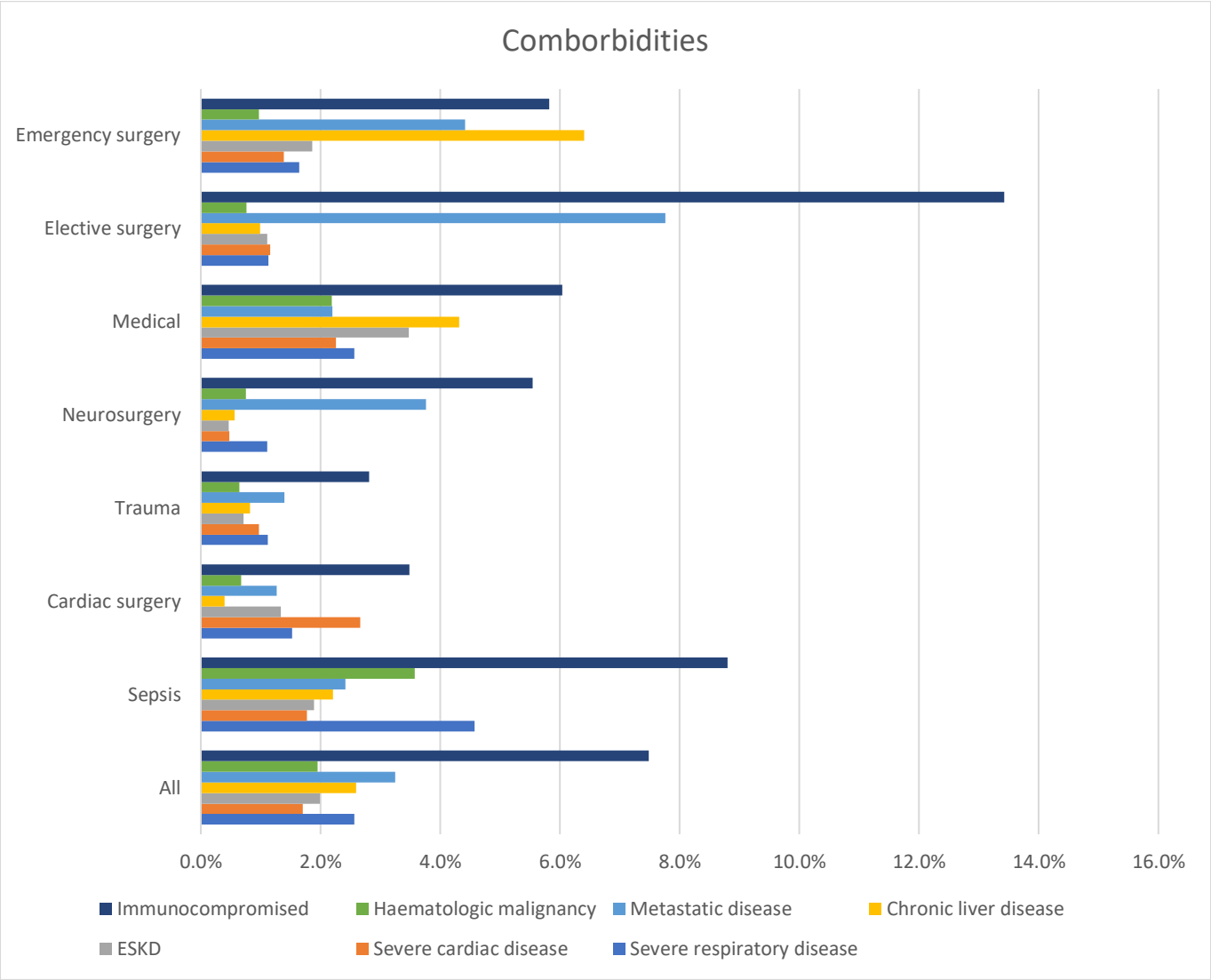
In the entire ICU cohort 90.2% of patients were white, 3.4% Asian, 2.2% Black and 4.1 Mixed /other.

5.4. eFigure 5. Distribution of dependency across specialties.



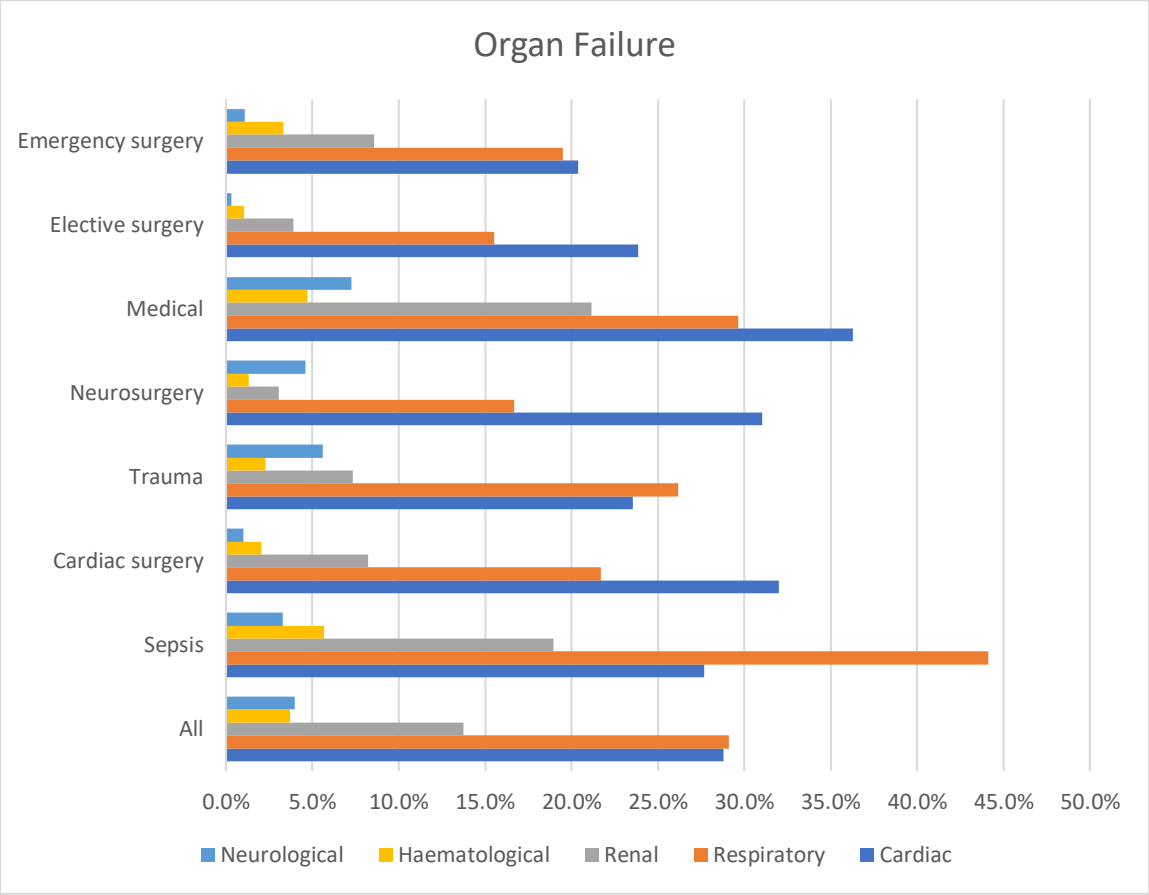
In the entire ICU cohort, 75.5% were fully independent, 22.8% required some assistance and 1% were fully dependant. Sepsis patients had the lowest levels of being fully independent (67.7%) compared with neurosurgery that has the higher share of fully independent patients (86.1%).

5.5. eFIgure 6. Variation in comorbidities across specialties.



Across the entire ICU cohort, 7.5% were immunocompromised, 3.2% had metastatic disease, 2.6% had severe respiratory disease. Elective surgery patients had the highest burden of immunocompromised state (13.4%) and metastatic disease (7.8%). Trauma patients had the lowest burden of co-morbidities.

5.6. eFIgure 7. Variation in organ failure across specialties.



Respiratory and cardiac failure (shock) were the most frequent organ failure,29.1% and 28.8%. In sepsis, respiratory failure occurred in 44.1% of patients. Elective surgery and trauma patients had the lowest burden of organ failure.

6.Subgroup analysis

eTable 1. A subgroup analysis of patients in the highest quartile of $ICNARC_{2018}$ score (N= 212,042).

Specialization measures	Absolute specialisation			Relative specialisation		
	OR	95%CI	P value	OR	95%CI	P value
Sepsis	1.00	[0.98,1.03]	0.843	1.00	[1.00,1.01]	0.396
Cardiac	0.98	[0.95,1.01]	0.240	1.00	[1.00,1.00]	0.937
Neurosurgery	0.99	[0.96,1.03]	0.663	1.00	[1.00,1.00]	0.817
Trauma	0.99	[0.96,1.02]	0.454	1.00	[1.00,1.00]	0.300
Medical	0.99	[0.96,1.02]	0.620	1.00	[0.99,1.01]	0.598
Elective surgery	0.98	[0.95,1.01]	0.195	1.00	[1.00,1.00]	0.854
Emergency surgery	0.99	[0.95,1.03]	0.508	1.00	[1.00,1.00]	0.796

OR= odds ratio, CI= confidence interval.

One unit represents a 10% change in specialisation

7. Sensitivity analysis

7.1 eTable 2. Table showing association between mortality and specialisation using the exponent of the share measure as the absolute measure and the relative difference as the measure of relative specialisation

Specialization measures	Exponent Absolute specialisation			Relative specialisation		
	OR	95%CI	P value	OR	95%CI	P value
Sepsis	1.00***	1.00, 1.00	<0.001	1.03**	1.01, 1.06	0.003
Cardiac	1.00	1.00, 1.00	0.538	1.00	0.97, 1.02	0.833
Neurosurgery	1.00	1.00, 1.00	0.260	1.02	0.99, 1.05	0.144
Trauma	1.00	1.00, 1.00	0.792	1.03*	1.00, 1.05	0.019
Medical	1.00	1.00, 1.00	0.698	1.02	1.00, 1.05	0.063
Elective surgery	1.00***	1.00, 1.00	<0.001	0.96**	0.94, 0.99	0.003
Emergency surgery	0.99**	0.99, 1.00	0.002	1.00	0.97, 1.03	0.997

OR= odds ratio, CI= confidence interval.

P<0.05=*. P<0.01=**P<0.001=***

8 Sensitivity analysis

8.1) eTable 3. Sensitivity analysis of year fixed effects.

	Absolute			Relative		
	OR	**95%CI**	**P value**	**OR**	**95%CI**	**P value**
Specialization measures						
• Sepsis	1.06***	1.04,1.08	<0.001	1.01***	1.01,1.02	<0.001
• Cardiac	1.00	0.97,1.03	0.932	1.00	1.00,1.00	0.724
• Neurosurgery	1.00	0.98,1.03	0.744	1.00	1.00,1.00	0.688
• Trauma	1.02	1.00,1.04	0.093	1.00**	1.00,1.01	0.003
• Medical	1.03**	1.01,1.06	0.007	1.01**	1.00,1.01	0.004
• Elective surgery	0.96**	0.94,0.98	0.001	0.99***	0.99,1.00	0.001
• Emergency surgery	1.01	0.98,1.03	0.715	1.00	1.00,1.00	0.554
Monthly total volume	0.99	0.98,1.00	0.065	0.99*	0.98,1.00	0.020
Functional state						
• Independent(reference)	1.00			1.00		
• Some assistance	1.12***	1.10,1.14	<0.001	1.12***	1.10,1.14	<0.001
• Total assistance	1.53***	1.44,1.63	<0.001	1.53***	1.44,1.63	<0.001
Ethnicity						
• White	1.00			1.00		
• Asian	0.95**	0.91,0.99	0.007	0.95**	0.91,0.99	0.007
• Black	0.83***	0.79,0.88	<0.001	0.83***	0.79,0.87	<0.001
• Mixed/other	0.99	0.96,1.03	0.789	0.99	0.96,1.03	0.788
Age in year	1.01***	1.01, 1.01	<0.001	1.01***	1.01, 1.01	<0.001
Male sex	1.11***	1.10,1.13	<0.001	1.11***	1.10,1.13	<0.001
ICNARC score	1.06***	1.06, 1.06	<0.001	1.06***	1.06, 1.06	<0.001
Co-morbidities						
• Severe respiratory disease	1.12***	1.08,1.17	<0.001	1.13***	1.08,1.17	<0.001
• Severe cardiovascular disease	1.30***	1.24,1.36	<0.001	1.30***	1.24,1.36	<0.001
• End-stage renal disease	1.25***	1.20,1.31	<0.001	1.25***	1.20,1.31	<0.001
• Severe liver disease	1.37***	1.32,1.42	<0.001	1.37***	1.32,1.42	<0.001
• Metastatic disease	1.20***	1.15,1.25	<0.001	1.20***	1.15,1.24	<0.001
• Haematological malignancy	1.18***	1.13,1.23	<0.001	1.18***	1.13,1.23	<0.001
• Immunocompromised	1.07***	1.04,1.10	<0.001	1.07***	1.04,1.10	<0.001
Organ failure						
• Shock	1.02**	1.01,1.04	0.008	1.02**	1.00,1.04	0.010
• Respiratory failure	1.20***	1.18,1.22	<0.001	1.20***	1.18,1.22	<0.001
• Acute renal failure	1.18***	1.16,1.20	<0.001	1.18***	1.16,1.20	<0.001
• Haematological failure	1.25***	1.21,1.30	<0.001	1.25***	1.21,1.30	<0.001
• Neurological failure	1.04	1.00,1.07	0.051	1.03	1.00,1.07	0.058
IMD	1.00***	1.00, 1.00	<0.001	1.00***	1.00, 1.00	<0.001
Monthly throughput	1.00	0.99,1.01	0.713	1.00	0.99,1.01	0.915
Academic affiliation						
• Non-university	1.00			1.00		
• University affiliated	0.98	0.91,1.07	0.683	0.99	0.91,1.07	0.765
• University	1.02	0.95,1.09	0.547	1.02	0.95,1.09	0.542
Year						
• 2010	1.00			1.00		
• 2011	0.92***	0.89,0.95	<0.001	0.91***	0.88,0.93	<0.001

• 2012	0.89***	0.86,0.92	<0.001	0.88***	0.85,0.90	<0.001
• 2013	0.86***	0.83,0.88	<0.001	0.85***	0.83,0.88	<0.001
• 2014	0.82***	0.80,0.85	<0.001	0.81***	0.79,0.84	<0.001
• 2015	0.85***	0.83,0.88	<0.001	0.83***	0.81,0.86	<0.001
• 2016	0.87***	0.84,0.90	<0.001	0.85***	0.83,0.88	<0.001

OR= odds ratio, CI= confidence interval.

One unit represents a 10% change in specialisation

P<0.05=*. P<0.01=**P<0.001=***

This analysis assumes that all ICUs in the same year deployed the same technologies and were subject to similar year effects. The results do not change the primary analysis.

8.2 eTable 4 The association between specialisation and ICU mortality

Specialization measures	Absolute specialisation			Relative specialisation		
	OR	95%CI	P value	OR	95%CI	P value
Sepsis	1.02	0.99,1.04	0.170	1.01*	1.00,1.01	0.049
Cardiac	1.01	0.98,1.04	0.524	1.00*	1.00,1.00	0.047
Neurosurgery	0.97	0.94,1.00	0.085	1.00	1.00,1.00	0.248
Trauma	1.00	0.97,1.02	0.728	1.00*	1.00,1.01	0.043
Medical	1.00	0.97,1.03	0.978	1.00	1.00,1.01	0.199
Elective surgery	0.97	0.95,1.00	0.075	1.00	1.00,1.00	0.902
Emergency surgery	1.00	0.97,1.04	0.981	1.00	1.00,1.00	0.204

OR= odds ratio, CI= confidence interval.

One unit represents a 10% change in specialisation

P<0.05=*. P<0.01=**P<0.001=***

This analysis uses ICU mortality as the outcome measure. This approach favourably allocates patients that survive to ICU discharge but subsequently die before hospital discharge. This approach offers the most optimistic analysis of the benefits of ICU specialisation. This analysis did not identify any benefits in terms of mortality to ICU specialisation

8.3 eTable 5. Subgroup of highest quartile ICNARC$_{2018}$ score treated by the ideal speciality ICU.

Subgroup	N	Absolute specialisation			Relative specialisation		
		OR	**95%CI**	**p-value**	**OR**	**95%CI**	**p-value**
Sepsis	102,838	1.01	0.97,1.05	0.555	1.00	0.99,1.01	0.870
Cardiac	6249	1.01	0.86,1.18	0.943	1.00	1.00,1,01	0.323
Neurosurgery	4,370	1.09	0.93,1.27	0.289	1.00	1.00,1,01	0.223
Trauma	15,845	0.90*	0.83, 0.98	0.016	0.99*	0.98,1.00	0.028
Medical	84,144	0.99	0.94,1.03	0.587	1.00	0.99, 1.02	0.502
Elective surgery	1,158	0.99	0.69, 1.41	0.935	1.00	0.99, 1.01	0.911
Emergency surgery	7,486	1.01	0.84, 1.20	0.932	1.01	0.98, 1.04	0.503

P<0.05=*. P<0.01=**P<0.001=***

This analysis examines the relationship between acute hospital mortality and "ideal" ICU specialisation for the most severely ill patients treated in the ideal ICU. i.e., the most severely ill with sepsis patients treated in the sepsis specialist ICU, the most severely ill with cardiac disease treated in the cardiac specialised ICU, the most severely ill patients with neurosurgical disease treated in the neurosurgery specialised ICU, the most severely ill patients with medical disease treated in a medicine specialised ICU, and the most severely ill patient requiring elective or emergency surgery treated in the respective "ideal" specialised ICU.

This analysis suggest that severely ill trauma patients may have lower acute hospital mortality when treated in a trauma specialised ICU. This analysis did not identify any other benefit to ICU specialisation.

8.4 eTable 6 Effect of absolute and relative specialisation on out-of-specialisation patients. The result for sepsis specialisation describes the odds ratio for acute hospital mortality for patients without sepsis. The result for cardiac specialisation describes the odds ratio for acute hospital mortality for non-cardiac patients. The result for neurosurgery describes the odds ratio for non-neurosurgical patients. The result for medical specialisation describes the acute hospital mortality for non -medical patients. The result for elective surgery specialisation describes the acute hospital mortality for non-elective surgery patients. The result for emergency surgery specialisation describes the acute hospital mortality for non-emergency surgery patients.

Specialization measures	**Absolute specialisation**			**Relative specialisation**		
	OR	95%CI	P value	OR	95%CI	P value
Sepsis	1.00	[0.98,1.03]	0.843	1.00	[1.00,1.01]	0.396
Cardiac	0.98	[0.95,1.01]	0.240	1.00	[1.00,1.00]	0.937
Neurosurgery	0.99	[0.96,1.03]	0.663	1.00	[1.00,1.00]	0.817
Trauma	0.99	[0.96,1.02]	0.454	1.00	[1.00,1.00]	0.300
Medical	0.99	[0.96,1.02]	0.620	1.00	[0.99,1.01]	0.598
Elective surgery	0.98	[0.95,1.01]	0.195	1.00	[1.00,1.00]	0.854
Emergency surgery	0.99	[0.95,1.03]	0.508	1.00	[1.00,1.00]	0.796

OR= odds ratio, CI= confidence interval; One unit represents a 10% change in specialisation

8.4 eTable 5. The effect of specialisation on non-specialised patients

Subgroup	N	Absolute specialisation			Relative specialisation		
		OR	**95%CI**	**p-value**	**OR**	**95%CI**	**p-value**
Sepsis specialisation on non-sepsis	612,397	1.04	1.01,1.07	0.003	1.01	1.00,1.02	0.029
Cardiac specialisation on non-cardiac	743,946	1.03	1.00,1.05	0.084	1.00	1.00,1,00	0.339
Neurosurgery specialisation on non-neurosurgical	825,198	1.01	0.99,1.04	0.305	1.00	1.00,1,00	0.279
Trauma specialisation on non-trauma	757,698	1.02	1.00,1.05	0.017	1.00	1.00,1.01	0.005
Medical specialisation on non-medical	642,049	1.03	1.01,1.06	0.017	1.00	1.00, 1.01	0.274
Elective surgery specialisation on non-elective surgery	717,914	1.00	0.97,1.02	0.887	1.00	1.00, 1.00	0.400
Emergency surgery specialisation on-emergency surgery	804,204	1.00	0.98, 1.03	0.696	1.00	1.00, 1.00	0.657

OR= odds ratio, CI= confidence interval; One unit represents a 10% change in specialisation

Patients without sepsis have an OR 1.04 (95% CI 1.01-1.07; p=0.003) and a OR 0f 1.01 (95%CI 1.00-1.02, p=0.029) for every 10% increase in absolute and relative sepsis specialisation. There is no significant effect observed for non-cardiac patients treated in the cardiac ICU, non-neurosurgical patients treated in the non-neurosurgical ICU and non-surgical patients treated in the surgical ICU.

References for Appendices

1. Arrow KJ: **The Economic Implications of Learning by Doing**. *The Review of Economic Studies* 1962, **29**(3):155-173.
2. Hall BHRN: **Handbook of the economics of innovation**. Amsterdam; Boston: North Holland; 2010.
3. Skinner W: **The Focused Factory**. *Harvard Business Review* 1974, **May-June**:114-121.
4. Bredenhoff E, van Lent WA, van Harten WH: **Exploring types of focused factories in hospital care: a multiple case study**. *BMC Health Serv Res* 2010, **10**:154.
5. Luft HS, Hunt SS, Maerki SC: **The volume-outcome relationship: practice-makes-perfect or selective-referral patterns?** *Health Serv Res* 1987, **22**(2):157-182.
6. Ketelaar NA, Faber MJ, Flottorp S, Rygh LH, Deane KH, Eccles MP: **Public release of performance data in changing the behaviour of healthcare consumers, professionals or organisations**. *Cochrane Database Syst Rev* 2011(11):CD004538.
7. Birkmeyer JD, Skinner JS, Wennberg DE: **Will volume-based referral strategies reduce costs or just save lives?** *Health Aff (Millwood)* 2002, **21**(5):234-241.
8. Stanak M, Strohmaier C: **Minimum volume standards in day surgery: a systematic review**. *BMC Health Serv Res* 2020, **20**(1):886.
9. Halm EA, Lee C, Chassin MR: **Is volume related to outcome in health care? A systematic review and methodologic critique of the literature**. *Ann Intern Med* 2002, **137**(6):511-520.
10. Christian CK, Gustafson ML, Betensky RA, Daley J, Zinner MJ: **The volume-outcome relationship: don't believe everything you see**. *World journal of surgery* 2005, **29**(10):1241-1244.
11. Birkmeyer JD, Siewers AE, Finlayson EV, Stukel TA, Lucas FL, Batista I, Welch HG, Wennberg DE: **Hospital volume and surgical mortality in the United States**. *N Engl J Med* 2002, **346**(15):1128-1137.
12. Papageorge MV, de Geus SWL, Woods AP, Ng SC, McAneny D, Tseng JF, Kenzik KM, Sachs TE: **The Effect of Hospital Versus Surgeon Volume on Short-Term Patient Outcomes After Pancreaticoduodenectomy: a SEER-Medicare Analysis**. *Ann Surg Oncol* 2022, **29**(4):2444-2451.
13. Enomoto LM, Gusani NJ, Dillon PW, Hollenbeak CS: **Impact of surgeon and hospital volume on mortality, length of stay, and cost of pancreaticoduodenectomy**. *J Gastrointest Surg* 2014, **18**(4):690-700.
14. Hamilton BH, Hamilton VH: **Estimating surgical volume--outcome relationships applying survival models: accounting for frailty and hospital fixed effects**. *Health Econ* 1997, **6**(4):383-395.
15. Gaynor M, Seider H, Vogt WB: **The Volume-Outcome Effect, Scale Economies, and Learning-by-Doing**. *American Economic Review* 2005, **95**(2):243-247.
16. Flood AB, Scott WR, Ewy W: **Does Practice Make Perfect? Part II: The Relation between Volume and and Outcomes and Other Hospital Characteristics**. *Medical Care* 1984, **22**(2):115-125.
17. Morche J, Renner D, Pietsch B, Kaiser L, Bronneke J, Gruber S, Matthias K: **International comparison of minimum volume standards for hospitals**. *Health Policy* 2018, **122**(11):1165-1176.
18. Ho V: **Learning by Doing**. In: *Encyclopedia of Health Economics.* Edited by Culyer AJ. San Diego: Elsevier; 2014: 141-145.

19. Sfekas A: **Learning, forgetting, and hospital quality: an empirical analysis of cardiac procedures in Maryland and Arizona**. *Health Econ* 2009, **18**(6):697-711.
20. Reader TW, Flin R, Mearns K, Cuthbertson BH: **Developing a team performance framework for the intensive care unit**. *Crit Care Med* 2009, **37**(5):1787-1793.
21. Wheelan SA, Burchill CN, Tilin F: **The link between teamwork and patients' outcomes in intensive care units**. *Am J Crit Care* 2003, **12**(6):527-534.
22. Pronovost PJ, Thompson DA, Holzmueller CG, Lubomski LH, Dorman T, Dickman F, Fahey M, Steinwachs DM, Engineer L, Sexton JB *et al*: **Toward learning from patient safety reporting systems**. *J Crit Care* 2006, **21**(4):305-315.
23. Rosen MA, DiazGranados D, Dietz AS, Benishek LE, Thompson D, Pronovost PJ, Weaver SJ: **Teamwork in healthcare: Key discoveries enabling safer, high-quality care**. *Am Psychol* 2018, **73**(4):433-450.
24. Gawande AA, Zinner MJ, Studdert DM, Brennan TA: **Analysis of errors reported by surgeons at three teaching hospitals**. *Surgery* 2003, **133**(6):614-621.
25. Keers RN, Williams SD, Cooke J, Ashcroft DM: **Causes of medication administration errors in hospitals: a systematic review of quantitative and qualitative evidence**. *Drug Saf* 2013, **36**(11):1045-1067.
26. Guzzo RASE: **Team effectiveness and decision making in organizations**. San Francisco: Jossey-Bass; 1995.
27. Heath C, Staudenmayer N: **Coordination Neglect: How Lay Theories of Organizing Complicate Coordination in Organizations**. *Research in Organizational Behavior* 2000, **22**:153-191.
28. Reader TW, Cuthbertson BH: **Teamwork and team training in the ICU: where do the similarities with aviation end?** *Crit Care* 2011, **15**(6):313.
29. Embriaco N, Papazian L, Kentish-Barnes N, Pochard F, Azoulay E: **Burnout syndrome among critical care healthcare workers**. *Curr Opin Crit Care* 2007, **13**(5):482-488.
30. Argote L, Epple D: **Learning Curves in Manufacturing**. *Science* 1990, **247**(4945):920-924.
31. Smith A: **Wealth of Nations. Books 1-3 Books 1-3**; 2009.
32. FW T: **American Council of Learned Societies. Scientific Management : Comprising Shop Management the Principles of Scientific Management and Testimony Before the Special House Committee.** . New York:: Harper;; 1947.
33. Hyer NL, Wemmerlöv U, Morris JA: **Performance analysis of a focused hospital unit: The case of an integrated trauma center**. *Journal of Operations Management* 2009, **27**(3):203-219.
34. Lindlbauer I, Schreyogg J: **The relationship between hospital specialization and hospital efficiency: do different measures of specialization lead to different results?** *Health Care Manag Sci* 2014, **17**(4):365-378.
35. H. T: **Economics and Information Theory.** ; 1967.
36. Daidone S, D'Amico F: **Technical efficiency, specialization and ownership form: evidences from a pooling of Italian hospitals**. *Journal of Productivity Analysis* 2009, **32**(3):203.
37. Huesch MD, Sakakibara M: **Forgetting the learning curve for a moment: how much performance is unrelated to own experience?** *Health Econ* 2009, **18**(7):855-862.
38. Chandra A, Staiger DO: **Productivity Spillovers in Healthcare: Evidence from the Treatment of Heart Attacks**. *J Polit Econ* 2007, **115**:103-140.

39. Liedtka JM, Whitten E: **Enhancing care delivery through cross-disciplinary collaboration: a case study**. *J Healthc Manag* 1998, **43**(2):185-203; discussion 203-185.
40. Young GJ, Charns MP, Heeren TC: **Product-Line Management in Professional Organizations: an Empirical Test of Competing Theoretical Perspectives**. *Academy of Management Journal* 2004, **47**(5):723-734.
41. Byrne MM, Charns MP, Parker VA, Meterko MM, Wray NP: **The effects of organization on medical utilization: an analysis of service line organization**. *Med Care* 2004, **42**(1):28-37.
42. Cram P, Vaughan-Sarrazin MS, Wolf B, Katz JN, Rosenthal GE: **A comparison of total hip and knee replacement in specialty and general hospitals**. *J Bone Joint Surg Am* 2007, **89**(8):1675-1684.
43. Girotra S, Lu X, Popescu I, Vaughan-Sarrazin M, Horwitz PA, Cram P: **The impact of hospital cardiac specialization on outcomes after coronary artery bypass graft surgery: analysis of medicare claims data**. *Circ Cardiovasc Qual Outcomes* 2010, **3**(6):607-614.
44. Greenwald L, Cromwell J, Adamache W, Bernard S, Drozd E, Root E, Devers K: **Specialty versus community hospitals: referrals, quality, and community benefits**. *Health Aff (Millwood)* 2006, **25**(1):106-118.
45. Carey K, Burgess JF, Jr., Young GJ: **Specialty and full-service hospitals: a comparative cost analysis**. *Health Serv Res* 2008, **43**(5 Pt 2):1869-1887.
46. E. HR: **Market-Driven Health Care : Who Wins Who Loses in the Transformation of America's Largest Service Industry. Reading Mass: Addison-Wesley Pub.** 1997.
47. Lott JP, Iwashyna TJ, Christie JD, Asch DA, Kramer AA, Kahn JM: **Critical illness outcomes in specialty versus general intensive care units**. *Am J Respir Crit Care Med* 2009, **179**(8):676-683.
48. Kramer AH, Zygun DA: **Do neurocritical care units save lives? Measuring the impact of specialized ICUs**. *Neurocrit Care* 2011, **14**(3):329-333.
49. Elf K, Nilsson P, Enblad P: **Outcome after traumatic brain injury improved by an organized secondary insult program and standardized neurointensive care**. *Crit Care Med* 2002, **30**(9):2129-2134.
50. Cram P, Rosenthal GE, Vaughan-Sarrazin MS: **Cardiac revascularization in specialty and general hospitals**. *N Engl J Med* 2005, **352**(14):1454-1462.
51. Berger E, Winkelmann J, Eckhardt H, Nimptsch U, Panteli D, Reichebner C, Rombey T, Busse R: **A country-level analysis comparing hospital capacity and utilisation during the first COVID-19 wave across Europe**. *Health Policy* 2022, **126**(5):373-381.
52. Cardona-Morrell M, Kim J, Turner RM, Anstey M, Mitchell IA, Hillman K: **Non-beneficial treatments in hospital at the end of life: a systematic review on extent of the problem**. *Int J Qual Health Care* 2016, **28**(4):456-469.
53. Chang DW, Shapiro MF: **Association Between Intensive Care Unit Utilization During Hospitalization and Costs, Use of Invasive Procedures, and Mortality**. *JAMA Intern Med* 2016, **176**(10):1492-1499.
54. Taylor RA: **Principles of the Economic Behavior of Hospitals**. *Nebraska Journal of Economics and Business* 1972, **11**(2):45-53.
55. McGuire A: **The theory of the hospital: a review of the models**. *Soc Sci Med* 1985, **20**(11):1177-1184.

56. Pauly M, Redisch M: **The Not-For-Profit Hospital as a Physicians' Cooperative**. *The American Economic Review* 1973, **63**(1):87-99.
57. Newhouse JP: **Toward a Theory of Nonprofit Institutions: An Economic Model of a Hospital**. *The American Economic Review* 1970, **60**(1):64-74.
58. Harris JE: **The Internal Organization of Hospitals: Some Economic Implications**. *Bell Journal of Economics* 1977, **8**(2):467-482.
59. Crilly T, Le Grand JJ: **The motivation and behaviour of hospital Trusts**. *Soc Sci Med* 2004, **58**(10):1809-1823.
60. Goldfarb M, Hornbrook M, Rafferty J: **Behavior of the Multiproduct Firm: A Model of the Nonprofit Hospital System**. *Medical Care* 1980, **18**(2):185-201.
61. Halpern NA, Pastores SM: **Critical Care Medicine Beds, Use, Occupancy, and Costs in the United States: A Methodological Review**. *Crit Care Med* 2015, **43**(11):2452-2459.
62. Wunsch H: **The outbreak that invented intensive care**. *Nature* 2020.
63. Everberg G, Ibsen B: **[Transport of polio patients with respiratory symptoms]**. *Ugeskr Laeger* 1954, **116**(29):1077-1078.
64. Lassen HC: **A preliminary report on the 1952 epidemic of poliomyelitis in Copenhagen with special reference to the treatment of acute respiratory insufficiency**. *Lancet* 1953, **1**(6749):37-41.
65. Takala J: **A History of Intensive Care Medicine**. In: *The Wondrous Story of Anesthesia.* Edited by Eger Ii EI, Saidman LJ, Westhorpe RN. New York, NY: Springer New York; 2014: 785-798.
66. Weil MH, Tang W: **From intensive care to critical care medicine: a historical perspective**. *Am J Respir Crit Care Med* 2011, **183**(11):1451-1453.
67. Rothen HU, Stricker K, Einfalt J, Bauer P, Metnitz PG, Moreno RP, Takala J: **Variability in outcome and resource use in intensive care units**. *Intensive Care Med* 2007, **33**(8):1329-1336.
68. Marshall JC, Bosco L, Adhikari NK, Connolly B, Diaz JV, Dorman T, Fowler RA, Meyfroidt G, Nakagawa S, Pelosi P *et al*: **What is an intensive care unit? A report of the task force of the World Federation of Societies of Intensive and Critical Care Medicine**. *J Crit Care* 2017, **37**:270-276.
69. Wilcox ME, Chong CA, Niven DJ, Rubenfeld GD, Rowan KM, Wunsch H, Fan E: **Do intensivist staffing patterns influence hospital mortality following ICU admission? A systematic review and meta-analyses**. *Crit Care Med* 2013, **41**(10):2253-2274.
70. Angus DC, Shorr AF, White A, Dremsizov TT, Schmitz RJ, Kelley MA, Committee on Manpower for P, Critical Care S: **Critical care delivery in the United States: distribution of services and compliance with Leapfrog recommendations**. *Crit Care Med* 2006, **34**(4):1016-1024.
71. Pronovost PJ, Angus DC, Dorman T, Robinson KA, Dremsizov TT, Young TL: **Physician staffing patterns and clinical outcomes in critically ill patients: a systematic review**. *JAMA* 2002, **288**(17):2151-2162.
72. Murthy S, Wunsch H: **Clinical review: International comparisons in critical care - lessons learned**. *Crit Care* 2012, **16**(2):218.
73. Martin JM, Hart GK, Hicks P: **A unique snapshot of intensive care resources in Australia and New Zealand**. *Anaesth Intensive Care* 2010, **38**(1):149-158.
74. Prin M, Wunsch H: **International comparisons of intensive care: informing outcomes and improving standards**. *Curr Opin Crit Care* 2012, **18**(6):700-706.

75. **Levels of care** [https://www.ics.ac.uk/Society/Guidance/PDFs/Levels_of_Care_25_3_21]
76. **Guidelines for intensive care unit admission, discharge, and triage**. *Critical Care Medicine* 1999, **27**(3):633-638.
77. Ward NS, Afessa B, Kleinpell R, Tisherman S, Ries M, Howell M, Halpern N, Kahn J, Members of Society of Critical Care Medicine Taskforce on ICUS: **Intensivist/patient ratios in closed ICUs: a statement from the Society of Critical Care Medicine Taskforce on ICU Staffing**. *Crit Care Med* 2013, **41**(2):638-645.
78. Gershengorn HB, Harrison DA, Garland A, Wilcox ME, Rowan KM, Wunsch H: **Association of Intensive Care Unit Patient-to-Intensivist Ratios With Hospital Mortality**. *JAMA Intern Med* 2017, **177**(3):388-396.
79. Wunsch H, Angus DC, Harrison DA, Collange O, Fowler R, Hoste EA, de Keizer NF, Kersten A, Linde-Zwirble WT, Sandiumenge A *et al*: **Variation in critical care services across North America and Western Europe**. *Crit Care Med* 2008, **36**(10):2787-2793, e2781-2789.
80. Guidet B, van der Voort PHJ, Csomos A: **Intensive care in 2050: healthcare expenditure**. *Intensive Care Medicine* 2017, **43**(8):1141-1143.
81. Gooch RA, Kahn JM: **ICU bed supply, utilization, and health care spending: an example of demand elasticity**. *JAMA* 2014, **311**(6):567-568.
82. Takala J, Moser A, Raj R, Pettila V, Irincheeva I, Selander T, Kiiski O, Varpula T, Reinikainen M, Jakob SM: **Variation in severity-adjusted resource use and outcome in intensive care units**. *Intensive Care Med* 2022, **48**(1):67-77.
83. Knaus WA, Wagner DP, Zimmerman JE, Draper EA: **Variations in mortality and length of stay in intensive care units**. *Ann Intern Med* 1993, **118**(10):753-761.
84. Churpek MM, Gupta S, Spicer AB, Parker WF, Fahrenbach J, Brenner SK, Leaf DE, Investigators S-C: **Hospital-Level Variation in Death for Critically Ill Patients with COVID-19**. *Am J Respir Crit Care Med* 2021, **204**(403-411).
85. Gillies MA, Power GS, Harrison DA, Fleming A, Cook B, Walsh TS, Pearse RM, Rowan KM: **Regional variation in critical care provision and outcome after high-risk surgery**. *Intensive Care Med* 2015, **41**(10):1809-1816.
86. Kadri SS, Sun J, Lawandi A, Strich JR, Busch LM, Keller M, Babiker A, Yek C, Malik S, Krack J *et al*: **Association Between Caseload Surge and COVID-19 Survival in 558 U.S. Hospitals, March to August 2020**. *Ann Intern Med* 2021, **174**(9):1240-1251.
87. Rapoport J, Teres D, Barnett R, Jacobs P, Shustack A, Lemeshow S, Norris C, Hamilton S: **A comparison of intensive care unit utilization in Alberta and western Massachusetts**. *Crit Care Med* 1995, **23**(8):1336-1346.
88. Pearse RM, Moreno RP, Bauer P, Pelosi P, Metnitz P, Spies C, Vallet B, Vincent JL, Hoeft A, Rhodes A *et al*: **Mortality after surgery in Europe: a 7 day cohort study**. *Lancet* 2012, **380**(9847):1059-1065.
89. Bekelman JE, Halpern SD, Blankart CR, Bynum JP, Cohen J, Fowler R, Kaasa S, Kwietniewski L, Melberg HO, Onwuteaka-Philipsen B *et al*: **Comparison of Site of Death, Health Care Utilization, and Hospital Expenditures for Patients Dying With Cancer in 7 Developed Countries**. *JAMA* 2016, **315**(3):272-283.
90. Angus DC, Barnato AE, Linde-Zwirble WT, Weissfeld LA, Watson RS, Rickert T, Rubenfeld GD, Robert Wood Johnson Foundation ICUE-O-LPG: **Use of intensive care at the end of life in the United States: an epidemiologic study**. *Crit Care Med* 2004, **32**(3):638-643.

91. Wunsch H, Linde-Zwirble WT, Harrison DA, Barnato AE, Rowan KM, Angus DC: **Use of intensive care services during terminal hospitalizations in England and the United States**. *Am J Respir Crit Care Med* 2009, **180**(9):875-880.
92. Howell MD, Stevens JP: **ICUs after surgery, mortality, and the Will Rogers effect**. *Intensive Care Med* 2015, **41**(11):1990-1992.
93. Rudd KE, Johnson SC, Agesa KM, Shackelford KA, Tsoi D, Kievlan DR, Colombara DV, Ikuta KS, Kissoon N, Finfer S *et al*: **Global, regional, and national sepsis incidence and mortality, 1990-2017: analysis for the Global Burden of Disease Study**. *Lancet* 2020, **395**(10219):200-211.
94. Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, Bellomo R, Bernard GR, Chiche JD, Coopersmith CM *et al*: **The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3)**. *JAMA* 2016, **315**(8):801-810.
95. Iwashyna TJ, Cooke CR, Wunsch H, Kahn JM: **Population burden of long-term survivorship after severe sepsis in older Americans**. *J Am Geriatr Soc* 2012, **60**(6):1070-1077.
96. Bauer M, Gerlach H, Vogelmann T, Preissing F, Stiefel J, Adam D: **Mortality in sepsis and septic shock in Europe, North America and Australia between 2009 and 2019-results from a systematic review and meta-analysis**. *Crit Care* 2020, **24**(1):239.
97. Seymour CW, Gesten F, Prescott HC, Friedrich ME, Iwashyna TJ, Phillips GS, Lemeshow S, Osborn T, Terry KM, Levy MM: **Time to Treatment and Mortality during Mandated Emergency Care for Sepsis**. *N Engl J Med* 2017, **376**(23):2235-2244.
98. Ferrer R, Artigas A, Levy MM, Blanco J, Gonzalez-Diaz G, Garnacho-Montero J, Ibanez J, Palencia E, Quintana M, de la Torre-Prados MV *et al*: **Improvement in process of care and outcome after a multicenter severe sepsis educational program in Spain**. *JAMA* 2008, **299**(19):2294-2303.
99. Rhee C, Yu T, Wang R, Kadri SS, Fram D, Chen HC, Klompas M, Program CDCPE: **Association Between Implementation of the Severe Sepsis and Septic Shock Early Management Bundle Performance Measure and Outcomes in Patients With Suspected Sepsis in US Hospitals**. *JAMA Netw Open* 2021, **4**(12):e2138596.
100. Park RE, Brook RH, Kosecoff J, Keesey J, Rubenstein L, Keeler E, Kahn KL, Rogers WH, Chassin MR: **Explaining variations in hospital death rates. Randomness, severity of illness, quality of care**. *JAMA* 1990, **264**(4):484-490.
101. Hatfield KM, Dantes RB, Baggs J, Sapiano MRP, Fiore AE, Jernigan JA, Epstein L: **Assessing Variability in Hospital-Level Mortality Among U.S. Medicare Beneficiaries With Hospitalizations for Severe Sepsis and Septic Shock**. *Crit Care Med* 2018, **46**(11):1753-1760.
102. Abe T, Ogura H, Kushimoto S, Shiraishi A, Sugiyama T, Deshpande GA, Uchida M, Nagata I, Saitoh D, Fujishima S *et al*: **Variations in infection sites and mortality rates among patients in intensive care units with severe sepsis and septic shock in Japan**. *J Intensive Care* 2019, **7**:28.
103. ColeBrook LP, A.W. : **Treatment of 106 cases of peurperal fever by sulphanilamide(streptocide).** *Lancet* 1937, **230**(5961):1237-1242.
104. Bauer J, Bruggmann D, Klingelhofer D, Maier W, Schwettmann L, Weiss DJ, Groneberg DA: **Access to intensive care in 14 European countries: a spatial analysis**

of intensive care need and capacity in the light of COVID-19. *Intensive Care Med* 2020, **46**(11):2026-2034.
105. OECD: **Beyond containment. Health systems responses to COVID-19 in the OECD**. In: *Tackling coronavirus (COVID-19): Contributing to a Global Effort.* 2020.
106. Metcalfe MA, Sloggett A, McPherson K: **Mortality among appropriately referred patients refused admission to intensive-care units**. *Lancet* 1997, **350**(9070):7-11.
107. Goldfrad C, Rowan K: **Consequences of discharges from intensive care at night**. *Lancet* 2000, **355**(9210):1138-1142.
108. Hutchings A, Durand MA, Grieve R, Harrison D, Rowan K, Green J, Cairns J, Black N: **Evaluation of modernisation of adult critical care services in England: time series and cost effectiveness analysis**. *BMJ* 2009, **339**:b4353.
109. Rosenthal GE, Sirio CA, Shepardson LB, Harper DL, Rotondi AJ, Cooper GS: **Use of intensive care units for patients with low severity of illness**. *Arch Intern Med* 1998, **158**(10):1144-1151.
110. Chen LM, Render M, Sales A, Kennedy EH, Wiitala W, Hofer TP: **Intensive care unit admitting patterns in the Veterans Affairs health care system**. *Arch Intern Med* 2012, **172**(16):1220-1226.
111. OECD: **UNDERSTANDING DIFFERENCES IN HEALTH EXPENDITURE BETWEEN THE UNITED STATES AND OECD COUNTRIES**. In.; 2022.
112. Rhodes A, Ferdinande P, Flaatten H, Guidet B, Metnitz PG, Moreno RP: **The variability of critical care bed numbers in Europe**. *Intensive Care Med* 2012, **38**(10):1647-1653.
113. Wunsch H, Angus DC, Harrison DA, Linde-Zwirble WT, Rowan KM: **Comparison of medical admissions to intensive care units in the United States and United Kingdom**. *Am J Respir Crit Care Med* 2011, **183**(12):1666-1673.
114. Carr BG, Addyson DK, Kahn JM: **Variation in critical care beds per capita in the United States: implications for pandemic and disaster planning**. *JAMA* 2010, **303**(14):1371-1372.
115. **Population and household estimates, England and Wales: Census 2021** [https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/populationandhouseholdestimatesenglandandwales/census2021]
116. **Critical Care Bed Capacity and Urgent Operations Cancelled** [https://www.england.nhs.uk/statistics/statistical-work-areas/critical-care-capacity/]
117. Bravata DM, Perkins AJ, Myers LJ, Arling G, Zhang Y, Zillich AJ, Reese L, Dysangco A, Agarwal R, Myers J *et al*: **Association of Intensive Care Unit Patient Load and Demand With Mortality Rates in US Department of Veterans Affairs Hospitals During the COVID-19 Pandemic**. *JAMA Netw Open* 2021, **4**(1):e2034266.
118. Wilcox ME, Rowan KM, Harrison DA, Doidge JC: **Does Unprecedented ICU Capacity Strain, As Experienced During the COVID-19 Pandemic, Impact Patient Outcome?** *Crit Care Med* 2022, **50**(6):e548-e556.
119. Black N, Payne M: **Directory of clinical databases: improving and promoting their use**. *Qual Saf Health Care* 2003, **12**(5):348-352.
120. Harrison DA, Brady AR, Rowan K: **Case mix, outcome and length of stay for admissions to adult, general critical care units in England, Wales and Northern Ireland: the Intensive Care National Audit & Research Centre Case Mix Programme Database**. *Crit Care* 2004, **8**(2):R99-111.

121. Young JD, Goldfrad C, Rowan K: **Development and testing of a hierarchical method to code the reason for admission to intensive care units: the ICNARC Coding Method. Intensive Care National Audit & Research Centre**. *Br J Anaesth* 2001, **87**(4):543-548.
122. Seymour CW, Liu VX, Iwashyna TJ, Brunkhorst FM, Rea TD, Scherag A, Rubenfeld G, Kahn JM, Shankar-Hari M, Singer M *et al*: **Assessment of Clinical Criteria for Sepsis: For the Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3)**. *JAMA* 2016, **315**(8):762-774.
123. Ferrando-Vivas P, Jones A, Rowan KM, Harrison DA: **Development and validation of the new ICNARC model for prediction of acute hospital mortality in adult critical care**. *J Crit Care* 2017, **38**:335-339.
124. Zimmerman JE, Kramer AA, McNair DS, Malila FM, Shaffer VL: **Intensive care unit length of stay: Benchmarking based on Acute Physiology and Chronic Health Evaluation (APACHE) IV**. *Crit Care Med* 2006, **34**(10):2517-2529.
125. de Keizer NF, Bonsel GJ, Goldfad C, Rowan KM: **The added value that increasing levels of diagnostic information provide in prognostic models to estimate hospital mortality for adult intensive care patients**. *Intensive Care Med* 2000, **26**(5):577-584.
126. Arts DG, Bosman RJ, de Jonge E, Joore JC, de Keizer NF: **Training in data definitions improves quality of intensive care data**. *Crit Care* 2003, **7**(2):179-184.
127. Chen LM, Martin CM, Morrison TL, Sibbald WJ: **Interobserver variability in data collection of the APACHE II score in teaching and community hospitals**. *Crit Care Med* 1999, **27**(9):1999-2004.
128. Simkins M, Iqbal A, Gronemeyer A, Konzen L, White J, Koenig M, Palmer C, Kerby P, Buckman S, Despotovic V *et al*: **Inter-Rater Reliability and Impact of Disagreements on Acute Physiology and Chronic Health Evaluation IV Mortality Predictions**. *Crit Care Explor* 2019, **1**(10):e0059.
129. Wunsch H, Brady AR, Rowan K: **Impact of exclusion criteria on case mix, outcome, and length of stay for the severity of disease scoring methods in common use in critical care**. *J Crit Care* 2004, **19**(2):67-74.
130. Harrison DA, Brady AR, Parry GJ, Carpenter JR, Rowan K: **Recalibration of risk prediction models in a large multicenter cohort of admissions to adult, general critical care units in the United Kingdom**. *Crit Care Med* 2006, **34**(5):1378-1388.
131. Harrison DA, Parry GJ, Rowan K: **Risk prediction models: recalibration**. *Crit Care Med* 2007, **35**(1):326-327.
132. Harrison DA, Ferrando-Vivas P, Shahin J, Rowan KM. In: *Ensuring comparisons of health-care providers are fair: development and validation of risk prediction models for critically ill patients.* Southampton (UK); 2015.
133. Drye EE, Normand SL, Wang Y, Ross JS, Schreiner GC, Han L, Rapp M, Krumholz HM: **Comparison of hospital risk-standardized mortality rates calculated by using in-hospital and 30-day models: an observational study with implications for hospital profiling**. *Ann Intern Med* 2012, **156**(1 Pt 1):19-26.
134. Luft HS, Bunker JP, Enthoven AC: **Should operations be regionalized? The empirical relation between surgical volume and mortality. 1979**. *Clin Orthop Relat Res* 2007, **457**:3-9.
135. Mesman R, Faber MJ, Berden B, Westert GP: **Evaluation of minimum volume standards for surgery in the Netherlands (2003-2017): A successful policy?** *Health Policy* 2017, **121**(12):1263-1273.

136. Bhattarai N, McMeekin P, Price C, Vale L: **Economic evaluations on centralisation of specialised healthcare services: a systematic review of methods**. *BMJ Open* 2016, **6**(5):e011214.
137. Metcalfe D, Bouamra O, Parsons NR, Aletrari MO, Lecky FE, Costa ML: **Effect of regional trauma centralization on volume, injury severity and outcomes of injured patients admitted to trauma centres**. *Br J Surg* 2014, **101**(8):959-964.
138. Morris S, Hunter RM, Ramsay AI, Boaden R, McKevitt C, Perry C, Pursani N, Rudd AG, Schwamm LH, Turner SJ *et al*: **Impact of centralising acute stroke services in English metropolitan areas on mortality and length of hospital stay: difference-in-differences analysis**. *BMJ* 2014, **349**:g4757.
139. Fitzsimons KJ, Mukarram S, Copley LP, Deacon SA, van der Meulen JH: **Centralisation of services for children with cleft lip or palate in England: a study of hospital episode statistics**. *BMC Health Serv Res* 2012, **12**:148.
140. Christiansen T, Vrangbaek K: **Hospital centralization and performance in Denmark-Ten years on**. *Health Policy* 2018, **122**(4):321-328.
141. Magnussen J, Hagen TP, Kaarboe OM: **Centralized or decentralized? A case study of Norwegian hospital reform**. *Soc Sci Med* 2007, **64**(10):2129-2137.
142. Nimptsch U, Mansky T: **Hospital volume and mortality for 25 types of inpatient treatment in German hospitals: observational study using complete national data from 2009 to 2014**. *BMJ Open* 2017, **7**(9):e016184.
143. Lindenauer PK, Behal R, Murray CK, Nsa W, Houck PM, Bratzler DW: **Volume, quality of care, and outcome in pneumonia**. *Ann Intern Med* 2006, **144**(4):262-269.
144. Ross JS, Normand SL, Wang Y, Ko DT, Chen J, Drye EE, Keenan PS, Lichtman JH, Bueno H, Schreiner GC *et al*: **Hospital volume and 30-day mortality for three common medical conditions**. *N Engl J Med* 2010, **362**(12):1110-1118.
145. Haut ER, Pronovost PJ, Schneider EB: **Limitations of administrative databases**. *JAMA* 2012, **307**(24):2589; author reply 2589-2590.
146. Kumbhani DJ, Fonarow GC, Heidenreich PA, Schulte PJ, Lu D, Hernandez A, Yancy C, Bhatt DL: **Association Between Hospital Volume, Processes of Care, and Outcomes in Patients Admitted With Heart Failure: Insights From Get With The Guidelines-Heart Failure**. *Circulation* 2018, **137**(16):1661-1670.
147. Joynt KE, Orav EJ, Jha AK: **The association between hospital volume and processes, outcomes, and costs of care for congestive heart failure**. *Ann Intern Med* 2011, **154**(2):94-102.
148. Angus DC, Black N: **Improving care of the critically ill: institutional and health-care system approaches**. *Lancet* 2004, **363**(9417):1314-1320.
149. Nguyen YL, Wallace DJ, Yordanov Y, Trinquart L, Blomkvist J, Angus DC, Kahn JM, Ravaud P, Guidet B: **The Volume-Outcome Relationship in Critical Care: A Systematic Review and Meta-analysis**. *Chest* 2015, **148**(1):79-92.
150. Viswanathan M, Patnode CD, Berkman ND, Bass EB, Chang S, Hartling L, Murad MH, Treadwell JR, Kane RL: **Recommendations for assessing the risk of bias in systematic reviews of health-care interventions**. *Journal of clinical epidemiology* 2018, **97**:26-34.
151. Galbraith RF: **A note on graphical presentation of estimated odds ratios from several clinical trials**. *Stat Med* 1988, **7**(8):889-894.
152. Higgins JP, Thompson SG, Deeks JJ, Altman DG: **Measuring inconsistency in meta-analyses**. *BMJ* 2003, **327**(7414):557-560.

153. Begg CB, Mazumdar M: **Operating characteristics of a rank correlation test for publication bias**. *Biometrics* 1994, **50**(4):1088-1101.
154. Shi L, Lin L: **The trim-and-fill method for publication bias: practical guidelines and recommendations based on a large database of meta-analyses**. *Medicine (Baltimore)* 2019, **98**(23):e15987.
155. **Service delivery and safety: Improving the prevention, diagnosis and clinical management of sepsis** [https://www.who.int/servicedeliverysafety/areas/sepsis/en/]
156. Ikoma N, Kim B, Elting LS, Shih YT, Badgwell BD, Mansfield P: **Trends in Volume-Outcome Relationship in Gastrectomies in Texas**. *Ann Surg Oncol* 2019, **26**(9):2694-2702.
157. Lee KC, Sethuraman K, Yong J: **On the Hospital Volume and Outcome Relationship: Does Specialization Matter More Than Volume?** *Health Serv Res* 2015, **50**(6):2019-2036.
158. O'Donnell TFX, Boitano LT, Deery SE, Lancaster RT, Siracuse JJ, Schermerhorn ML, Scali ST, Patel VI: **Hospital Volume Matters: The Volume-Outcome Relationship in Open Juxtarenal AAA Repair**. *Ann Surg* 2020, **271**(1):184-190.
159. Sewalt CA, Wiegers EJA, Venema E, Lecky FE, Schuit SCE, Den Hartog D, Lingsma HF: **The volume-outcome relationship in severely injured patients: A systematic review and meta-analysis**. *J Trauma Acute Care Surg* 2018, **85**(4):810-819.
160. Luft HS, Bunker JP, Enthoven AC: **Should operations be regionalized? The empirical relation between surgical volume and mortality**. *N Engl J Med* 1979, **301**(25):1364-1369.
161. Alsfasser G, Leicht H, Gunster C, Rau BM, Schillinger G, Klar E: **Volume-outcome relationship in pancreatic surgery**. *Br J Surg* 2016, **103**(1):136-143.
162. Wright JD: **The Volume-Outcome Paradigm for Gynecologic Surgery: Clinical and Policy Implications**. *Clin Obstet Gynecol* 2020.
163. Gu WJ, Wu XD, Zhou Q, Zhang J, Wang F, Ma ZL, Gu XP: **Relationship between Annualized Case Volume and Mortality in Sepsis: A Dose-Response Meta-analysis**. *Anesthesiology* 2016, **125**(1):168-179.
164. Walkey AJ, Wiener RS: **Hospital case volume and outcomes among patients hospitalized with severe sepsis**. *Am J Respir Crit Care Med* 2014, **189**(5):548-555.
165. Shahin J, Harrison DA, Rowan KM: **Relation between volume and outcome for patients with severe sepsis in United Kingdom: retrospective cohort study**. *BMJ* 2012, **344**:e3394.
166. Peelen L, de Keizer NF, Peek N, Scheffer GJ, van der Voort PH, de Jonge E: **The influence of volume and intensive care unit organization on hospital mortality in patients admitted with severe sepsis: a retrospective multicentre cohort study**. *Crit Care* 2007, **11**(2):R40.
167. Engel C, Brunkhorst FM, Bone HG, Brunkhorst R, Gerlach H, Grond S, Gruendling M, Huhle G, Jaschinski U, John S *et al*: **Epidemiology of sepsis in Germany: results from a national prospective multicenter study**. *Intensive Care Med* 2007, **33**(4):606-618.
168. Reinikainen M, Karlsson S, Varpula T, Parviainen I, Ruokonen E, Varpula M, Ala-Kokko T, Pettila V: **Are small hospitals with small intensive care units able to treat patients with severe sepsis?** *Intensive Care Med* 2010, **36**(4):673-679.
169. Needham DM, Bronskill SE, Rothwell DM, Sibbald WJ, Pronovost PJ, Laupacis A, Stukel TA: **Hospital volume and mortality for mechanical ventilation of medical and**

surgical patients: a population-based analysis using administrative data. *Crit Care Med* 2006, **34**(9):2349-2354.
170. Hawks L, Himmelstein DU, Woolhandler S, Bor DH, Gaffney A, McCormick D: **Trends in Unmet Need for Physician and Preventive Services in the United States, 1998-2017**. *JAMA Intern Med* 2020, **180**(3):439-448.
171. Dudley RA, Johansen KL, Brand R, Rennie DJ, Milstein A: **Selective referral to high-volume hospitals: estimating potentially avoidable deaths**. *JAMA* 2000, **283**(9):1159-1166.
172. Joseph B, Morton JM, Hernandez-Boussard T, Rubinfeld I, Faraj C, Velanovich V: **Relationship between hospital volume, system clinical resources, and mortality in pancreatic resection**. *J Am Coll Surg* 2009, **208**(4):520-527.
173. Macomber CW, Shaw JJ, Santry H, Saidi RF, Jabbour N, Tseng JF, Bozorgzadeh A, Shah SA: **Centre volume and resource consumption in liver transplantation**. *HPB (Oxford)* 2012, **14**(8):554-559.
174. Nguyen YL, Milbrandt EB, Weissfeld LA, Kahn JM, Chiche JD, Aegerter P, Clermont G, Kellum JA, Guidet B, Angus DC: **Intensive care unit renal support therapy volume is not associated with patient outcome**. *Crit Care Med* 2011, **39**(11):2470-2477.
175. Vincent JL, Marshall JC, Namendys-Silva SA, Francois B, Martin-Loeches I, Lipman J, Reinhart K, Antonelli M, Pickkers P, Njimi H *et al*: **Assessment of the worldwide burden of critical illness: the intensive care over nations (ICON) audit**. *Lancet Respir Med* 2014, **2**(5):380-386.
176. Morris S, Ramsay AIG, Boaden RJ, Hunter RM, McKevitt C, Paley L, Perry C, Rudd AG, Turner SJ, Tyrrell PJ *et al*: **Impact and sustainability of centralising acute stroke services in English metropolitan areas: retrospective analysis of hospital episode statistics and stroke national audit data**. *BMJ* 2019, **364**:l1.
177. Harrison DA, Lone NI, Haddow C, MacGillivray M, Khan A, Cook B, Rowan KM: **External validation of the Intensive Care National Audit & Research Centre (ICNARC) risk prediction model in critical care units in Scotland**. *BMC Anesthesiol* 2014, **14**:116.
178. Harrison DA, Parry GJ, Carpenter JR, Short A, Rowan K: **A new risk prediction model for critical care: the Intensive Care National Audit & Research Centre (ICNARC) model**. *Crit Care Med* 2007, **35**(4):1091-1098.
179. Harrison DA, Welch CA, Eddleston JM: **The epidemiology of severe sepsis in England, Wales and Northern Ireland, 1996 to 2004: secondary analysis of a high quality clinical database, the ICNARC Case Mix Programme Database**. *Crit Care* 2006, **10**(2):R42.
180. Vandenbroucke JP, von Elm E, Altman DG, Gotzsche PC, Mulrow CD, Pocock SJ, Poole C, Schlesselman JJ, Egger M, Initiative S: **Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration**. *Epidemiology* 2007, **18**(6):805-835.
181. Royston P: **Model Selection for Univariable Fractional Polynomials**. *The Stata Journal* 2017, **17**(3):619-629.
182. Lusa L, Ahlin C: **Restricted cubic splines for modelling periodic data**. *PLoS One* 2020, **15**(10):e0241364.
183. Rabe-Hesketh S, Skrondal A, Skrondal A: **Multilevel and longitudinal modeling using stata**, 2nd edn. College Station, Tex.: Stata Press Publication; 2008.

184. Austin PC, Wagner P, Merlo J: **The median hazard ratio: a useful measure of variance and general contextual effects in multilevel survival analysis**. *Stat Med* 2017, **36**(6):928-938.
185. Royston P: **Model selection for univariable fractional polynomials**. *Stata J* 2017, **17**(3):619-629.
186. Binder H, Sauerbrei W: **Adding local components to global functions for continuous covariates in multivariable regression modeling**. *Stat Med* 2010, **29**(7-8):808-817.
187. Ding P, VanderWeele TJ: **Sensitivity Analysis Without Assumptions**. *Epidemiology* 2016, **27**(3):368-377.
188. VanderWeele TJ, Ding P: **Sensitivity Analysis in Observational Research: Introducing the E-Value**. *Ann Intern Med* 2017, **167**(4):268-274.
189. Schunck R, Perales F: **Within- and Between-cluster Effects in Generalized Linear Mixed Models: A Discussion of Approaches and the Xthybrid command**. *The Stata Journal* 2017, **17**(1):89-115.
190. Panageas KS, Schrag D, Riedel E, Bach PB, Begg CB: **The effect of clustering of outcomes on the association of procedure volume and surgical outcomes**. *Ann Intern Med* 2003, **139**(8):658-665.
191. Abbenbroek B, Duffield CM, Elliott D: **The intensive care unit volume-mortality relationship, is bigger better? An integrative literature review**. *Aust Crit Care* 2014, **27**(4):157-164; quiz 165.
192. Finks JF, Osborne NH, Birkmeyer JD: **Trends in hospital volume and operative mortality for high-risk surgery**. *N Engl J Med* 2011, **364**(22):2128-2137.
193. Maharaj R, McGuire A, Street A: **Association of Annual Intensive Care Unit Sepsis Caseload With Hospital Mortality From Sepsis in the United Kingdom, 2010-2016**. *JAMA Netw Open* 2021, **4**(6):e2115305.
194. Mesman R, Westert GP, Berden BJ, Faber MJ: **Why do high-volume hospitals achieve better outcomes? A systematic review about intermediate factors in volume-outcome relationships**. *Health Policy* 2015, **119**(8):1055-1067.
195. Nicholas LH, Stith SS: **Are Transplant Centers That Meet Insurer Minimum Volume Requirements Better Quality?** *Med Care Res Rev* 2021, **78**(5):502-510.
196. Leung S, Pastores SM, Oropello JM, Lilly CM, Galvagno SM, Jr., Badjatia N, Jacobi J, Herr DL, Oliveira JD, Academic Leaders in Critical Care Medicine Task Force of the Society of Critical Care M: **Regionalization of Critical Care in the United States: Current State and Proposed Framework From the Academic Leaders in Critical Care Medicine Task Force of the Society of Critical Care Medicine**. *Crit Care Med* 2022, **50**(1):37-49.
197. Gaffney AW: **ICU Equity and Regionalization in the COVID-19 Era**. *Ann Am Thorac Soc* 2022.
198. Cooper Z, Gibbons S, Jones S, McGuire A: **Does Hospital Competition Save Lives? Evidence from the English NHS Patient Choice Reforms**. *Econ J (London)* 2011, **121**(554):F228-F260.
199. Gaynor M, Seider H, Vogt WB: **The Volume-Outcome Effect, Scale Economies, and Learning-by-Doing**. *The American Economic Review* 2005, **95**:243-247.
200. Van Gestel R, Muller T, Bosmans J: **Does my high blood pressure improve your survival? Overall and subgroup learning curves in health**. *Health Econ* 2017, **26**(9):1094-1109.

201. Ho V: **Learning and the evolution of medical technologies: the diffusion of coronary angioplasty**. *J Health Econ* 2002, **21**(5):873-885.
202. Ho SS, Chan YL, Yeung DK, Metreweli C: **Blood flow volume quantification of cerebral ischemia: comparison of three noninvasive imaging techniques of carotid and vertebral arteries**. *AJR Am J Roentgenol* 2002, **178**(3):551-556.
203. Avdic D, Lundborg P, Vikstrom J: **Estimating returns to hospital volume: Evidence from advanced cancer surgery**. *J Health Econ* 2019, **63**:81-99.
204. World Health A: **Improving the prevention, diagnosis and clinical management of sepsis**. In. Geneva: World Health Organization; 2017.
205. Walton NT, Mohr NM: **Concept review of regionalized systems of acute care: Is regionalization the next frontier in sepsis care?** *J Am Coll Emerg Physicians Open* 2022, **3**(1):e12631.
206. Vandenbroucke JP, von Elm E, Altman DG, Gotzsche PC, Mulrow CD, Pocock SJ, Poole C, Schlesselman JJ, Egger M, Initiative S: **Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration**. *Int J Surg* 2014, **12**(12):1500-1524.
207. Benkard CL: **Learning and Forgetting: The Dynamics of Aircraft Production**. *American Economic Review* 2000, **90**(4):1034-1054.
208. **Causality, learning and forgetting in surgery.** [http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.183.4869.]
209. Ervin JN, Kahn JM, Cohen TR, Weingart LR: **Teamwork in the intensive care unit**. *Am Psychol* 2018, **73**(4):468-477.
210. Gowrisankaran G, Olin JM: **Causality , Learning and Forgetting in Surgery**. In*: 2005*; 2005.
211. Facchini G: **Forgetting-by-not-doing: The case of surgeons and cesarean sections**. *Health Econ* 2022, **31**(3):481-495.
212. Shuhaiber JH, Goldsmith K, Nashef SA: **Impact of cardiothoracic resident turnover on mortality after cardiac surgery: a dynamic human factor**. *Ann Thorac Surg* 2008, **86**(1):123-130; discussion 130-121.
213. Bozic KJ, Maselli J, Pekow PS, Lindenauer PK, Vail TP, Auerbach AD: **The influence of procedure volumes and standardization of care on quality and efficiency in total joint replacement surgery**. *J Bone Joint Surg Am* 2010, **92**(16):2643-2652.
214. Huguet M: **Centralization of care in high volume hospitals and inequalities in access to care**. *Soc Sci Med* 2020, **260**:113177.
215. Littek W: **Labor, Division of**. In: *International Encyclopedia of the Social & Behavioral Sciences.* Edited by Smelser NJ, Baltes PB. Oxford: Pergamon; 2001: 8220-8226.
216. Ricardo DMJR: **The works of David Ricardo**. London: J. Murray; 1888.
217. Eastaugh SR: **Hospital costs and specialization: benefits of limiting the number of product lines**. *J Health Care Finance* 2009, **36**(2):24-34.
218. Zwanziger J, Melnick GA, Simonson L: **Differentiation and specialization in the California hospital industry 1983 to 1988**. *Med Care* 1996, **34**(4):361-372.
219. Lee KS, Chun KH, Lee JS: **Reforming the hospital service structure to improve efficiency: urban hospital specialization**. *Health Policy* 2008, **87**(1):41-49.
220. **Health at a glance. Europe 2020**
221. Wunsch H, Gershengorn H, Scales DC: **Economics of ICU organization and management**. *Critical care clinics* 2012, **28**(1):25-v.

222. Evans L, Best C: **Managing malnutrition in patients with dementia**. *Nurs Stand* 2015, **29**(28):50-57.
223. Barro JR, Huckman RS, Kessler DP: **The effects of cardiac specialty hospitals on the cost and quality of medical care**. *J Health Econ* 2006, **25**(4):702-721.
224. Grasic K, Mason AR, Street A: **Paying for the quantity and quality of hospital care: the foundations and evolution of payment policy in England**. *Health Econ Rev* 2015, **5**(1):50.
225. Gaughan J, Gutacker N, Grasic K, Kreif N, Siciliani L, Street A: **Paying for efficiency: Incentivising same-day discharges in the English NHS**. *J Health Econ* 2019, **68**:102226.
226. Schneider JE, Miller TR, Ohsfeldt RL, Morrisey MA, Zelner BA, Pengxiang L: **The economics of specialty hospitals**. *Med Care Res Rev* 2008, **65**(5):531-553; discussion 554-563.
227. Barro JR, Huckman RS, Kessler DP: **The Effects of Cardiac Specialty Hospitals on the Cost and Quality of Medical Care**. *Health Economics eJournal* 2006.
228. KC DS, Terwiesch C: **The Effects of Focus on Performance: Evidence from California Hospitals**. *Management Science* 2011, **57**(11):1897-1912.
229. Huckman RS, Zinner DE: **Does Focus Improve Operational Performance? Lessons from the Management of Clinical Trials**. *Strategic Management Journal* 2008, **29**(2):173-193.
230. W. S: **The focused factory**. *Harvard Business Review* 1974, **52**:113-120.
231. Kuntz L, Scholtes S, Sülz S: **Separate and Concentrate: Accounting for Patient Complexity in General Hospitals**. *Management Science* 2019, **65**(6):2482-2501.
232. Hagen TP, Vaughan-Sarrazin MS, Cram P: **Relation between hospital orthopaedic specialisation and outcomes in patients aged 65 and older: retrospective analysis of US Medicare data**. *BMJ* 2010, **340**:c165.
233. Schneider JE, Miller TR, Ohsfeldt RL, Morrisey MA, Zelner BA, Li P: **The Economics of Specialty Hospitals**. *Medical Care Research and Review* 2008, **65**(5):531-553.
234. Regge MD, Trybou J, Gemmel P, Duyck P, Annemans L: **Physician-owned specialized facilities: focused factories or destructive competition?: a systematic review**. In*: 2013*; 2013.
235. Kramer AH, Zygun DA: **Neurocritical care: why does it make a difference?** *Current Opinion in Critical Care* 2014, **20**(2):174-181.
236. Goldfarb M, van Diepen S, Liszkowski M, Jentzer JC, Pedraza I, Cercek B: **Noncardiovascular Disease and Critical Care Delivery in a Contemporary Cardiac and Medical Intensive Care Unit**. *Journal of Intensive Care Medicine* 2019, **34**(7):537-543.
237. Cooper Z, Gibbons S, Skellern M: **Does competition from private surgical centres improve public hospitals' performance? Evidence from the English National Health Service**. *Journal of Public Economics* 2018, **166**:63-80.
238. Regge MD, Trybou J, Gemmel P, Duyck P, Annemans L: **Physician-Owned Specialized Facilities: Focused Factories or Destructive Competition?** *Academy of Management Proceedings* 2013, **2013**(1):13667.
239. Swanson A: **Physician investment in hospitals: Specialization, selection, and quality in cardiac care**. *J Health Econ* 2021, **80**:102519.

240. Gabel JR, Fahlman C, Kang R, Wozniak G, Kletke P, Hay JW: **Where do I send thee? Does physician-ownership affect referral patterns to ambulatory surgery centers?** *Health Aff (Millwood)* 2008, **27**(3):w165-174.
241. Clark JR, Huckman RS: **Broadening Focus: Spillovers, Complementarities, and Specialization in the Hospital Industry**. *Management Science* 2012, **58**(4):708-722.
242. Conrad Kobel ET: **Hospital specialisation within a DRG-Framework: The Austrian Case**. In: *Faculty of Economics and Statistics, University of Innsbruck.* Edited by Papers W; 2013.
243. Lindlbauer I, Schreyögg J: **The relationship between hospital specialization and hospital efficiency: do different measures of specialization lead to different results?** *Health Care Management Science* 2014, **17**(4):365-378.
244. Buzzelli M, Su J: **Multi-level modelling in health research: a caution and rejoinder on temporally mismatched data**. *Soc Sci Med* 2006, **62**(5):1215-1218.
245. Bell A, Fairbrother M, Jones K: **Fixed and random effects models: making an informed choice**. *Quality & Quantity* 2019, **53**(2):1051-1074.
246. Ferrando-Vivas P, Jones A, Rowan KM, Harrison DA: **Development and validation of the new ICNARC model for prediction of acute hospital mortality in adult critical care**. *Journal of Critical Care* 2017, **38**:335-339.
247. Mielke D, Malinova V, Moerer O, Suntheim P, Voit M, Rohde V: **Does the subspecialty of an intensive care unit (ICU) has an impact on outcome in patients suffering from aneurysmal subarachnoid hemorrhage?** *Neurosurgical Review* 2019, **42**(1):147-153.
248. Kelly FE, Fong K, Hirsch N, Nolan JP: **Intensive care medicine is 60 years old: the history and future of the intensive care unit**. *Clin Med (Lond)* 2014, **14**(4):376-379.
249. Costa DK, Kahn JM: **Organizing Critical Care for the 21st Century**. *JAMA* 2016, **315**(8):751-752.
250. Collaborators C-EM: **Estimating excess mortality due to the COVID-19 pandemic: a systematic analysis of COVID-19-related mortality, 2020-21**. *Lancet* 2022, **399**(10334):1513-1536.
251. Gaffney AW: **Intensive Care Unit Equity and Regionalization in the COVID-19 Era**. *Ann Am Thorac Soc* 2022, **19**(5):717-719.
252. McCrum ML, Lipsitz SR, Berry WR, Jha AK, Gawande AA: **Beyond volume: does hospital complexity matter?: an analysis of inpatient surgical mortality in the United States**. *Med Care* 2014, **52**(3):235-242.
253. Arabi YM, Azoulay E, Al-Dorzi HM, Phua J, Salluh J, Binnie A, Hodgson C, Angus DC, Cecconi M, Du B *et al*: **How the COVID-19 pandemic will change the future of critical care**. *Intensive Care Med* 2021, **47**(3):282-291.
254. de Cruppe W, Malik M, Geraedts M: **Minimum volume standards in German hospitals: do they get along with procedure centralization? A retrospective longitudinal data analysis**. *BMC Health Serv Res* 2015, **15**:279.
255. Subramanian MP, Yang Z, Chang SH, Willis D, Zhang J, Semenkovich TR, Heiden BT, Kozower BD, Kreisel D, Meyers BF *et al*: **Minimum Volume Standards for Surgical Care of Early-Stage Lung Cancer: A Cost-Effectiveness Analysis**. *Ann Thorac Surg* 2022.
256. MacKenzie EJ, Rivara FP, Jurkovich GJ, Nathens AB, Frey KP, Egleston BL, Salkever DS, Scharfstein DO: **A national evaluation of the effect of trauma-center care on mortality**. *N Engl J Med* 2006, **354**(4):366-378.

257. Birkmeyer JD, Finlayson EV, Birkmeyer CM: **Volume standards for high-risk surgical procedures: potential benefits of the Leapfrog initiative**. *Surgery* 2001, **130**(3):415-422.
258. Staebler S: **Regionalized systems of perinatal care: health policy considerations**. *Adv Neonatal Care* 2011, **11**(1):37-42.
259. Davie JH, R.M.; Mountford, J.; Morris, S. : **London's Hyperacute Stroke Units Improve Outcomes and Lower Costs**. *Harvard Business Review* 2013.
260. Barnato AE, Kahn JM, Rubenfeld GD, McCauley K, Fontaine D, Frassica JJ, Hubmayr R, Jacobi J, Brower RG, Chalfin D *et al*: **Prioritizing the organization and management of intensive care services in the United States: the PrOMIS Conference**. *Crit Care Med* 2007, **35**(4):1003-1011.
261. FICM: **WORKFORCE DATABANK for Adult Critcal Care**. In. Churchill House | 35 Red Lion Square | London | WC1R 4SG: The Faculty of Intensive Care Medicine 2021.
262. Prescott HC, Langa KM, Liu V, Escobar GJ, Iwashyna TJ: **Increased 1-year healthcare use in survivors of severe sepsis**. *Am J Respir Crit Care Med* 2014, **190**(1):62-69.
263. Corrales-Medina VF, Alvarez KN, Weissfeld LA, Angus DC, Chirinos JA, Chang CC, Newman A, Loehr L, Folsom AR, Elkind MS *et al*: **Association between hospitalization for pneumonia and subsequent risk of cardiovascular disease**. *JAMA* 2015, **313**(3):264-274.
264. Iwashyna TJ, Ely EW, Smith DM, Langa KM: **Long-term cognitive impairment and functional disability among survivors of severe sepsis**. *JAMA* 2010, **304**(16):1787-1794.
265. Hsia RY, Krumholz H, Shen YC: **Evaluation of STEMI Regionalization on Access, Treatment, and Outcomes Among Adults Living in Nonminority and Minority Communities**. *JAMA Netw Open* 2020, **3**(11):e2025874.
266. Mullen MT, Wiebe DJ, Bowman A, Wolff CS, Albright KC, Roy J, Balcer LJ, Branas CC, Carr BG: **Disparities in Accessibility of Certified Primary Stroke Centers**. *Stroke* 2014, **45**(11):3381-3388.
267. Goldstein H: **Hierarchical Data Modeling in the Social Sciences**. *Journal of Educational and Behavioral Statistics* 1995, **20**(2):201-204.
268. Goldstein H, Spiegelhalter DJ: **League Tables and Their Limitations: Statistical Issues in Comparisons of Institutional Performance**. *Journal of the Royal Statistical Society Series A (Statistics in Society)* 1996, **159**(3):385-443.
269. Orsini N, Greenland S: **A procedure to tabulate and plot results after flexible modeling of a quantitative covariate**. *Stata Journal* 2011, **11**(1):1-29.
270. Smith OM, Chant C, Burns KEA, Kaur M, Ashraf S, DosSantos CC, Hwang SW, Friedrich JO: **Characteristics, clinical course, and outcomes of homeless and non-homeless patients admitted to ICU: A retrospective cohort study**. *PLoS One* 2017, **12**(6):e0179207.
271. Ginde AA, Moss M, Shapiro NI, Schwartz RS: **Impact of older age and nursing home residence on clinical outcomes of US emergency department visits for severe sepsis**. *J Crit Care* 2013, **28**(5):606-611.
272. **English indices of deprivation** [https://www.gov.uk/government/collections/english-indices-of-deprivation]
273. Knaus WA, Draper EA, Wagner DP, Zimmerman JE: **APACHE II: a severity of disease classification system**. *Crit Care Med* 1985, **13**(10):818-829.

274. Harrison DA, Rowan KM: **Outcome prediction in critical care: the ICNARC model**. *Curr Opin Crit Care* 2008, **14**(5):506-512.
275. Levy MM, Artigas A, Phillips GS, Rhodes A, Beale R, Osborn T, Vincent JL, Townsend S, Lemeshow S, Dellinger RP: **Outcomes of the Surviving Sepsis Campaign in intensive care units in the USA and Europe: a prospective cohort study**. *Lancet Infect Dis* 2012, **12**(12):919-924.
276. Hastie T, Friedman J, Tisbshirani R: **The Elements of statistical learning : data mining, inference, and prediction**. New York: Springer; 2018.
277. Gelfand AE, Dey DK: **Bayesian Model Choice: Asymptotics and Exact Calculations**. *Journal of the Royal Statistical Society Series B (Methodological)* 1994, **56**(3):501-514.
278. Cameron AC, Triverdi PK: **Microeconometrics using Stata**. College Station, Tex: Stata Press; 2010.
279. Allison PD: **Fixed Effects Regression Models**. Thousand Oaks, Calif.: Sage Publications; 2009.