USING SURVIVAL ANALYSIS TO PREDICT SYSTEM FAILURE IN NHS TRUSTS

Can survival Analysis predict system strain and support proactive intervention in NHS Cancer Care?

Abstract

The UK's National Health Service (NHS) is under sustained pressure, reflected in its persistent failure to meet the 62-day cancer Referral to Treatment (RTT) standard. This study applied a quantitative, survival-based approach, using a Cox Proportional Hazards (CoxPH) model on 410,378 NHS trust-month observations (2009–2020) to quantify the hazard of service failure. The analysis showed that cancer delays act as a dependable proxy for deep structural fragility. Results confirmed a critical treatment bottleneck, with admitted cases experiencing a 38 percent higher breach rate, and highlighted several contradictory systemic behaviours. Activity Diversity emerged as the strongest risk factor (HR 2.94), challenging the assumption that service diversification improves resilience and instead pointing toward resource fragmentation. In contrast, operational efficiency metrics, such as referral conversion, were strongly protective. The predictive artefact developed in this study offers a new, evidence-led framework for proactive NHS governance, enabling resource allocation based on measurable risk rather than reactive performance thresholds.

Nima Osman

No23451

Table of Contents

Declaration

I hereby declare that this dissertation, submitted for the master's degree in data science, is entirely my own work and has not been submitted in whole or in part for any degree or professional qualification at this or any other university.

Statement of originality and academic integrity:

1. I confirm that all material incorporated into this dissertation that is not my own original work is acknowledged and referenced in accordance with the University's regulations regarding plagiarism and academic conduct.
2. I confirm that the data used for the quantitative analysis (Cancer Waiting Times, Hospital Activity, and Workforce data) were sourced from publicly available NHS digital datasets, and no primary data collection involving human participants was undertaken.

Acknowledgment

This dissertation marks the culmination of an intensive period of study and independent research. I owe my deepest gratitude to the individuals and institutions that made this project possible. Foremost, I would like to express my sincere appreciation to my supervisor, Nawaz Khan, for their guidance, patience and invaluable advice. Their direction was essential for navigating the complexities of NHS datasets and for shaping the analytical rigour of the Cox Proportional Hazards model.

I also wish to thank the NHS Digital for making the necessary public data available, which served as the foundation for this research.

Chapter 1

Introduction

1.1 The problem

The UK's National Health Service (NHS) is currently under extraordinary pressure, raising urgent questions about its long-term sustainability and its ability to maintain consistent standards of care. These systemic challenges are no longer abstract concerns; patients are experiencing real harm as care quality increasingly falls short of clinical and public expectations. Immediate, coordinated action is needed to address the dual crises of workforce shortages and capacity limitations, with both short-term mitigation and long-term structural reform essential to safeguarding public health (Cooksley et al., 2023).

A clear and measurable manifestation of this systemic strain is the NHS's diminishing ability to meet its own time-based performance targets for critical treatments. Nowhere is this more consequential than in oncology, where delays in diagnosis and treatment are directly correlated with worsened clinical outcomes and higher mortality rates (OECD, 2020). This project focuses specifically on these delays within cancer services as a proxy for wider NHS dysfunction.

Since the introduction of the NHS Cancer Plan in 2000 and its full implementation in 2005, specific standards have guided the delivery of timely cancer care. Two key benchmarks include the 31-day Decision to First Treatment (DTT) standard, which measures the time from treatment decision to commencement for both new and recurrent cancers, and the 62-day Referral to Treatment (RTT) standard, which tracks the time from urgent GP referral to treatment initiation for new cancer diagnoses.

However, adherence to these targets has sharply deteriorated. In 2010/11, 96% of patients with urgent GP referrals were seen by specialists within two weeks. By 2022/23, this figure had fallen to 85%. The 62-day RTT target saw an even steeper decline from 88% compliance (95% CI: 85–93) in 2010/11 to just 64% (95% CI: 53–73) in 2022/23 (Liaqat et al., 2022). Most concerning is the system's persistent failure to meet the 85% RTT threshold, with over 24,555 patients in

September 2023 waiting beyond 62 days for treatment, a dramatic rise from 11,000 in 2020 and a peak of 34,000 in May 2020 (Quinn et al., 2024).

These failures stem from a confluence of systemic factors, including chronic understaffing, inadequate budgets, overwhelmed care pathways, and insufficient diagnostic infrastructure. The consequences are stark: worsened health outcomes, increased mortality, and diminished patient trust, particularly among vulnerable populations such as ethnic minorities and the elderly.

Cancer, as a disease, demands timely intervention. It is estimated to affect 38.9% of men and women during their lifetime, and the World Health Organisation reported in 2020 that it accounted for 1 in 6 deaths globally (WHO, 2025; National Cancer Institute, 2024). Timely diagnosis and treatment are the most effective strategies for improving prognosis, making persistent NHS delays not only unacceptable but dangerous.

When a healthcare system repeatedly fails to meet its performance standards, it constitutes a service failure, a term often used in business but increasingly applicable to health systems facing repeated underperformance (Sands et al., 2020). Because oncology is typically prioritised within hospital care pipelines, delays in cancer services signal a broader crisis: if even high-priority conditions cannot be managed effectively, the situation for other conditions is likely far worse.

A stark example of the consequences of healthcare collapse can be seen in Greece during the early 2010s. Austerity-led funding cuts of up to 40% in public hospitals, coupled with staff shortages and resource limitations, led to a 40% increase in cancer mortality over just a few years. Greek oncology departments experienced mass burnout, unsafe working conditions, and a collapse in timely care delivery, highlighting how economic and structural strain can lead to systemic failure (Pittaka et al., 2022).

1.2 Motivation and Rationale

This project is driven by the critical need to shift from reactive management of NHS cancer care delays to a proactive, data-driven approach**.** Delays in oncology are not merely administrative lapses; they directly influence patient survival. Despite clear benchmarks, such as the 62-day RTT standard, breaches continue, pointing to systemic dysfunction that endangers lives and erodes public confidence.

These treatment delays are more than isolated clinical incidents; they are predictive signals of broader health system failure. The Greek oncology crisis serves as a powerful warning of what can occur when healthcare systems ignore or inadequately respond to accumulating pressure. A failure in a high-priority area like cancer care is a “canary in the coal mine,” signalling deep structural cracks that could precipitate collapse across the wider system.

The availability of publicly accessible data greatly supports the feasibility of this research, as monthly NHS data on cancer waiting times at the trust level. This allows for granular analysis without the delays of Freedom of Information requests and supports rigorous modelling of localised system performance.

This study aims not just to describe the problem, but to drive policy impact**.** Existing monitoring frameworks often lack the precision to support targeted interventions, and threshold-driven systems may even encourage “gaming” behaviour rather than genuine improvement (Li & Evans, 2022). Moreover, most research has focused on clinical outcomes or retrospective audit data, largely overlooking how organisational behaviour and resource limitations contribute to persistent underperformance.

By developing a predictive early-warning model using survival analysis and machine learning, this project fills a crucial research gap. It will go beyond surface-level trend analysis to simulate the real-world effects of interventions such as workforce reallocation on reducing treatment delays. These simulations can inform trust-specific, actionable recommendations**.**

1.3 Research question

Can survival analysis predict systemic strain and support proactive intervention in NHS cancer care?

1.4 Aim and Objectives

Project Aims:

To predict the risk of critical failure in NHS cancer services by analysing how key factors, hospital activity, referral volumes, emergency-to-elective ratios, and workforce levels, affect 62-day treatment delays.

To generate actionable insights that help policymakers allocate resources proactively and support high-risk trusts.

Supporting Objectives:

1. Compile and integrate datasets on 62-day RTT cancer waiting times, workforce levels, and hospital activity across NHS trusts.
2. Implement a Cox Proportional Hazards model to define and analyse service failure, where a trust exceeds the 62-day target by more than 20%.
3. Develop and validate a predictive survival model to quantify each trust's relative risk of breaching the standard.
4. Use the model for policy simulations, adjusting key covariates (e.g., workforce or referral rates) to generate data-driven recommendations to mitigate service failures and enhance system resilience.

1.5 Scope of this research

- The study examines NHS England using public cancer waiting time data (2010-2023).
- Focuses on the 62-day Referral to Treatment (RTT) standard as a measure of systemic strain.
- Analyses key covariates: workforce, referral volumes, and hospital activity.
- Uses only secondary, anonymised public data, no primary data collection or participant interaction.
- Excludes financial data, race, and socioeconomic factors due to limited availability.

- Workforce morale, satisfaction, and burnout are not assessed for the same reason.
- Aims to identify underperforming trusts and inform policy-based intervention, not clinical ones.

1.6 Significance and Contributions

It will utilise the Cox Proportional Hazards (CoxPH) model to address the strain on NHS cancer services and its direct impact on patient outcomes. By focusing on cancer treatment delays as a proxy for systemic pressure, this approach offers a new methodological lens for understanding systemic fragility. This project will generate actionable insights by identifying underperforming trusts and quantifying the influence of key covariates (e.g., workforce and referral volumes). This will then allow it to target resource allocation and support the development of a personalised method for targeting system strain in NHS trusts. The shift from reactive problem-solving to proactive prevention aligns with a critical need identified in the current healthcare policy debate.

1.7 Dissertation structure

This dissertation will be structured into five main chapters. Chapter 1 will introduce the critical state of NHS cancer care, outline the research problem, and detail the study's motivation. Chapter 2 will present a comprehensive literature review, synthesising existing research on cancer delays, contributing factors, and current strategies. Chapter 3 will detail the methodology, including data collection, survival analysis, counterfactual simulations, and machine learning approaches. Chapter 4 will present the results, including model findings, comparative analyses, and policy impact assessments. Finally, Chapter 5 will discuss the implications of the findings, their contributions, limitations, and future research directions.

Chapter 2

Literature review

2.1 AI/ML and Technology in Healthcare System Optimisation

Machine learning and related technologies are transforming healthcare systems by improving efficiency, predictions, and personalisation. These methods enable rapid processing of complex datasets to support clinical and operational decision-making, marking a shift towards data-driven, proactive models of care (Bajwa et al., 2021; Maleki Varnosfaderani & Forouzanfar, 2024).

2.1.1 Definition of Artificial Intelligence (AI)and Machine learning in Healthcare

Machine learning is a core component of artificial intelligence that enables systems to learn from experience and identify patterns. In healthcare, ML models analyse volumes of clinical data to support predictions, recommendations and diagnostic decisions (Esmaeilzadeh, 2024; Mennella *et al.*, 2024). Deep learning, a subset of ML, uses layered neural networks for tasks such as image and speech recognition, while natural language processing (NLP) enables computes to interpret medical text and notes (Maleki Varnosfaderani & Forouzanfar, 2024). Collectively, these methods enhance access to care, diagnostic accuracy, and operational efficiency (Dicuonzo *et al.*, 2022).

2.1.2 Applications of AI/ML in Healthcare

ML and DL have been applied across clinical and operational areas:

A. Diagnostics and Decision Support:

AI/ML algorithms analyse medical images and patient data to detect disease and predict outcomes, often with accuracy comparable to that of clinical experts (Varnosfaderani & Forouzanfar, 2024).

These models assist in identifying cancer types, assessing readmission risks and discovering new biomarkers (Li *et al.*, 2024; Sai *et al.*, 2024).

B. Personalised Medicine

They support tailored treatment by analysing genomic and clinical data to predict drug response and disease progression. It also accelerates drug discovery by modelling treatment effects and patients' reactions (Maleki Varnosfaderani & Forouzanfar, 2024).

C. Patient Monitoring and Support:

AI/ML-driven wearables and virtual assistants continuously monitor patient data, issue alerts, and provide real-time support, thereby improving engagement and early detection of complications (Maleki Varnosfaderani & Forouzanfar, 2024; Lyon *et al.*, 2021).

D. Operational optimisation:

Within hospital systems, AI/ML streamlines scheduling, predicts admission trends, and manages resources to reduce waiting times and enhance patient flow (Maleki Varnosfaderani & Forouzanfar, 2024; Dicuonzo *et al.*, 2022).

2.1.3 Benefits and challenges

Machine learning enhances healthcare by improving diagnostic precision, personalising treatment, and optimising system performance (Bajwa *et al.*, 2021; Li *et al.*, 2024). It reduces administrative burdens through automation and strengthens healthcare resilience by enabling faster, data-driven responses to demand fluctuations (Rane *et al.*, 2024).

However, challenges remain around data quality, transparency and integration. AI and ML models rely on large, accurate datasets, yet healthcare data are often fragmented and inconsistent (Esmaeilzadeh, 2024). Algorithmic bias, especially when training data lacks diversity, can lead to unequal performance across patient groups (Li *et al.*, 2024; Pasricha, 2023). The "black box' nature of deep learning models also raises concerns about interpretability and trust (Esmaeilzadeh, 2024).

Furthermore, integration with legacy systems and regulatory oversight requirements slows large-scale adoption (Li *et al.*, 2024; Maleki Varnosfaderani & Forouzanfar, 2024).

2.1.4 Ethical and Regulatory Considerations:

The ethical deployment of ML in healthcare depends on maintaining transparency, accountability, and patient autonomy. Systems must be explainable, secure and rigorously tested to prevent harm and ensure trust (Mennella *et al.*, 2024; Pasricha, 2023). Clear frameworks are needed to define responsibility when errors occur, particularly in clinical decision-making contexts (Esmaeilzadeh, 2024).

2.2 The Concept of Healthcare System Collapse and Early Warning Indicators

A healthcare system enters a state of failure when it consistently fails to meet its performance standards, a situation comparable to service failure in business. Sands et al. (2020) argue that service failures are inherent in human-led systems, due to their susceptibility to error and variance. In complex systems like healthcare, early warning indicators (EWIs) are measurable signals that highlight emerging risks or potential breakdowns, enabling timely intervention. Unlike simple leading indicators, EWIs involve a network of adaptive variables that shift roles during different system phases, offering a dynamic and comprehensive prediction of vulnerability (Eurostat, 2025).

Long wait times are one of the most prominent EWIs in universal healthcare systems like the NHS. As Limiri (2025) notes, these delays signify inefficiencies and poor system sustainability, often aligning with a decline in the quality of care. Contributing factors such as workforce shortages, limited resources, inadequate funding, and care pathway interdependencies further reflect internal fragility (Limiri, 2025). The outcomes of prolonged waiting include worsening health, higher mortality rates, and patient dissatisfaction, signalling a system under immense stress. This chronic strain leads to care fragmentation and burnout among healthcare workers, intensifying the risk of systemic collapse. Despite various interventions, the persistence of these issues underscores a fundamental lack of resilience and adaptation capacity (Limiri, 2025).

Aggarwal et al. (2024) describe NHS cancer services as being at a “tipping point,” where small disruptions could lead to large-scale consequences. They stress that technological innovation alone cannot solve the systemic issues in cancer care, which are deeply intertwined with broader NHS capacity, funding models, and policy structures. Specific points of pressure include treatment backlogs, underfunded staffing, and outdated reimbursement mechanisms. These bottlenecks, combined with the UK’s lagging cancer survival rates and widening healthcare inequalities, paint a stark picture of a system nearing collapse (Aggarwal et al., 2023) (Aggarwal et al., 2024).

The collapse of Greece’s oncology services provides a real-world example of system failure. Pittaka et al. (2022) detail how austerity measures following the 2010 economic crisis led to severe staffing shortages, supply deficits, and widespread burnout. With healthcare professionals underpaid and public hospitals underfunded, Greece prioritised fiscal policy over health equity and functionality, ultimately leading to systemic breakdown.

2.3 The NHS Cancer Waiting Times Landscape: Policy, Performance, and Pressures

The National Health service (NHS) in England operates under a policy framework for cancer waiting times standards, introduced through the NHS Cancer Plan in 2000 and implemented in 2005. These policies set maximum waiting periods for patients to receive care (Hayes et al., 2024).

Similar to other European systems, these targets are used for financial penalties, policy planning and guiding patient choice (OECD, 2020). However, the effectiveness and fairness of these targets, especially under growing system pressure, require critical consideration.

In October 2023, new policy changes simplified the previous ten standards into three main measures: the 28-day Faster Diagnosis Standard (FDS), the 62-day Referral to Treatment (RTT), and the 31-day Decision to Treatment (DTT) (NHS England, 2023; O'Dowd, 2023). The well-known 2-week wait (2WW) standard was removed to align with straight-to-test pathways (NHS England, 2023). Although the reform aims to streamline performance management, it has raised concerns that simplifying targets may hide deeper problems, such as capacity shortages, rather than solving them (O'Dowd, 2023).

The NHS has consistently struggled to meet these targets. The 62-day RTT standard has not been achieved since December 2015 (Kirk-Wade et al., 2024). By May 2024, only 65.8% of patients were treated within this timeframe, far below the 85% target (Kirk-Wade et al., 2024' Limiri, 2025). The national hospital waiting list reached 7.8 million in September 2023, continuing a steady rise since 2021 ((Kirk-Wade et al., 2024; Warner & Zaranko, 2024). These persistent delays, which began before and worsened after the COVID-19 pandemic (Warner & Zaranko, 2024), have serious consequences. Over one in three cancer patients now wait more than 62 days for treatment, increasing their risk of preventable death (Lawler et al., 2024).

Several factors contribute to these ongoing delays:

- **Workforce and Resource Constraints:** Staffing shortages and limited funding remain central challenges (Limiri, 2025). The UK faces a 15% shortfall in clinical oncologists and a projected shortage of 4,000 nurses by 2030 (Aggarwal et al., 2024a). These pressures lead to ethical dilemmas in prioritising care (Dodhia et al., 2023) and rising staff burnout due to years of underinvestment (Aggarwal et al., 2023). The rising costs of medical technology further strain finances (Aggarwal et al., 2024a).
- **Operational Interdependencies:** Delays in one area, such as A&E or social care, create a ripple effect across the system, leading to more extended hospital stays and treatment bottlenecks (Limiri, 2025). Hospital networks are interconnected, meaning backlogs in one trust often affect neighbouring ones (Cima & Almeida, 2024).
- **COVID-19 Pandemic Impact:** The Pandemic heavily disrupted cancer care, reducing compliance with 62-day targets (Fox et al., 2022). Diagnostic referrals for breast cancer fell

by 27% in the first half of 2020 (Gathani et al., 2021), highlighting long-standing weaknesses in capacity and resilience.

These combined pressures demonstrate the fragile state of NHS services. Addressing them requires a coordinated, evidence-based approach to improve system resilience and ensure timely equitable care for all patients.

2.4 Survival Analysis in Healthcare and Public Policy

Survival analysis, also known as time-to-event or event history analysis, is a statistical method used to examine the time until a specific event occurs (Denfeld et al., 2023; Indrayan & Tripathi, 2022). It is particularly valuable in medical research, where data often include censored observations, cases in which the event has not yet occurred or follow-up is incomplete, and event times are typically skewed (Indrayan & Tripathi, 2022).

This method uniquely considers both whether an event occurs (binary outcome) and when it occurs (continuous outcome), making it ideal for analysing event timing and occurrence together (Denfeld et al., 2023). Unlike logistic or linear regression, survival analysis explicitly models time and handles incomplete or non-normally distributed data effectively (Denfeld et al., 2023).

More recently, survival analysis has been applied to assess system-level factors influencing health outcomes. Gibbs et al. (2023) developed a framework to evaluate how waiting times affect health and inequality in the NHS. Their model uses Quality-Adjusted Life Years (QALYs) to measure health impact and accounts for disease progression, patient mortality or withdrawal from waiting lists, and socioeconomic disparities (Gibbs et al., 2023). This framework provides policymakers with evidence to improve outcomes and reduce inequalities, particularly in managing delayed elective procedures following the COVID-19 pandemic.

Monikapreethi S K et al. (Preetha et al., 2024) applied the Cox Proportional Hazards Model to predict individual patient outcomes. Their aim was to assess the model’s predictive performance and measure the effect of different clinical parameters on survival time, supporting more personalised treatment planning (Monikapreethi S K et al., 2024).

The studies by Gibbs et al. (2023) and Monikapreethi S K et al. (Preetha et al., 2024) provide complementary perspectives on healthcare system performance. Gibbs et al. focus on the broader, system-level impact of waiting times on population health and inequality, while Monikapreethi S K et al. use CoxPH to predict patient-level survival outcomes. Together, they connect macro-level policy analysis with micro-level clinical outcomes. This integration helps identify high-risk areas within the system (Gibbs et al., 2023) and the patients most affected within those areas (Monikapreethi S K et al., 2024). Methodologically, Gibbs et al.'s framework could be enhanced by incorporating CoxPH techniques. In contrast, Monikapreethi S K et al. 's models could incorporate system-level factors, such as waiting-time exposure, to better link individual outcomes to overall system performance.

Survival analysis uses several key methods to study time-to-event data. The Kaplan-Meier(K-M) estimator is a non-parametric technique that estimates the unadjusted probability of survival over time and effectively visualises event patterns, even with censored data (Denfeld et al., 2023; Indrayan & Tripathi, 2022). The Log-Rank test compares survival curves between groups to determine whether differences in event timing are statistically significant across the study period (Denfeld et al., 2023; Indrayan & Tripathi, 2022).

The Cox Proportional Hazards (CoxPH) model, however, remains the most versatile and widely used survival method. As a semi-parametric multivariate regression model, it assesses the independent effect of several covariates on the hazard, or risk, of an event (Denfeld et al., 2023; Indrayan & Tripathi, 2022). By producing interpretable hazard ratios, CoxPH identifies which factors most strongly influence survival outcomes. The proportional hazards assumption, that hazard ratios remain constant over time, makes it suitable for both prediction and interpretation. Compared to other methods for describing or comparing survival patterns, CoxPH offers greater analytical power and practical applicability, making it the leading model for extending current research and linking systemic healthcare pressures to individual patient outcomes.

2.5 Synthesis and Research Gap

The reviewed literature highlights several critical gaps that necessitate this research. While AI offers significant potential for healthcare optimisation, there is a recognised lack of real-life implementation and rigorous evaluation of complex AI/ML algorithms in actual healthcare environments for systemic prediction and optimisation (More *et al.*, 2025). Furthermore, persistent challenges in integrating AI with existing, often conservative, healthcare systems and their complex workflows continue to impede seamless adoption (Li *et al.*, 2024; Dicuonzo *et al.*, 2023). Regulatory frameworks for dynamic AI algorithms also remain in nascent stages, posing hurdles for widespread deployment (Li *et al.*, 2024; Pasricha, 2023). From an economic standpoint, despite acknowledging the negative impacts of austerity and inflation, there is a distinct gap in understanding the direct and granular impact of specific healthcare reforms and chronic economic stressors on individual health system components (Guccio *et al.*, 2024; Movsisyan *et al.*, 2024). Finally, while the importance of healthcare resilience is gaining traction, its precise operationalisation and the detailed mechanisms contributing to it, particularly in contexts of impending systemic strain, are not yet fully elucidated (Lyng *et al.*, 2021; Lyng *et al.*, 2022).

The specific gap addressed by this research is the absence of a comprehensive, predictive model capable of quantitatively linking specific, operational pressures, namely hospital activity, staffing levels, and an increase in workload, to systemic strain or failure indicators within individual NHS trusts, particularly within the critical context of cancer care delays. This research aims to develop and validate an early warning indicator system to facilitate proactive intervention.

This research builds upon previous work by:

- **Leveraging advanced analytical methodologies in a novel context:** It extends existing AI/ML capabilities by applying survival analysis (Cox Proportional Hazards model) to model the "time to oncology collapse" at the trust level, a sophisticated approach to predicting systemic failure in a high-priority service.
- **Focusing on a critical, real-world NHS context:** By using cancer treatment delays as a direct proxy for wider NHS dysfunction, the study provides a measurable and impactful

assessment of healthcare resilience within a clinically urgent and publicly significant area, informed by publicly available data.

This research diverges from previous work by:

- **Pioneering a "collapse proxy" for systemic strain:** The novel application of cancer treatment delays as a direct indicator of broader systemic dysfunction or "oncology collapse" at the individual trust level offers a unique methodological lens for understanding healthcare system fragility.
- **Developing proactive, trust-level policy insights:** This study aims to move beyond retrospective audits or broad clinical outcomes by generating specific, actionable recommendations for policymakers at the individual NHS trust level. The goal is to prevent systemic failure and facilitate a critical shift from reactive problem-solving to proactive prevention.
- **Integrating historical precedent for early warning:** The research explicitly incorporates lessons from the Greek oncology crisis, using this stark example of systemic failure under economic strain as a contextual backdrop and a source of early warning signals for the NHS.

In conclusion, this study offers a novel approach to predicting and understanding systemic strain in NHS cancer care. It builds upon existing knowledge of AI's predictive capabilities and the impact of economic factors while directly addressing critical gaps in granular, actionable insights for policy intervention. By doing so, it seeks to significantly enhance the resilience and efficiency of the NHS's vital cancer services, moving towards a more anticipatory and resilient healthcare system.

Chapter 3

Methodology

3.1. Research Philosophy and Approach

This research is grounded in a positivist philosophy, which holds that knowledge is derived from empirical observation and measurement, aiming to quantify phenomena, establish causal relationships, and generalise findings. This approach is inherently aligned with a quantitative design, focusing on numerical data to test a hypothesis and draw objective conclusions (Ghanad, 2023). The choice of a quantitative methodology is driven by the research questions, which aim to quantify the influence of specific variables, such as staffing levels and breaches, on a measurable outcome: the duration until an NHS trust fails to meet its cancer waiting time targets. This aligns with a focus on testing theory and is influenced by empiricist and positivist ideologies, aiming to establish a predictive model through logical, systematic data analysis. The use of large-scale, secondary datasets and advanced statistical modelling, such as survival analysis, is central to this approach, enabling generalisation of findings across a diverse and representative sample of NHS trusts.

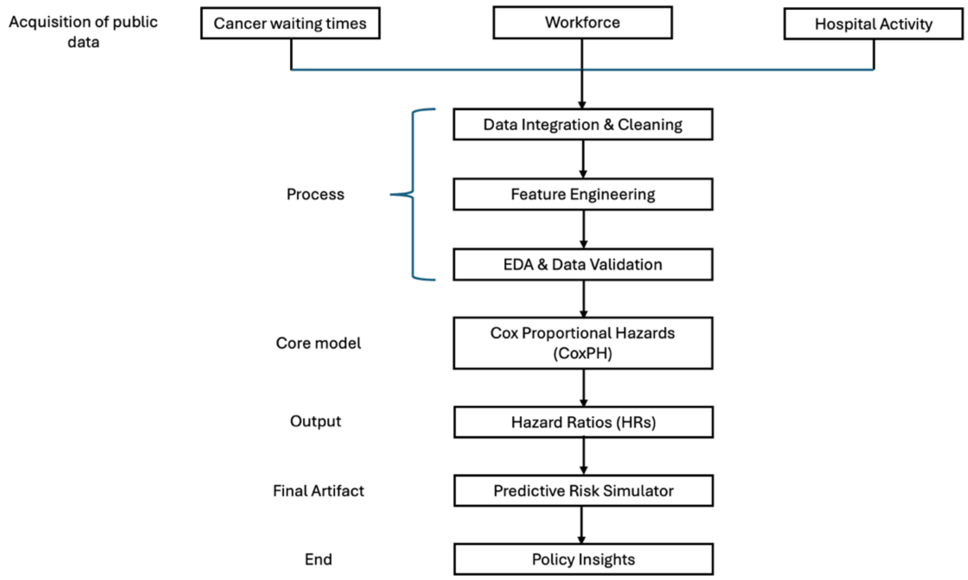


Figure 1: Flowchart of the methodology.

3.2. Data Sources and Acquisition Strategy

Cancer Waiting Times Data: 15 annual datasets from NHS Digital (2009-2023)

Link: https://www.england.nhs.uk/statistics/statistical-work-areas/cancer-waiting-times/historic-cwt-data-before-september-2023/

Hospital Activity Data: Monthly Activity Returns (MAR) from NHS Digital (2009-2020)

Link: https://www.england.nhs.uk/statistics/statistical-work-areas/hospital-activity/monthly-hospital-activity/mar-data/

Workforce Data: NHS workforce statistics (2009-2023)

Link: https://digital.nhs.uk/data-and-information/publications/statistical/nhs-workforce-statistics/september-2024

The foundation of this study's analytical framework is the integration of three comprehensive datasets from NHS Digital, covering the period from 2009 to 2023. These sources were selected for their granularity and direct relevance to the study's focus on systemic strain at the trust level. The datasets utilised include Cancer Waiting Times Data, a series of 15 annual files from NHS England's Cancer Waiting Times Provider-level statistics, formatted as Excel (.xlsx, .xlsb) files, which provide monthly performance data for 221 NHS trusts. This comprises 580,209 individual treatment records.

The Hospital Activity Data consists of Monthly Activity Returns (MAR) from NHS Digital, comprising 89 separate Excel files that capture comprehensive operational metrics from 2009 to 2020. This extensive collection comprises 37,782 trust-month observations across 570 healthcare organisations, detailing variables including elective and emergency admissions, outpatient activity, and GP referral patterns. Lastly, the Workforce Data was sourced from NHS Digital workforce statistics, providing Full-Time Equivalent (FTE) staffing data by trust for the 2009-2023 period.

3.2.1. Data Acquisition and Quality Assurance

To ensure the reproducibility and integrity of the data pipeline, all datasets were acquired systematically using automated download protocols. This approach was implemented to minimise the potential for human error during data acquisition and to maintain a complete audit trail of all source materials used in the analysis.

A key aspect of this process was the strategic alignment of the datasets over time. The temporal alignment strategy prioritised overlapping periods across all data sources to maximise analytical depth while maintaining temporal consistency. This led to the definition of a core analytical period from 2009 to 2020, representing 127 months of integrated healthcare system data. This temporal scope was a conscious decision to establish baseline patterns for understanding normal system dynamics before the confounding period of extraordinary disruptions caused by the COVID-19 pandemic.

3.3. Data Integration and Preprocessing Framework

The process of integrating and preprocessing these heterogeneous datasets was a significant challenge that required an automated framework.

3.3.1. Multi-Year Cancer Data Standardisation

The cancer waiting times data, compiled from 15 annual datasets, demanded extensive standardisation due to inconsistencies in NHS reporting over time. Key challenges included variations in column naming (e.g. 'cancer type' versus 'cancer type/treatment), disparate data formats, and shifts in organisational structures. To address this, a comprehensive standardisation protocol was developed, involving the systematic mapping of equivalent variables across different years and the conversion of mixed date formats (Excel serial numbers and standard datetime strings) using dual-pathway processing. For this analysis, the focus was on the 62-day cancer treatment standards, which are considered the most clinically significant and capacity-sensitive metrics.

3.3.2. Hospital Activity Data Integration

The integration of the hospital activity data presented significant technical challenges due to the sheer scale and heterogeneity of the 89 individual monthly activity returns (MAR) files, which often featured variable header positions and inconsistent formatting. To overcome this, an automated header detection algorithm was developed. This Python-based approach automatically identifies the correct header row across disparate file formats.

The algorithm iterates through the data to find a row containing at least one of a set of known header names, then sets that row as the official header. This represents a good methodological contribution to healthcare data processing, enabling efficient and scalable data integration. A subsequent framework implemented systematic naming conventions, including converting to lowercase and removing special characters, to allow for successful concatenation across all 89 source files. Financial year data (April-March) was then converted to calendar years using a rule-based algorithm to ensure temporal consistency for time-series analysis.

3.3.3. *Workforce Data Processing*

The workforce data required a wide-to-long transformation. This process was necessary to convert annual staffing snapshots into a monthly time-series format suitable for integration with the other datasets, ensuring that staffing levels could be analysed alongside cancer waiting times and hospital activity data.

3.4. Feature Engineering and Variable Construction

The problem with raw numbers is that they don't account for the size of an NHS trust, e.g., a small hospital in a more rural area vs. a larger hospital in London.

The feature engineering approach was guided by healthcare systems theory, aiming to derive indicators that reflect operational dynamics and systemic strain rather than simply relying on raw activity volumes.

3.4.1. *Granular Activity Decomposition*

Hospital activity data was systematically decomposed into operationally distinct components to provide a more nuanced view than aggregate metrics. Rather than treating all hospital admissions as a single measure, they were separated into categories that reflect different operational demands and resource requirements.

For example, elective admissions were divided into:

- Overnight admissions – planned procedures requiring inpatient stays
- Daycase admissions – Scheduled procedures completed within a single day
- Emergency admissions – unplanned acute care requiring immediate attention.

Additionally, referral pathways were tracked by measuring both the number of referrals made by general practitioners and the conversion rate from referrals to actual patient appointments.

This decomposition approach recognised that different types of hospital activity create distinct operational pressures. Emergency surges may force cancellation of elective procedures, while outpatient capacity constraints can create referral backlogs. Since cancer services operate within this broader hospital system, understanding these specific pressures enables more precise identification of conditions that may compromise cancer delivery. By monitoring these disaggregated activity patterns, we can detect early indicators of system strain that may impact cancer services before formal performance targets are breached, providing opportunities for proactive intervention.

3.4.2. Derived System Indicators

Four derived metrics were constructed to capture key system dynamics (Table 1):

1. Care complexity indicator: The 'daycase_to_overnight_ratio' measures the shift towards ambulatory care. A higher ratio indicates efficient care, while a lower ratio may signal capacity constraints.
2. System Pressure Indicators: The 'emergency-to-elective' quantifies acute demand pressure. A high ratio suggests the system is under pressure, which could disrupt elective services such as cancer treatments.
3. System Efficiency Metric: The 'referral_conversion_rate' measures the system's responsiveness to primary care demand. Lower rates indicate capacity constraints affecting cancer referral pathways.
4. Activity Diversity Index: This Shannon entropy-based measure assesses the breadth of a trust's service portfolio. A higher diversity index suggests greater system resilience, while a lower index may point to a more specialised but potentially vulnerable service model.

Variable	Meaning	Why	Reasoning
Daycase_to_overnight_ratio	How many patients go home the same day vs stay overnight	Higher ratio = hospital is efficient Lower ratio = Hospital might be struggling	- Getting sicker patients - Having capacity problems

			- Becoming less efficient - Cancer service might be affected
Emergency_to_elective_ratio	How many emergency patient's vs planned patients	Higher ratio = hospital is overwhelmed with emergencies Lower ratio = hospital has good balance of planned care.	- Hospital cancels planned operations - Cancer treatments delayed - System under stress
Referral_conversion_rate	How many GP referrals get seen	Higher rate = hospital seeing most referred patients quickly Lower rate = Hospital has massive waiting lists	- Outpatient clinics are overwhelmed - Cancer referrals get delayed - Patients wait longer for diagnosis
Activity_diveristy_index	How many different types of services the hospital provides	Higher diversity = Hospital can adapt and reallocate resources Lower diversity = Hospital is specialised but vulnerable	- Less ability to shift resources when needed - Higher risk if their main service gets overwhelmed.

Table 1: Table displaying the engineered variables and their descriptions.

Ratios predict cancer service problems because raw numbers tell you what is happening. Ratios tell you how well the system is working. And when the system isn't working well, cancer patients suffer. These ratios help us predict and prevent that suffering by spotting problems early. e.g.

- If the emergency ratio rises, the hospital gets overwhelmed with urgent cases.
- If the referral conversion drops, cancer referrals start getting delayed

- If the day case ratio falls, Operations take longer, less efficiency
- If diversity drops, there is less flexibility to cope with pressure.

As a result, cancer patients face delays because the whole hospital system is under strain.

3.4.3. Data Type Standardisation

A final preprocessing step involved a systematic numeric conversion pipeline to handle mixed data types in activity columns. This process involved string standardisation, null-value mapping, and regular-expression cleaning to ensure mathematical validity for the modelling phases. This approach was crucial to maximise data retention while preparing the data for analysis.

3.5. Temporal-Spatial Data Integration Strategy

The data integration strategy employed a two-dimensional merge using both organisational identifiers, ‘org code’, and temporal keys, ‘date’, to combine all datasets.

The final integrated dataset covers the period from 2009 to 2020, comprising 127 months of data, or 76% of the total available timeline. This strategic decision to prioritise analytical depth over temporal breadth was deliberate, allowing the inclusion of crucial hospital activity and workforce context while avoiding the extraordinary disruptions of the COVID-19 pandemic. The pre-pandemic data is essential for establishing baseline patterns and understanding normal system dynamics before developing interventions.

3.6. Data Quality Assurance and Validation

Quality control was applied to the final dataset. This involved removing rows with temporal discontinuities, performing a 100% validation check on all variables, and conducting a complete-case analysis to ensure no missing values. The final analytical dataset consists of 45,127 observational units, 23 variables, and is 100% complete for all essential variables.

3.7. EDA on the Final dataset

An exploratory data analysis (EDA) was performed to investigate the fundamental characteristics of the combined dataset. This critical initial phase aimed to validate the study's core assumptions, identify key relationships between variables, and prepare the data for subsequent advanced modelling. This aligns with the principles of quantitative research, which emphasise understanding data through statistical summarisation and visualisation.

The initial EDA began by generating a correlation matrix to assess the quantitative linear relationships among all variables. The matrix revealed complex, nuanced system dynamics, particularly among core performance metrics.

Further exploratory analysis was conducted using box plots on key variables, including ‘total treated’, ‘breaches’, and ‘workforce (FTE)’. These visualisations revealed a significant number of extreme values.

The data for these variables were normalised to a standard scale, a necessary preprocessing step given the wide range of values resulting from different trust sizes. This process ensures the stability and comparability of coefficients in the subsequent survival analysis and machine learning models, as many algorithms are sensitive to the scale of input features.

Finally, time-series analysis was performed by plotting key variables over the integrated period. This step enabled the visualisation of historical trends and shifts in performance, providing empirical evidence for the systemic deterioration of the 62-day RTT target that underpins this research.

The insights from this comprehensive EDA will provide a solid foundation for implementing survival analysis and machine learning models, ensuring that all findings are grounded in a deep, critical understanding of the dataset's characteristics.

3.8. Cox Proportional Hazards Model Development

3.8.1. Defining the Time-to-Event Variable

To construct the survival model, the 'Date' column was utilised as the temporal variable. A reference start date was defined as the earliest date present in the dataset, representing the beginning of the observation period.

For each observation, the time to event was calculated as the number of months from the reference date until either a trust breaches the 62-day target by more than 20% (event = 1) or the last available observation (event = 0, indicating censoring). This approach ensured that all entries were measured on a consistent temporal scale, allowing meaningful comparison across the dataset.

Given that the data was aggregated monthly, using months rather than days provided a more accurate and interpretable representation of the time-to-event process.

3.8.2. Data Pre-processing

Prior to model construction, the dataset underwent several pre-processing steps to ensure quality, consistency, and suitability for survival analysis.

1. Data Cleaning:
Missing values were handled using appropriate estimation methods, and duplicate records were removed. The dataset was reviewed for inconsistencies in variable formats and corrected where necessary.

2. Removal of Perfect Predictors:
The variable 'breaches' was identified as a perfect predictor of the event, resulting in a singular matrix during model fitting. To prevent computational issues and overfitting, it was excluded from the Cox model.

3. Feature Standardisation:

Continuous variables were standardised to improve numerical stability and ensure that each variable contributed proportionately to the model estimation.

4. Categorical Encoding:

Categorical variables were encoded using dummy variables, allowing them to be appropriately incorporated into the regression framework.

3.8.3. Model development

The model was developed using an iterative approach. Initially, all relevant predictors were included. Variables with high Variance Inflation Factors (VIF) or insignificant p-values were gradually removed to achieve an optimal balance between model complexity and explanatory power. Model selection was guided by the Akaike Information Criterion (AIC) and theoretical justification for inclusion of variables.

3.8.4. Model Evaluation and Validation

The model's performance and reliability were assessed through several measures:

- **Concordance Index (C-index):** Evaluated the model's discriminative ability to order event times correctly.
- **AIC and Log-likelihood:** Used to compare model fit and identify the most parsimonious model.
- **Residual Analysis:** Martingale and deviance residuals were examined to identify influential observations and potential outliers.
- **Cross-validation:** The model's stability was assessed using *k*-fold cross-validation, ensuring results were not dependent on a single data partition.

3.8.5. Software

All analyses were conducted using Python (version 3.12.12). The libraries used are widely available and will aid in the preprocessing, cleaning, analysis, visualisation and model building.

Please find the four coding worksheets in this GitHub repository.
https://github.com/Nima-Osman/Dissertation

Chapter 4

Results

4.1 Descriptive statistics

4.1.1 Correlation Matrix

A correlation matrix was the first assessment conducted on the data (figure 2).

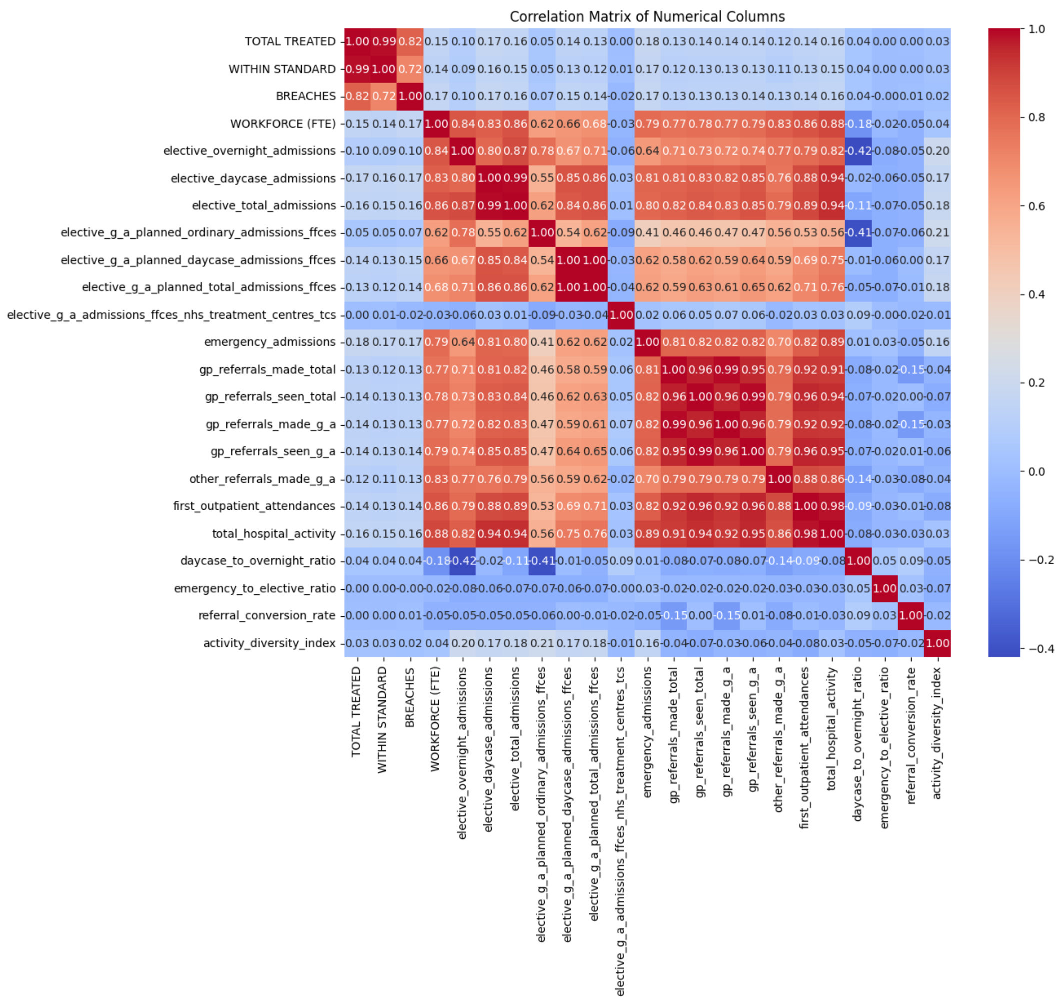


Figure 2: This is the correlation matrix that illustrates the Pearson correlation coefficients between the 24 variables of the dataset.

The variables ‘total treated’, ‘within standard’, and ‘breaches’ exhibit a strong positive correlation with one another, ranging from 0.72 (between ‘within standard’ and ‘breaches’) to 0.99 (between ‘total treated’ and ‘within standard’) (Figure 2). This indicates that, on average, the number of patients treated tends to align closely with the number treated within the standard. However, as overall activity increases, a central tension within the system becomes evident; a proportional increase in breaches of the standard accompanies higher treatment volumes.

The strong correlation between ‘within standard’ and ‘breaches’ suggests that these two measures are not mutually exclusive at the aggregate level. As trusts treat more patients, both the number of successful (‘within standard’) treatments and the number of breaches tend to rise simultaneously. This implies that while productivity increases, maintaining compliance with the standard becomes more challenging.

Additional insights from the correlation matrix:

- A strong positive correlation (0.84) exists between 'workforce (FTE)’ and 'elective_overnight_admissions', indicating that larger 'workforce' sizes are associated with higher overnight admission volumes. A similar pattern is observed with 'daycase_admissions' (0.83). Overall, hospitals with more staff tend to record higher patient admissions. This relationship is further reinforced by the correlation between 'workforce (FTE)’ and 'elective_total_admissions' (0.86). Additionally, 'elective_total_admissions' and 'elective_overnight_admissions' demonstrate a strong positive relationship (0.87).
- ‘gp_referrals_made_total’ also shows a strong positive correlation (0.77) with ‘workforce (FTE)’, suggesting that higher staffing levels are linked to increased referral activity and service demand.

4.1.2 Descriptive Analysis

Across all NHS trusts, the average number of cancer patients treated per month was 7.2, of which 5.9 were treated within the standard, resulting in an average of 1.2 breaches per hospital. The average workforce across trusts was 5,358.67 staff members per month, while the average total hospital activity reached 18,194.95 patients per hospital per month. No missing values were identified in the dataset, ensuring the reliability and completeness of the analysis.

Seasonal Trends

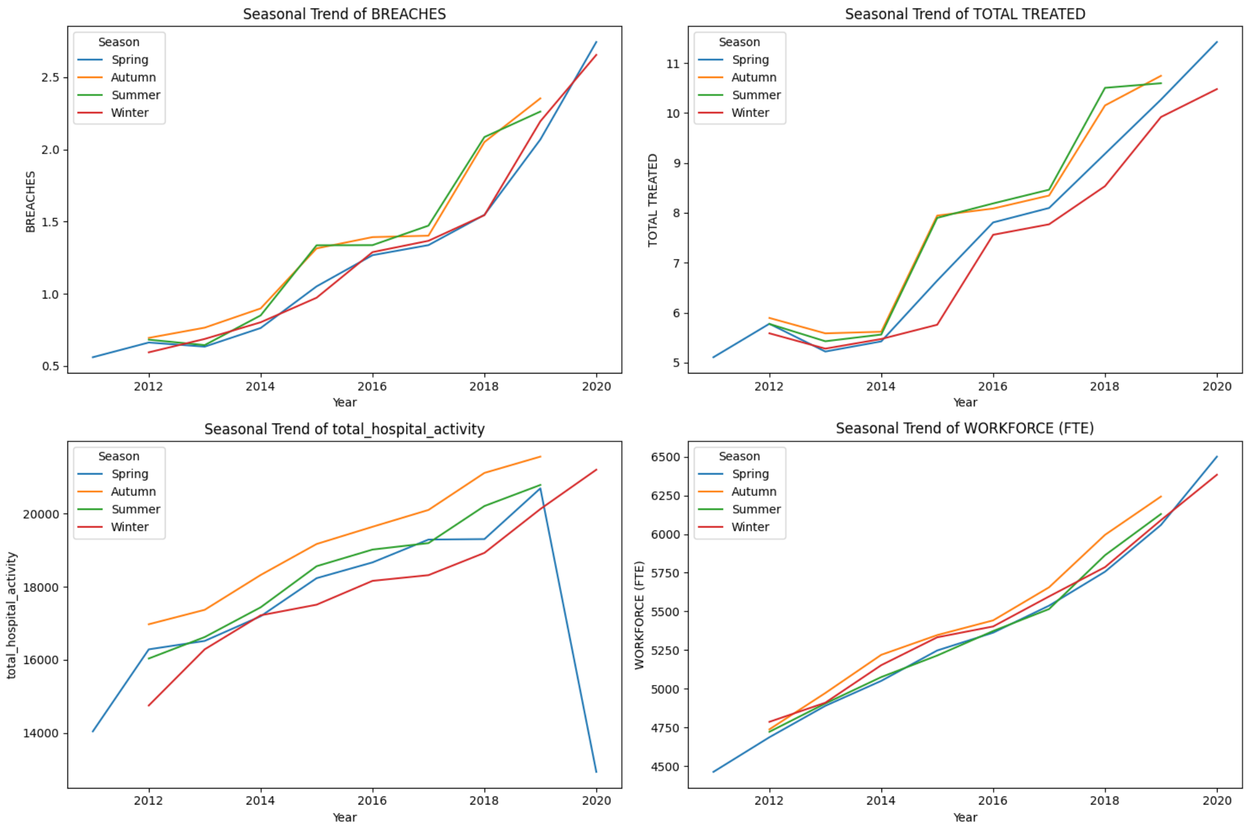


Figure 3: The four graphs illustrate the key variables in the dataset and their seasonal variations.

An examination of seasonal trends for ‘breaches’ across all NHS trusts indicates a steady increase in breaches over the years, with autumn and summer recording higher breach levels than spring and winter. A similar pattern is observed in the total number of patients treated, with winter consistently showing the lowest number of cancer treatments, while summer and autumn record the highest figures (Figure 3).

The trend in total hospital activity was unexpected. It was anticipated that the winter months, typically associated with higher hospital admissions due to flu and seasonal respiratory illnesses, would show the highest activity. However, the data reveal the lowest activity levels during winter, with autumn showing the highest hospital activity, followed by summer (Figure 3).

This pattern can be attributed to the increased pressure hospitals face during winter, primarily due to surges in emergency admissions linked to seasonal illnesses. The heightened demand for acute care can strain finite resources, leading to the cancellation or postponement of non-urgent elective procedures, including cancer treatments. Consequently, while emergency activity may rise, the overall hospital activity (which combines elective and emergency care) may decline due to deferred planned procedures.

Furthermore, staff sickness during winter months may exacerbate workforce shortages, directly impacting hospitals' capacity to carry out treatments and manage patient flow efficiently. Factors such as stress, resilience, and coping mechanisms among healthcare workers may further influence system performance and patient outcomes during this period.

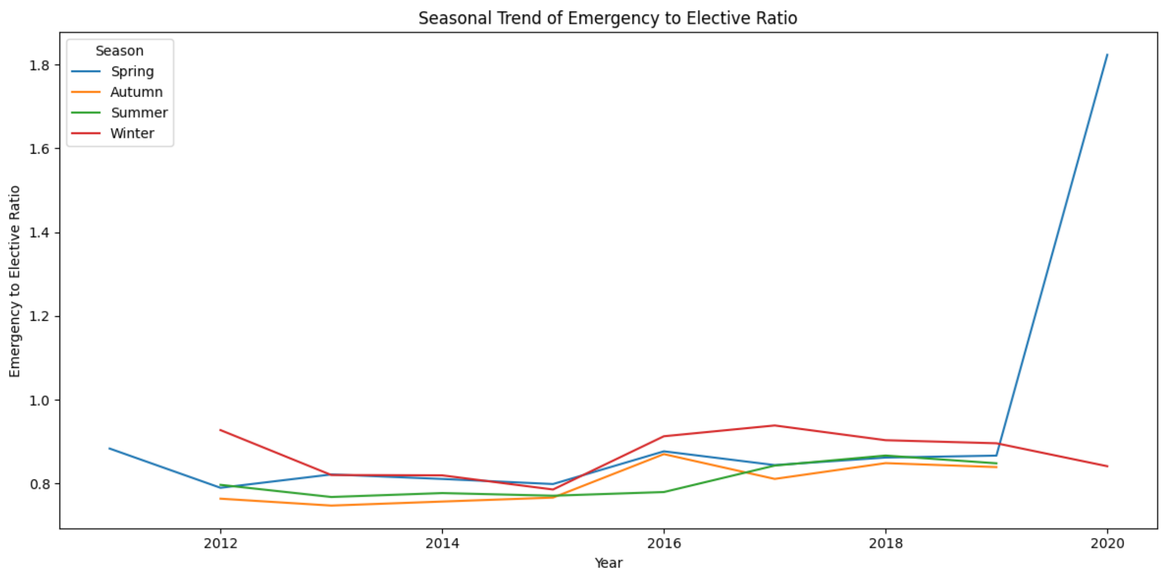


Figure 4: This line graph shows how the Emergency to Elective Ratio *fluctuates across different seasons.*

A closer examination of the Emergency to Elective Ratio (Figure 4), which compares the number of emergency patients to the number of planned (elective) patients, reveals that a higher ratio indicates greater emergency pressure. This provides a clearer understanding of hospital dynamics, particularly during the winter months.

As illustrated in the graph below, the ratio peaks during winter, suggesting that hospitals experience a higher influx of emergency cases during this period. This increased emergency demand places additional strain on hospital capacity and resources. Consequently, the overall hospital activity observed in the previous figure appears lower, likely due to the postponement or reduction of elective procedures as hospitals prioritise urgent care.

4.1.3 Breach patterns

While the available data does not directly explain the high number of breaches, it is possible to explore other variables in the dataset to identify potential contributing factors.

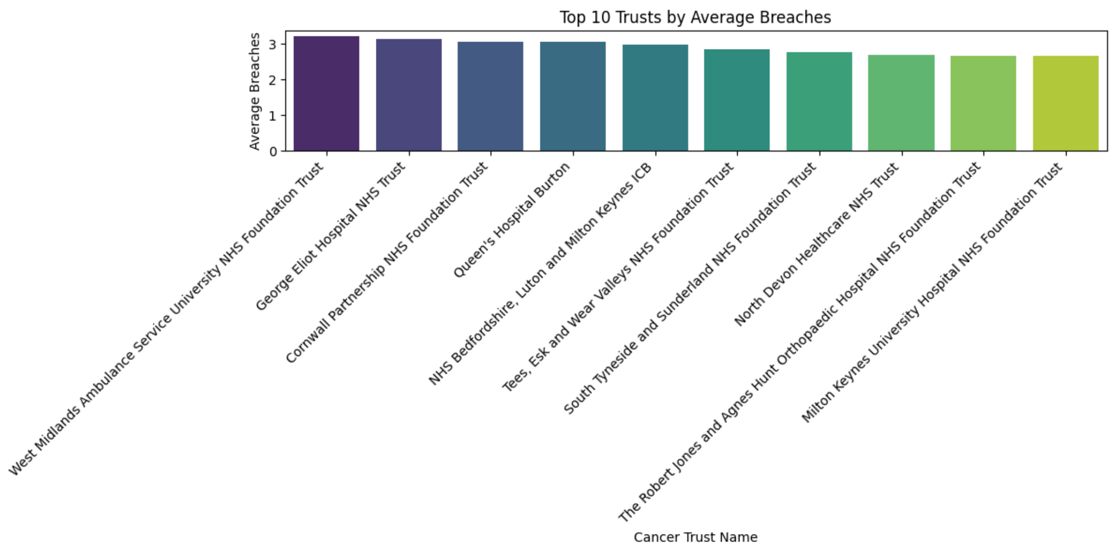


Figure 5: This illustrates the top 10 trusts with the highest number of average breaches of the timespan of 2012-2020.

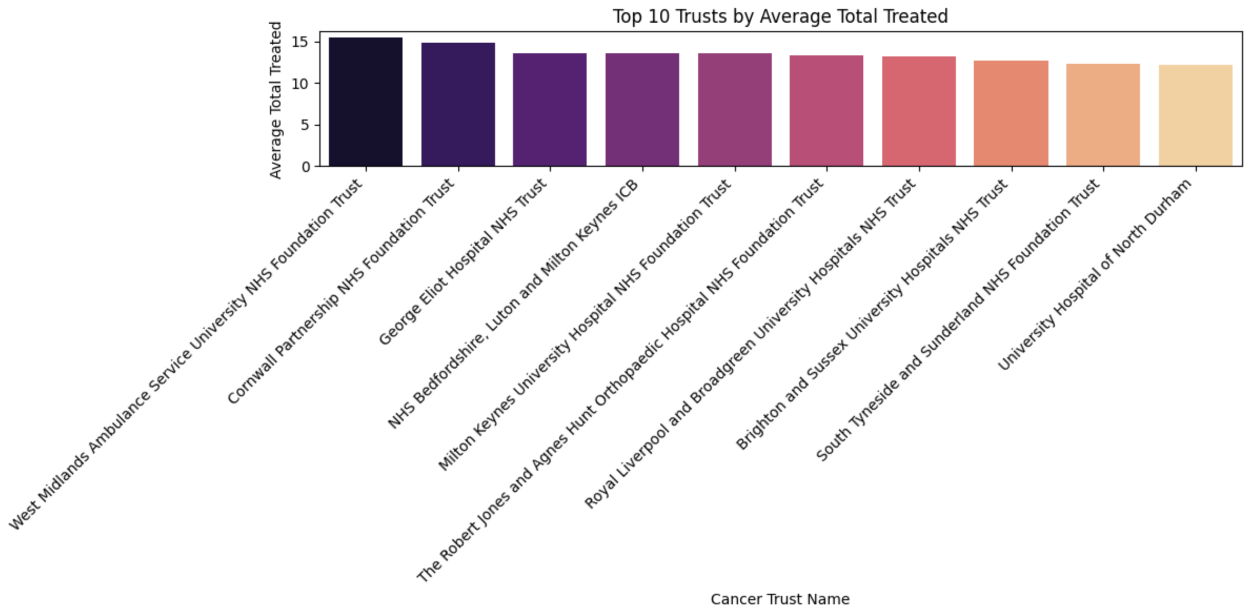


Figure 6: This illustrates the top 10 trusts with the highest average total treated over the timespan of 2012-2020.

While the available data does not directly explain the high number of breaches, it is possible to explore other variables in the dataset to identify potential contributing factors. Areas for further investigation include whether trusts with higher breaches also exhibit:

- Lower workforce (FTE) levels
- Higher total hospital activity
- Lower referral conversion rates
- Lower activity diversity index values

A notable finding is this significant overlap between the top ten trusts for average breaches (Figure 5) and those for average total treated (Figure 6). Trusts such as the following listed appear in both categories:

- West Midlands Ambulance Service University NHS Foundation Trust
- George Eliot Hospital NHS Trust
- Cornwall Partnership NHS Foundation Trust

- NHS Bedfordshire, Luton and Milton Keynes ICB
- Milton Keynes University Hospital NHS Foundation Trust, The Robert Jones and Agnes Hunt Orthopaedic Hospital NHS Foundation Trust
- and South Tyneside and Sunderland NHS Foundation

The overlap suggests that trusts treating higher patient volumes also tend to experience more breaches, potentially due to increased pressure on available resources, staff capacity, or infrastructure.
However, some trusts feature in only one list. For example, Queen's Hospital Burton and North Devon Healthcare NHS Trust rank among the top 10 for breaches but not for total treated, whereas Royal Liverpool and Broadgreen University Hospitals NHS Trusts, Brighton and Sussex University Hospitals NHS Trusts, and University Hospital of North Durham rank highly for total treated but not for breaches. This distinction indicates that High patient volume alone is not the sole determinant of breaches, and other underlying factors, such as operational efficiency, case complexity, or local workforce composition, are likely to contribute.

To gain deeper insights, it is essential to examine trends over time and consider metrics such as breach rate per total treated. It should also be acknowledged that the dataset may not capture all relevant factors influencing breach levels, such as specific resource limitations, staffing for key specialities, or changes in clinical pathways.

4.1.4 Detecting outliers

The detection and interpretation of these outliers are paramount because they represent the extreme, real-world operational challenges of the NHS that the predictive model aims to address. Failure to account for the wide variation in trust sizes and activity levels could introduce numerical instability into the modelling algorithms and skew the results.

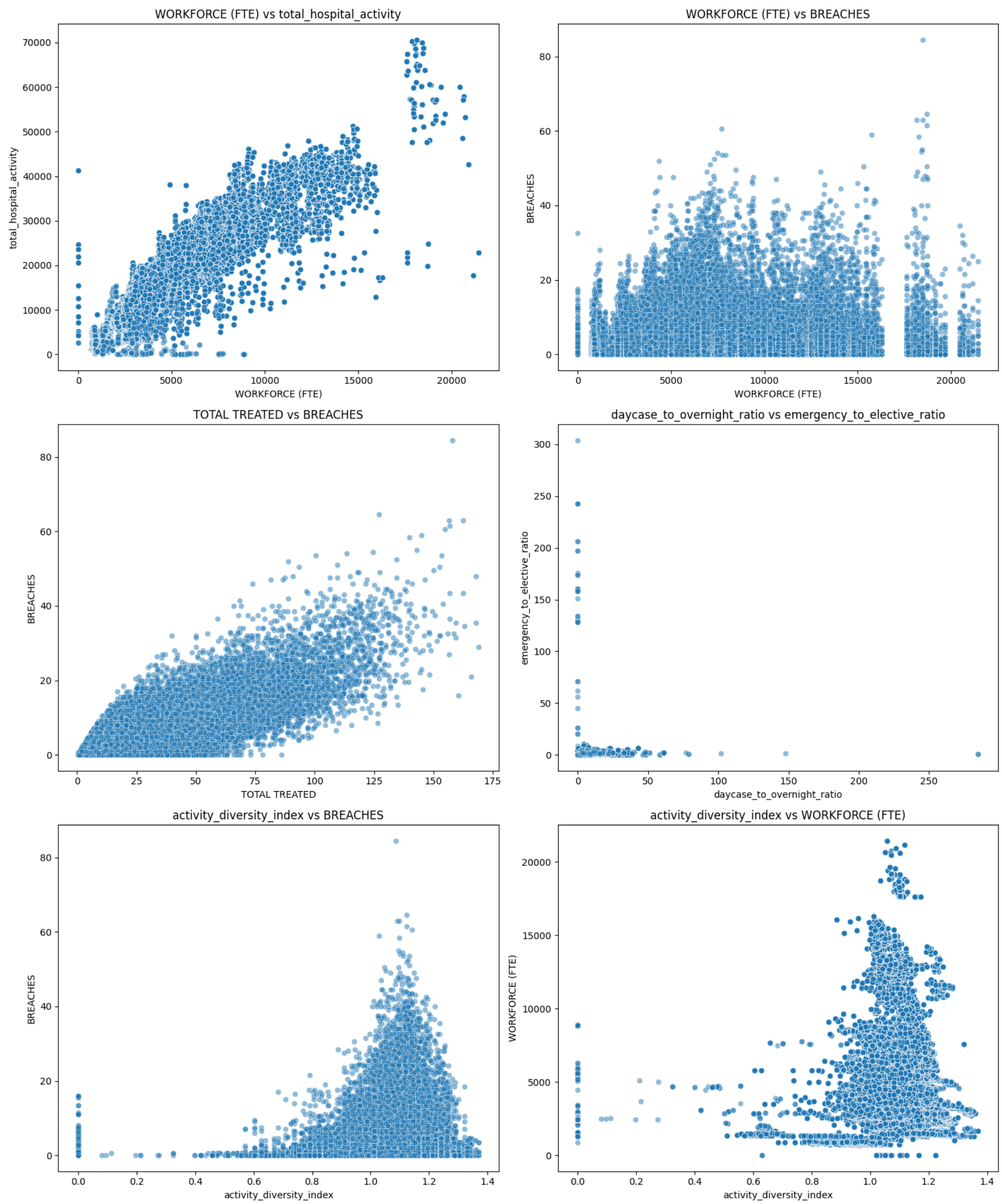


Figure 7: Scatterplots illustrating the correlations between selected variables.

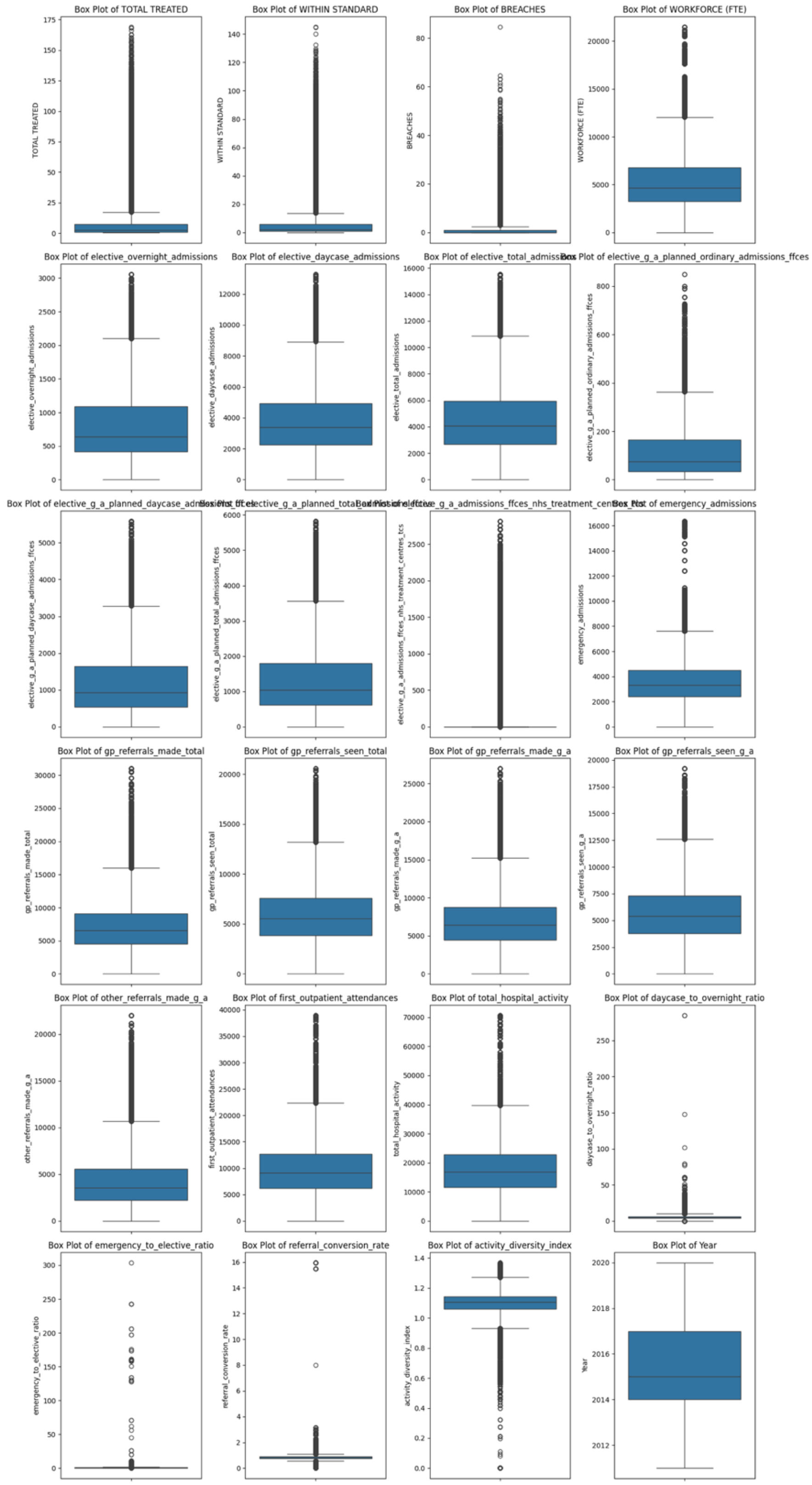


Figure 8: Boxplots displaying the distribution of key variables in the dataset, used to identify potential outliers.

Figure 7 displays scatterplots of the key variables. These scatterplots suggest that an increased workforce also leads to increased hospital activity. The more patients are treated, the greater the risk of breaches. Furthermore, the higher the activity diversity index, the more breaches occur. The activity diversity index is a measure of the number of service types the hospital offers, suggesting that the number of services available can affect whether the hospital experiences a breach.

The scatterplots themselves have many outliers, which can be further investigated using boxplots.

Outliers were identified using boxplots to visualise the distribution of key variables within the dataset. A considerable number of extreme values and outliers were observed, which may be attributed to several underlying factors:

1. <u>Trust size:</u> NHS trusts vary significantly in scale, ranging from small local hospitals to large multi-site teaching hospitals
2. <u>Specialised services:</u> Some trusts serve as regional or national centres for complex cancer treatments, leading to higher patient volumes or case complexity.
3. <u>Referrals volume spikes:</u> Periodic surges in referrals may temporarily increase total hospital activity and the number of breaches.
4. <u>Workforce fluctuations:</u> Changes in staffing levels can directly affect service delivery and capacity.
5. <u>Data reporting anomalies:</u> Inconsistencies or delays in data reporting may introduce artificial variation.
6. <u>Systemic pressure:</u> Periods of intense operational strain can result in simultaneous increases in breaches and the total number of cases treated.

The most plausible explanation is that these outliers reflect the heterogeneous nature of the NHS, where trusts differ substantially in size, capacity, and service specialisation. The variation in trust sizes from small local hospitals to large multi-site teaching hospitals naturally results in a wide range of activity levels and workforce numbers. These extreme values, therefore, represent the high-volume, high-pressure contexts central to the study's investigation of systemic strain. Their presence confirms the need for robust analytical models that are not overly skewed by these contextual differences.

Although the Cox Proportional Hazards (COXPH) model estimates relative risk through hazard ratios, which focus on the effect of a one-unit change in a variable rather than its absolute magnitude, data normalisation remains advisable in this context. The wide variation in trust sizes and activity levels can introduce numerical instability in optimisation algorithms used to fit the model, particularly when combined with other machine learning models.

Normalising the data to a standard scale (e.g., using min-max scaling or z-score standardisation) helps to mitigate these issues. It also facilitates more meaningful comparisons between coefficients, such as comparing the impact of a one-standard-deviation increase in 'workforce (FTE)' versus 'referral_conversion_rate'. This is particularly important for generating policy rankings, as it enables a consistent assessment of how different operational factors influence systemic pressure.

Furthermore, although the COXPH model itself is not sensitive to the magnitudes of individual variables, the size of a trust may influence how its hazard rate evolves over time. Normalisation, therefore, provides an additional layer of comparability, ensuring that observed differences reflect actual effects rather than disparities in scale.

4.1.5 Time Series Analysis

The following time-series graphs show the relationships between breaches and several key variables over time.

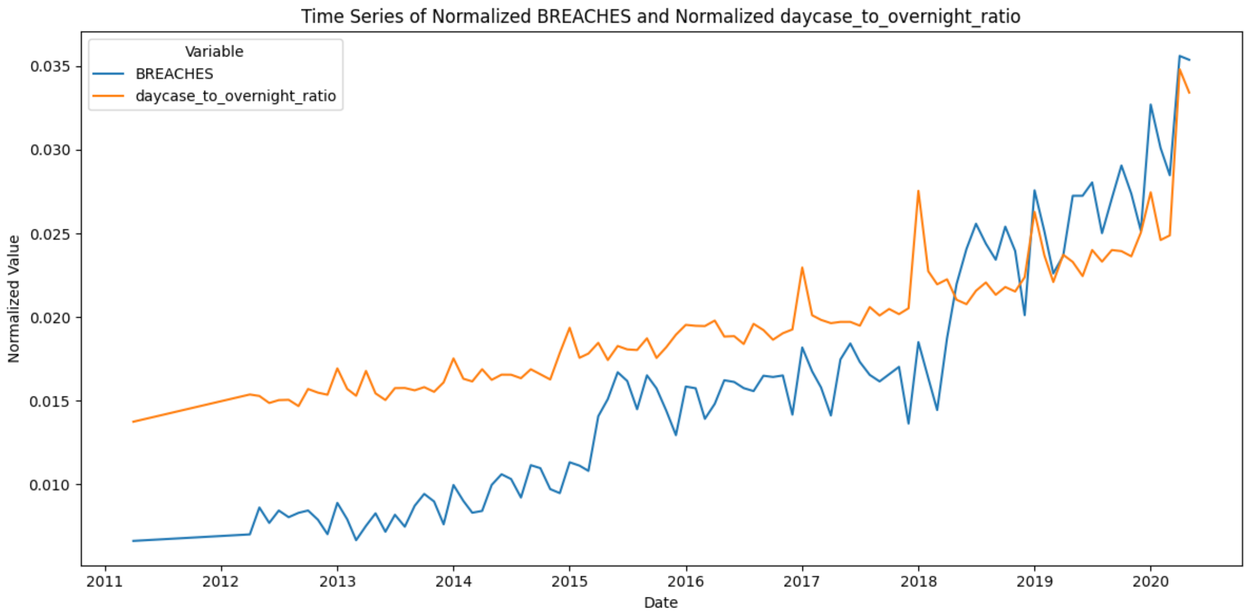


Figure 8: Time series graph comparing changes in the number of breaches over time with the Daycase to Overnight Ratio.

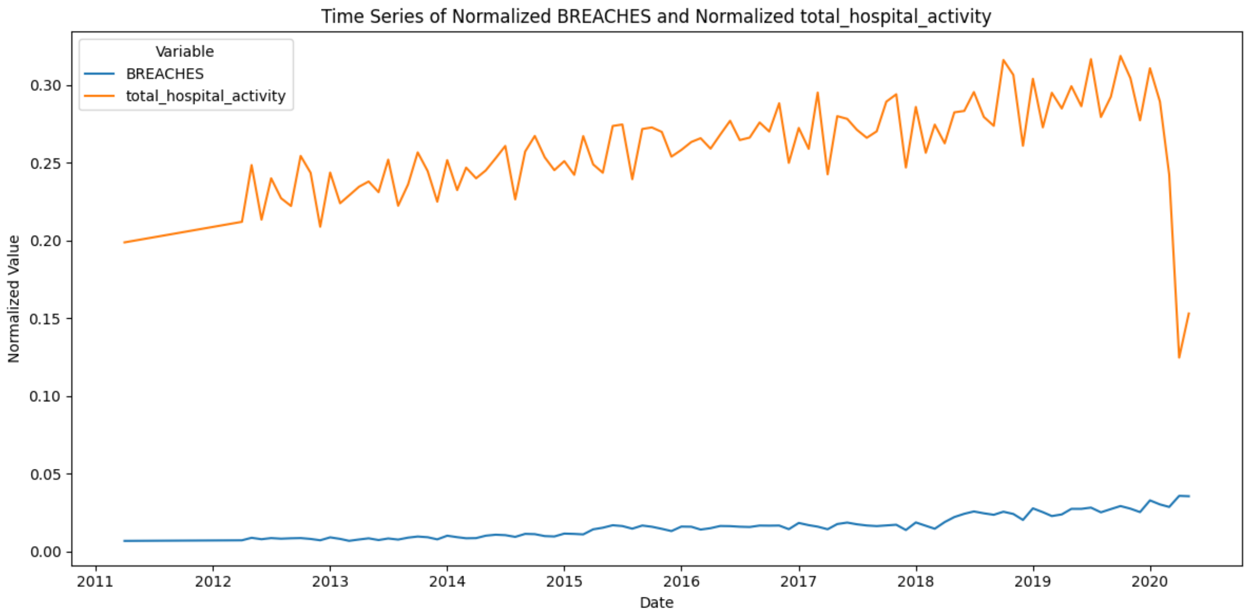


Figure 9: Time series graph comparing changes in the number of breaches over time with the Total hospital activity.

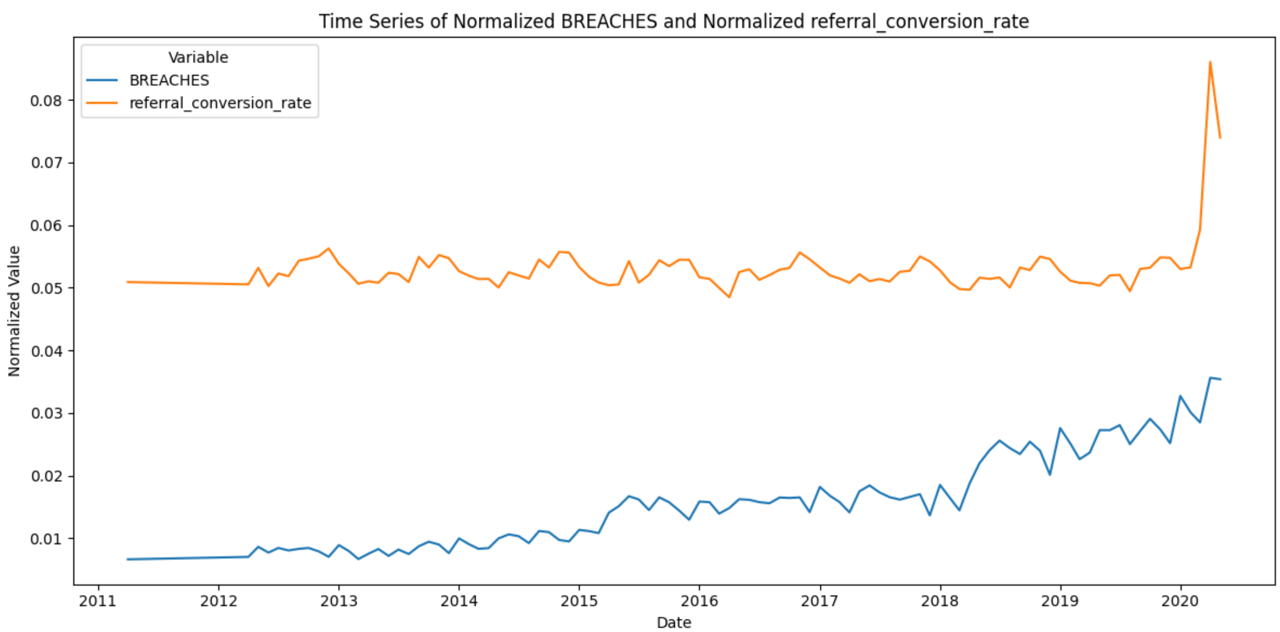


Figure 10: Time series graph comparing changes in the number of breaches over time with the Referral conversion rate.

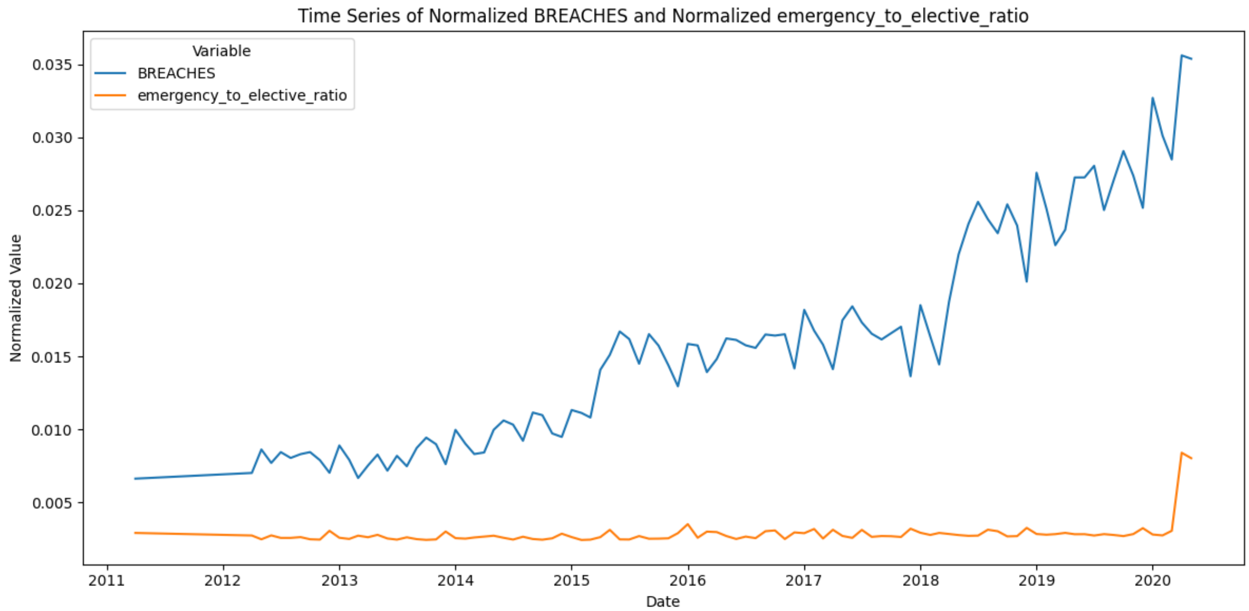


Figure 11: Time series graph comparing changes in the number of breaches over time with the emergency to elective ratio.

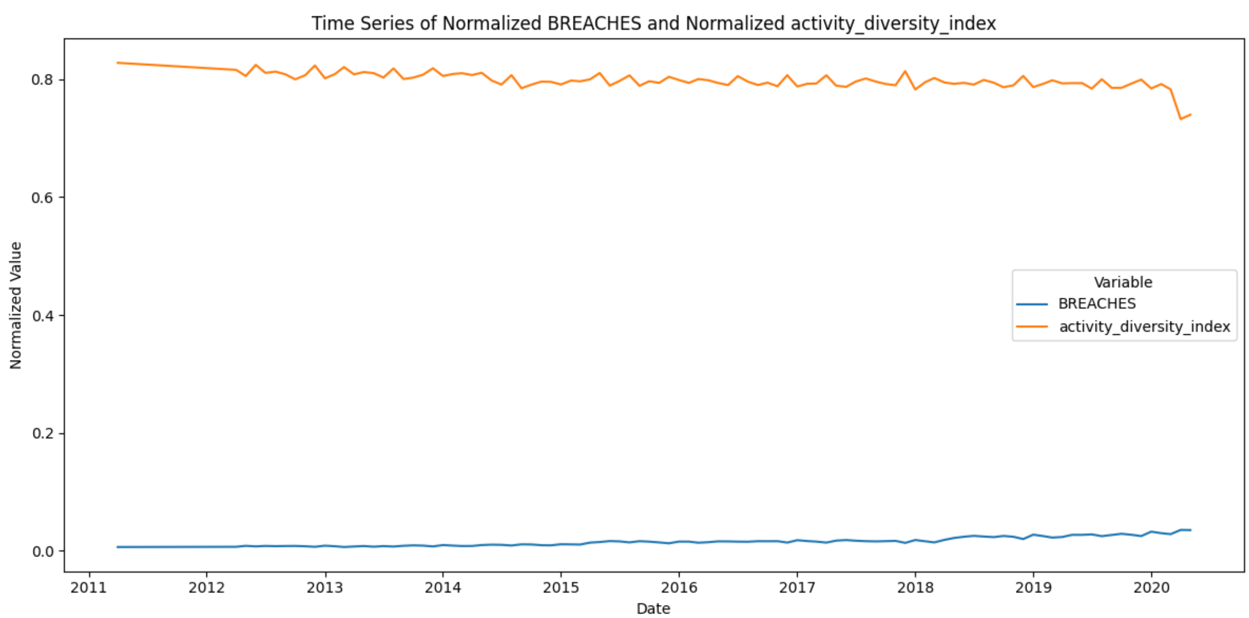


Figure 12: Time series graph comparing changes in the number of breaches over time with the activity diversity index.

The following time-series graphs show the relationships between breaches and several key variables. The main observations are summarised below:

Breaches vs Day Case to Overnight Ratio (Figure 8):

Between 2011 and 2018, both variables increased at a relatively steady rate. However, after 2018, the number of breaches rose sharply while the Day Case to Overnight Ratio continued to increase at a slower pace. This suggests that the ratio may have influenced the number of violations before 2018, but additional factors likely contributed to the rapid escalation observed thereafter.

Breaches vs Total Hospital Activity (Figure 9):

Both breaches and total hospital activity show a steady upward trend over time. This indicates a potential relationship in which a proportional rise in breaches accompanies an increase in overall hospital activity.

Breaches vs Referral Conversion Rate (Figure 10):

The Referral Conversion Rat*e* fluctuates around 0.05, showing no clear upward or downward trend, while breaches steadily increase, particularly after 2018. This pattern suggests that the referral conversion rate does not significantly affect the number of breaches.

Breaches vs Emergency to Elective Ratio (Figure 11):

The Emergency to Elective Ratio remains stable, primarily around 0.00025, with slight variation over time, while the number of breaches continues to rise. This implies that changes in the emergency-to-elective balance may not directly influence breach levels.

Breaches vs Activity Diversity Index (Figure 12):

The Activity Diversity Index shows a gradual decline over the observed period, while breaches continue to increase. This inverse relationship may suggest that trusts offering a narrower range of services are more vulnerable to breaches, potentially due to reduced adaptability and limited capacity to redistribute resources under strain.

Other variables were also examined through time series analysis; however, these are presented in the Appendices (1-7)**.** The variables highlighted above were selected for inclusion in the model-building phase**,** as they demonstrated the most relevant or interpretable relationships with breach trends.

4.1.6 Admitted vs non-admitted

Assessing the dataset, the results are also split into admitted and non-admitted.

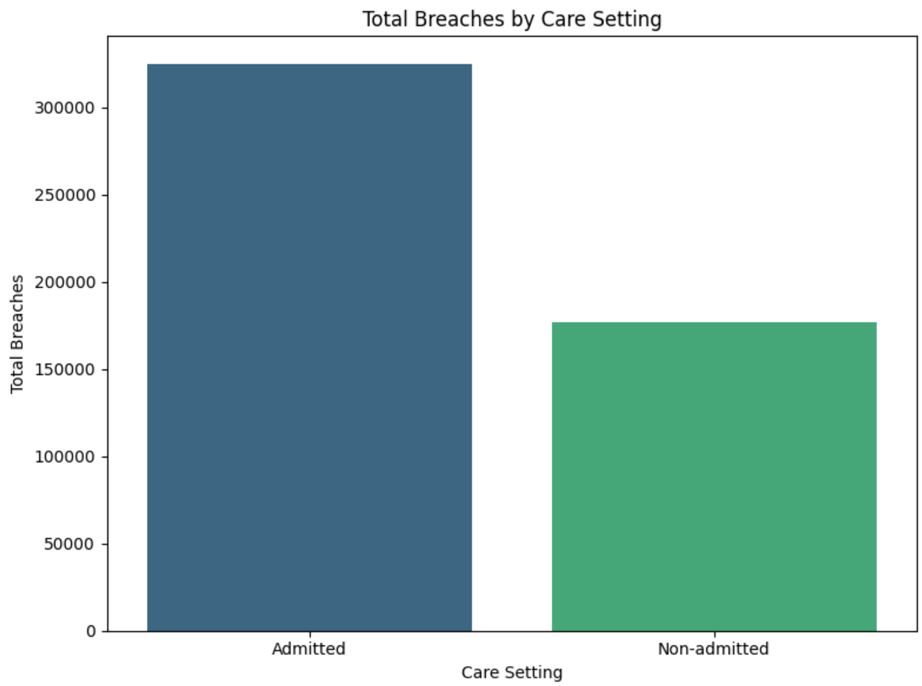


Figure 13: Bar chart illustrating the total number of breaches by care setting.

Figure 13 clearly illustrates this distinction, showing that admitted cases experience approximately twice as many breaches as non-admitted cases. When examined over time **(**Figure 14**),** both care settings display a similar upward trend, with the number of breaches steadily increasing across the observed years. However, there is a notable spike between 2014 and 2015, during which breaches rose sharply, especially among admitted cases, nearly doubling in magnitude compared to non-admitted cases. This sudden increase suggests a significant system-wide disruption or capacity constraint that disproportionately affected admitted patients during this period.

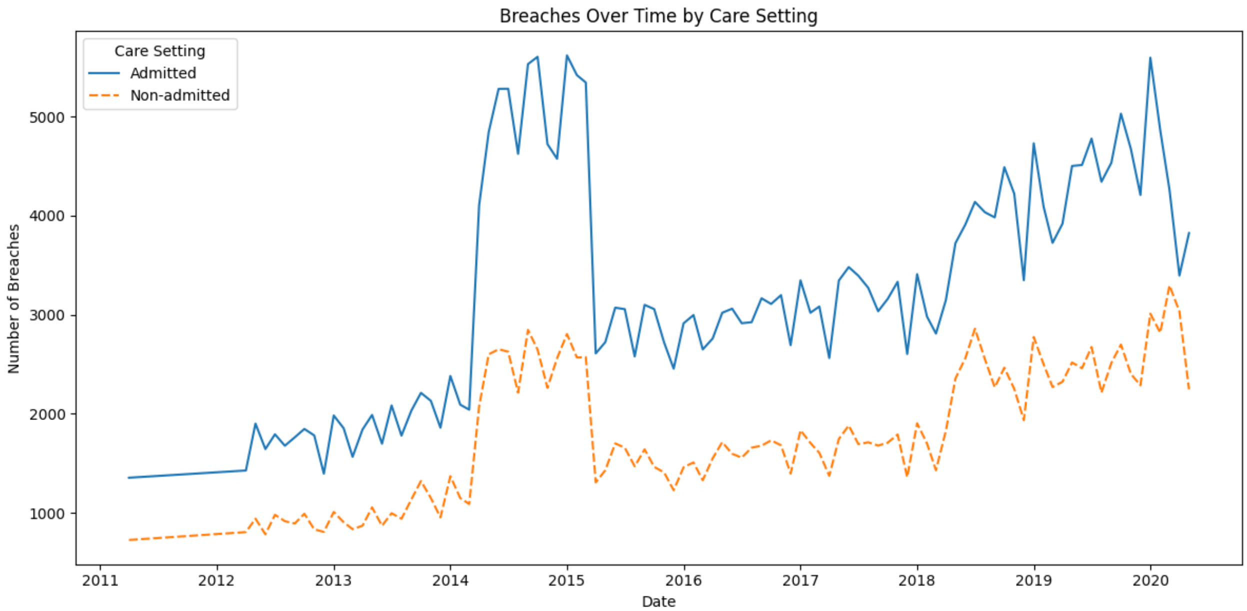


Figure 14: Time series graph comparing the trends in admitted and non-admitted cases over time.

To determine whether there is a meaningful difference between admitted and non-admitted cases, the breach rate was calculated (Figure 13). This approach normalises the data, accounting for differences in the volumes of admitted and non-admitted patients.

Results:

- **Admitted Breach Rate:**

$$\text{Total Admitted Breaches}/\text{Admitted Count} = 324{,}603/234{,}682 = 1.38$$

- **Non-Admitted Breach Rate:**

$$\text{Total Non-Admitted Breaches}/\text{Non-Admitted Count} = 177{,}060.5/175{,}696 = 1.01$$

These results reveal a clear and significant difference between the two categories. The admitted breach rate is approximately 38% higher than the non-admitted breach rate. In practical terms, for every 100 admitted patients, there were around 138 breaches, compared with 101 breaches per 100 non-admitted patients. This indicates that patients who are admitted to the hospital are considerably more likely to experience a breach.

Admitted breaches refer to cases where patients have been admitted for treatment, such as surgery, chemotherapy, or radiotherapy. A breach occurs if the time from the decision to treat to the start of the treatment (the 31-day DDT standard) is not met. These patients are already within the hospital system and typically require inpatient care.

Non-admitted breaches, by contrast, relate to patients who have been referred with a suspected cancer diagnosis but have not yet been admitted for treatment. A breach in this case represents a failure to meet the 62-day referral to treatment (RTT) standard, which measures the period from initial GP referral to the start of treatment. This phase broadly covers diagnostics and outpatient activity, including imaging, consultations, and testing.

The observed difference is clinically and operationally significant, suggesting that system strain becomes particularly acute during the treatment phase, after diagnosis has already been made. Potential contributing factors include:

- Limited surgical or procedural capacity
- Bed shortages are impacting inpatient admissions
- Scheduling delays for specific therapies (e.g., radiotherapy or chemotherapy)

Overall, the findings indicate that while the diagnostic phase appears to be managed relatively effectively, the delivery of treatment represents a key pressure point within the cancer care pathway.

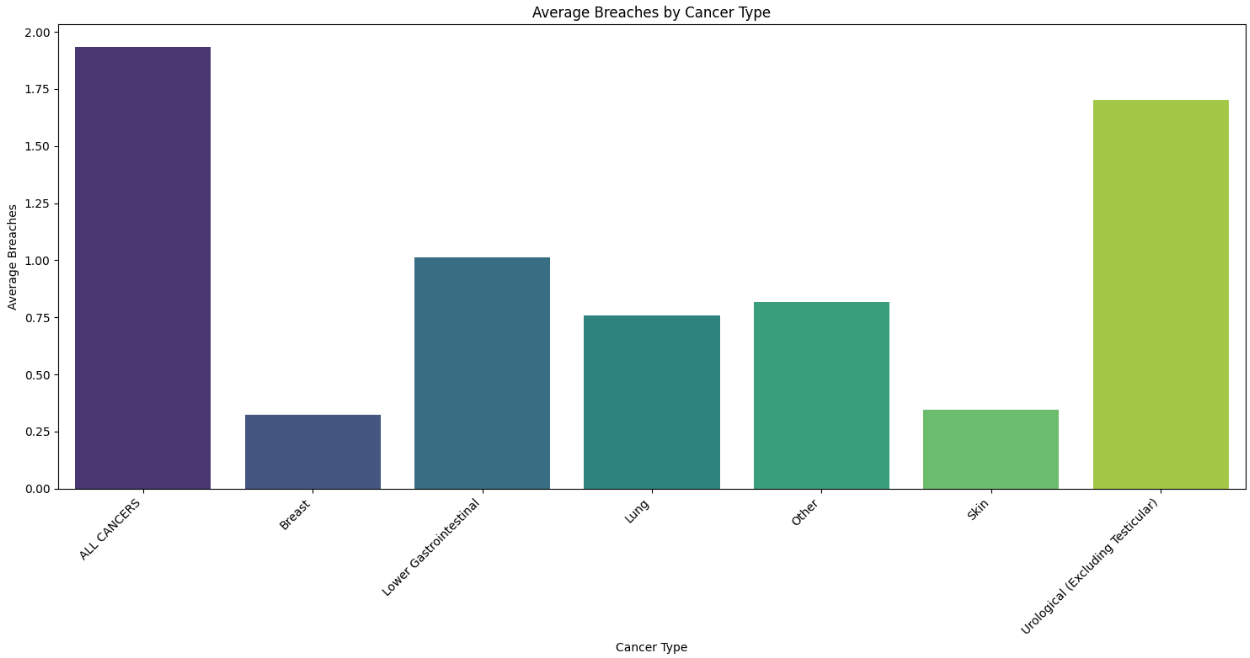


Figure 15: The graph compares the breaches by cancer type.

Urology demonstrates the highest number of breaches, whereas breast cancer records the lowest. Several factors may influence this disparity. The extensive public awareness campaigns and charity initiatives surrounding breast cancer have historically attracted substantial media attention and policy support, resulting in greater investment in diagnostic infrastructure, screening programmes, and treatment pathways. Early detection through established national screening programmes, such as routine mammography, may also contribute to more timely diagnoses and interventions, thereby reducing breast cancer rates. In contrast, urology cancers may not benefit from comparable levels of public awareness or early screening, often leading to later-stage presentations and more complex treatment pathways, both of which can increase the likelihood of breaches.

4.2 Definition of survival variables

The final phase of this analysis involved developing a predictive model to quantify the risk of service failure across NHS trusts. Given that the outcome variables are time-to-event (the duration until a trust breaches the 62-day cancer standard), the appropriate statistical framework is survival analysis. Specifically, the Cox Proportional Hazards (CoxPH) model was selected, as it assesses the effect of covariates on the hazard rate, allowing for the calculation of an interpretable relative risk (Hazard Ratio).

Before model specification, the dataset was defined according to the survival analysis convention:

1. Time-to-event (T): The number of months elapsed from the start of the observation period (April 2011) until the event occurred.
2. Event Indicator (E): A binary variable where 1 denotes the event (breach of the 62-day standard) and 0 denotes censoring (The trust had not yet breached by the end of the study period or was lost to follow-up).
3. Covariates (X): The 5 selected features used to predict the hazard rate.

The data revealed a counterintuitive but significant finding: hospital activity did not spike in winter months as might be expected with seasonal illnesses. This finding is interpreted as a manifestation of the system's capacity constraints. Acute demand from emergencies forces the cancellation of elective procedures, leading to a net decrease in overall activity, a direct indicator of system strain.

4.3 Model Specification and Feature Selection

The model was specified using the lifelines library in Python, estimating the baseline hazard using the Breslow method.

The final model specification was determined through an iterative feature selection process aimed at maximising the predictive power while maintaining the critical principles of statistical independence and parsimony.

Feature selection rationale:

Two core methodological constraints guided the process:

1. Collinearity resolutions: Raw variables used to construct engineered ratios were systematically removed to eliminate multicollinearity, ensuring that the final model's coefficients genuinely represent the effect of systemic strain captured by the ratios, rather than simply volume effects. High correlation among predictor variables can distort the interpretation of regression coefficients. A correlation matrix was examined to identify multicollinearity. Strongly correlated variables were carefully reviewed, and redundant predictors were removed where necessary.
   In addition, Principal Component Analysis (PCA) was considered to reduce dimensionality. PCA combines correlated variables into a smaller set of uncorrelated components, preserving most of the dataset's variance while mitigating multicollinearity. This step ensured that the final model remained both interpretable and statistically robust.

2. Model Improvement: A comprehensive search was performed to identify the combination of features that yielded the most significant increase in predictive accuracy (Concordance index)

This process culminated in the selection of the final 5-feature model, which included four core engineered ratios and the squared term of the workforce variables (workforce_squared), included to test for potential non-linear effects of capacity (Figure 16).

```
TOP 15 FEATURES (with prediction diversity ≥5.0):
==========================================================================================
#   Feature                                     Conc     Imp      Range   Std     Corr
------------------------------------------------------------------------------------------
✓ 1  workforce_squared                          0.661   +2.1pp    42.0    8.8  +0.156
✓ 2  WORKFORCE (FTE)_squared                    0.661   +2.1pp    42.0    8.8  +0.156
✓ 3  WORKFORCE (FTE)                            0.660   +2.0pp    42.0    8.6  +0.167
✓ 4  workforce_zscore                           0.660   +2.0pp    42.0    8.6  +0.167
✓ 5  emergency_x_workforce                      0.656   +1.6pp    42.0    8.7  +0.167
+ 6  emergency_admissions_squared               0.649   +0.9pp    42.0    8.7  +0.157
+ 7  emergency_squared                          0.649   +0.9pp    42.0    8.7  +0.157
+ 8  TOTAL TREATED                              0.648   +0.8pp    51.0   10.8  +0.817
+ 9  emergency_x_elective                       0.648   +0.8pp    43.0    8.8  +0.172
+ 10 other_referrals_made_g_a                   0.647   +0.7pp    43.0    8.8  +0.132
+ 11 elective_g_a_planned_ordinary_admissio     0.646   +0.6pp    43.0    8.7  +0.066
+ 12 WITHIN STANDARD                            0.646   +0.6pp    52.0   10.9  +0.715
+ 13 emergency_zscore                           0.644   +0.4pp    43.0    8.8  +0.170
+ 14 emergency_admissions                       0.644   +0.4pp    43.0    8.8  +0.170
+ 15 WORKFORCE (FTE)_log                        0.643   +0.3pp    44.0    8.8  +0.126

==========================================================================================
RECOMMENDED 5TH FEATURE:
==========================================================================================
Feature: workforce_squared
Concordance: 0.661 (66.1%)
Improvement: +2.1 percentage points
Prediction diversity: 8.8 (good - meets 5.0 threshold)
Prediction range: 42.0 months
Correlation with BREACHES: +0.1556
```

Figure 16: Screenshot of the output of code assessing the final five features of the CoxPH model.

4.4 Model Fit and Hazard Estimation

The final CoxPH model was fitted to the 410,378 observations.

The concordance index (c-index) measures the model’s discrimination, i.e., its ability to rank subjects by their estimated risk or predicted time-to-event correctly. A value of 1 indicates a perfect ranking, while 0 reflects a perfectly incorrect ranking. Importantly, the c-index evaluates only the *order* of predictions, not the size of the errors.

In Figure 16, the *improvement (imp)* column shows the gain in c-index when each feature is added relative to a baseline model. The *standard deviation (std)* ranges from 8.6 to 10.9, indicating that the features contribute enough variability to be informative. The 42-month prediction range represents the difference between the minimum and maximum predicted survival times. All variables show positive correlations, suggesting that as the feature's value increases, the instantaneous risk of the event (a breach) increases.

Overall, the listed features are reasonable predictors based on their concordance indices. The model assigns a risk score to each trust, and its ability to correctly rank time-to-breach yields a c-index of 0.661, meaning it correctly orders the risks 66.1% of the time. To assess the predictive power of each variable individually, the CoxPH model in Figure 16 is fitted once per feature. The resulting single-variable c-index allows the variables to be ranked: those with higher c-index values are stronger predictors. This approach helps separate genuine predictive signals from patterns driven by chance or collinearity. Variables with identical c-index scores are likely the same measure or highly collinear, in which case only one needs to be retained.

For example, workforce (FTE) has a higher c-index (0.66) than emergency admissions (0.64), indicating it is the stronger single predictor.

Finally, the model's coefficients are used to calculate relative risk. The Hazard Ratio (HR = exp(coef)) represents the proportional increase or decrease in the instantaneous risk of breaching the 62-day target for a one-unit change in the covariate.

FINAL COX PROPORTIONAL HAZARDS MODEL STATISTICAL SUMMARY

covariate	Coefficient (β)	Hazard Ratio (HR)	P-value	95% CI
daycase_to_overnight_ratio	-0.129619	0.87843	0	[0.8767 - 0.8801]
emergency_to_elective_ratio	-0.184634	0.831408	0	[0.8239 - 0.8390]
referral_conversion_rate	-0.117156	0.889446	2.9373e-26	[0.8704 - 0.9089]
activity_diversity_index	1.07942	2.94298	0	[2.8075 - 3.0849]
workforce_squared	-2.15788e-09	1	0	[1.0000 - 1.0000]

Figure 17: Table showing coefficients, Hazard Ratios, and P-values for the five selected covariates.

A hazard ratio is the proportional change in risk for a one-unit increase in the covariate. So, for each one-unit increase in ratio, the risk of breach is multiplied by the hazard ratio (Figure 17). The top three features all have hazard ratios less than one. This means that increasing these metrics reduces the risk of a breach. They are therefore protective factors. The activity diversity index, however, has an HR of 2.942, indicating that increasing by 1 unit almost triples the risk of a breach. Workforce squared as a HR of 1. This is unusual and suggests that after accounting for the other four variables, it no longer has a proportional effect on the hazard. This could be due to severe collinearity among multiple variables. All variables have p-values less than 0.05, some as low as 0, indicating statistical significance.

4.5 Addressing Proportional Hazards Assumptions

The Final CoxPH model is statistically valid as the p-value remained below 0.05 and hence, satisfies the proportional hazards (PH) assumption. The PH assumption is that the HR for any given covariate must remain constant over the study period. For example, if increasing the 'activity diversity index' triples the risk of service failure on day 10, it must also do it on day 100 and 300. The baseline risk itself can change over time, but the relative effect of the covariate must stay the same.

4.6 Model output

The model was able to take user input to make personalised predictions, provide a risk analysis and risk mitigations (Figure 18-23).

```
# Initialize the predictor with your data
predictor = BreachRiskPredictor(
    predictions_df=results_final,  # Your predictions DataFrame
    model=final_cph,  # Your trained 66% concordance model
    latest_data_df=latest_trust_data,  # Latest trust data with all features
    dataset_start_date='2011-04-01'  # First date in your dataset
)

# Start interactive session
predictor.interactive_session()
```

```
================================================================================
NHS TRUST CANCER WAITING TIME BREACH RISK PREDICTOR
================================================================================

This tool predicts 62-day cancer waiting time target breach risk
based on operational metrics and historical performance.

Please enter the NHS Trust name or code (or 'quit' to exit):
```

Figure 18: Screenshot of the final output of the model. It depicts the model prompting for the first input, the NHS trust name or code.

```
================================================================================
NHS TRUST CANCER WAITING TIME BREACH RISK PREDICTOR
================================================================================

This tool predicts 62-day cancer waiting time target breach risk
based on operational metrics and historical performance.

Please enter the NHS Trust name or code (or 'quit' to exit): southampton

--------------------------------------------------------------------------------
❌ Trust 'southampton' not found in dataset.

Please try again with a different trust name.

Please enter the NHS Trust name or code (or 'quit' to exit):
```

Figure 19: After the input of "Southampton' the model signals that the name does not appear in the dataset, and to enter a different trust name.

```
================================================================================
NHS TRUST CANCER WAITING TIME BREACH RISK PREDICTOR
================================================================================

This tool predicts 62-day cancer waiting time target breach risk
based on operational metrics and historical performance.

Please enter the NHS Trust name or code (or 'quit' to exit): southampton

--------------------------------------------------------------------------------
❌ Trust 'southampton' not found in dataset.

Please try again with a different trust name.

Please enter the NHS Trust name or code (or 'quit' to exit): dartford

--------------------------------------------------------------------------------
Trust: Dartford and Gravesham NHS Trust
Status: Breach already occurred on May 01, 2020

Would you like to see the risk factors? (yes/no):
```

Figure 20: After inputting "Dartford' resulted in the model identifying Dartford and Gravesham NHS Trust. It then also gave the status of the breach. The user also has the option to see risk factors if they choose too.

```
📊 RISK FACTOR ANALYSIS:
--------------------------------------------------------------------------------

Top risk factors (ordered by impact):

1. activity_diversity_index
   Current value: 0.647
   High diversity - can adapt resources (GOOD)

2. referral_conversion_rate
   Current value: 0.117
   Low conversion rate - referral backlog (RISK)

3. daycase_to_overnight_ratio
   Current value: 0.006
   Low daycase ratio - inefficient, patients staying longer (RISK)

4. emergency_to_elective_ratio
   Current value: 0.003
   Good balance between emergency and planned care (GOOD)

5. workforce_squared
   Current value: 0.561
   Adequate workforce capacity (GOOD)

Would you like to see mitigation strategies? (yes/no):
```

Figure 21: The risk factor analysis is available to the user if they choose to see it. This is followed by the model giving the user the option to see mitigation strategies for the risks.

```
💡 RECOMMENDED MITIGATION STRATEGIES:
================================================================================

1. Low referral conversion rate
   Current rate: 0.09
   Target rate: >0.80

   Recommended actions:
   1) Increase outpatient clinic capacity
   2) Streamline referral pathways
   3) Implement direct booking systems
   4) Review and remove referral bottlenecks

2. Low daycase-to-overnight ratio
   Current ratio: 0.06
   Target ratio: >0.60

   Recommended actions:
   1) Review patient discharge protocols
   2) Implement enhanced recovery pathways
   3) Increase day surgery capacity
   4) Train staff in ambulatory care procedures

--------------------------------------------------------------------------------

Please enter the NHS Trust name (or 'quit' to exit):
```

Figure 22: When the user chooses to see the mitigation strategies of the risks. Once those have been presented, the user is able to either input another NHS trust name or quit the program.

```
================================================================================
NHS TRUST CANCER WAITING TIME BREACH RISK PREDICTOR
================================================================================

This tool predicts 62-day cancer waiting time target breach risk
based on operational metrics and historical performance.

Please enter the NHS Trust name (or 'quit' to exit): royal national orthopaedic

--------------------------------------------------------------------------------
⚠️ Multiple trusts found matching 'royal national orthopaedic':
  1. The Royal National Orthopaedic Hospital NHS Trust
  2. Royal National Orthopaedic Hospital NHS Trust

Enter the number of the correct trust (or 'no' to try again): 1
Trust: The Royal National Orthopaedic Hospital NHS Trust
Status: Breach predicted to occur on April 10, 2024 (in 48 months from May 2020)

Would you like to see the risk factors? (yes/no):
```

Figure 23: In the event of a hospital having similar names, there is the option to choose the one that the user means. The model also provides outputs for the predicted breach date and the number of months until it occurs.

Chapter 5

Discussion

5.1 The urgent need for a predictive paradigm

The persistent failure of the NHS to meet its cancer waiting time targets demonstrates that threshold-based, reactive management systems are insufficient. A fundamentally new approach is required, one grounded in proactive, data-driven prediction, to achieve true system resilience. The Cox Proportional Hazards (CoxPH) model developed here, which attained a concordance of 66.1% using 410,378 observations, moves beyond retrospective auditing by quantifying the risk of service failure at the individual trust level. This quantitative methodology provides the precise levers necessary for micro-level policy interventions, yielding valuable insights even during the exploratory phases, including the need to personalise aid for trusts due to the wide range of operational scales (outliers). As Sands et al. (2020) argue, viewing the health system through the lens of service failure and utilising early warning indicators (EWIs) is an urgent mandate.

The CoxPH model brought together specific theories and aspects from Gibbs et al. (2023) and Monikapreethi S K et al. (2024). It focused on the broader system-level impact of waiting times on population health and examined how the wider system influenced whether it experienced a breach. It took the idea of using CoxPH model to predict patient-level survival outcomes and applied them to the individual NHS trust.

5.2 Validation of the treatment bottleneck

The initial descriptive analysis validates the study's focus on the 62-day breach as a signal of structural failure. While the system appears highly productive, reflected in the strong correlation between "total treated" and "within standard", this productivity rises alongside an increase in breaches. On one level this proportional rise is expected, since treating more patients naturally places greater pressure on services. However, 'expected' does not mean 'acceptable'. The pattern suggests that although treatment volumes have increased, the resources required to support this growth have not kept pace. Trusts are managing higher demand with static or insufficient capacity, and policy has not adapted to ensure that increased patient throughput is matched by corresponding investment. The result is a system operating at or beyond its limits without the support needed to prevent breaches.

Access to detailed resource data would therefore be invaluable. If trusts could track whether proportional increases matched increases in patient volume in staffing, beds, or treatment capacity, targeted interventions would become far more precise. Having granular data on specific resources would strengthen this even further, allowing support to be tailored to the exact pressures each trust faces. This can be conceptualised in a similar way to the CoxPH patient-level framework proposed by Monikapreethi S K et al. (2024). In that analogy, the trust behaves like a patient, its resource indicators function like biomarkers, and the breach acts as the predicted clinical outcome. Effective treatment depends on accurately identifying which biomarkers require intervention, reinforcing the importance of detailed operational data.

The finding that the admitted breach rate is 38 percent higher than the non-admitted rate is clinically and operationally significant. The admitted rate was approximately 1.38 breaches per patient record compared to 1.01 for non-admitted cases. This disparity shows that the constraint is not diagnosis, which corresponds to the non-admitted portion of the 62-day pathway, but treatment. Breaches arise most frequently during the Decision to Treat 31-day standard, which applies to admitted patients who rely on fixed, high-capital resources such as surgery, chemotherapy or radiotherapy. These results demonstrate that the 62-day breach is, at its core, a measure of physical and human resource strain. The system can diagnose, but it cannot deliver. Shortages in surgical slots, inpatient beds, specialist staff and treatment capacity form the critical bottleneck in the cancer pathway. This highlights the need for targeted investment in post-diagnostic resources rather than broad, undifferentiated policy reforms.

5.3 Quantifying resilience

The Hazard Ratios derived from the CoxPH model identified several organisational factors that actively reduced the instantaneous risk of service failure and can therefore be considered markers of structural resilience. Two variables were statistically significant protective factors.

The first is the referral conversion rate, which produced an HR of 0.8894 ($p < 0.001$). This indicates that a trust with a unit increase in its efficiency in converting GP referrals into attended appointments experiences roughly an 11.1 percent reduction in the hazard of a breach. High conversion rates help maintain smooth patient flow at the diagnostic gateway, reducing backlogs and preventing delays from amplifying along the pathway. This suggests that referral conversion should be routinely assessed across all NHS trusts and that policies designed to strengthen this metric could meaningfully reduce breach risk.

The second protective factor is the daycase to overnight ratio, which has an HR of 0.8784 ($p < 0.001$). This reflects an approximate 12.2 percent reduction in the hazard for trusts shifting towards ambulatory models of care. This finding reinforces the value of investing in streamlined clinical pathways that minimise dependence on inpatient beds. By reducing pressure on acute bed capacity, trusts are better able to preserve the integrity of high-priority cancer services. Together, these protective HRs show that structural investment in workflow optimisation can materially reduce the risk of service failure.

The model also highlights the foundational role of human capital. The protective effect of the workforce buffer (represented by the squared workforce term) aligns with findings from external crises. The collapse of Greek oncology services was driven primarily by burnout and staff shortages (Pittaka et al., 2022), underscoring that staffing capacity is ultimately the ceiling of system resilience. The positive correlation between workforce levels and efficiency metrics, when considered alongside Liu and Liu’s (2022) evidence that incentives improve job performance and morale, points toward a clear policy direction. Supporting staff wellbeing, ensuring adequate workforce levels and providing meaningful incentives are not simply ethical responsibilities. They are also practical strategies that improve trust-level resilience by sustaining the efficiency indicators (HR below 1) that protect the cancer pathway.

It is important to acknowledge, however, that incentives can raise ethical concerns and have not always been implemented responsibly. This will be explored further in section 5.6.

5.4 The Activity Diversity Paradox

The model's finding that the activity diversity index is the single most significant risk factor is the central counter-intuitive finding of this research. It yielded a highly substantial Hazard Ratio of 2.943 (p-value <0.001), meaning that a unit increase in the breadth of service offered by a trust striples the instantaneous hazard of a breach. This result stands in direct opposition to the theoretical benefit of diversification (i.e. that it should improve resilience by spreading risk).

Instead, the evidence suggests a state of resource fragmentation. Specialised cancer services require the concentration of costly and scarce resources, dedicated oncology staff, high-capital equipment and dedicated surgical theatre time. When trusts attempt to sustain a wide range of diverse services, these essential capital and human resources are spread too thinly across competing service lines. The HR of 2.943 serves as a quantitative measure of this operational compromise: it shows that the attempt to maintain high diversity introduces a perpetual fragility into the high-priority cancer pathways. This conclusion fundamentally challenges conventional wisdom in healthcare management and should inform primary policy debates regarding the financial sustainability of generalised vs specialised healthcare services.

5.5 Interpreting the Emergency Contradictions

The second major contradiction in the model is the finding that the emergency-to-elective ratio has a protective HR of 0.8314 ($p < 0.001$). At face value, this implies that higher emergency pressure lowers the risk of a cancer breach, which conflicts with the standard operational assumption that emergency surges create hospital-wide delays. This assumption is well supported, including by the seasonal trend data in this study and by Sartina et al. (2022), who show that emergency department overcrowding negatively affects the entire hospital system.

However, the interpretation of this variable is limited by the data available. We do not know whether the emergency to elective ratios in this dataset genuinely indicate overcrowding, because we lack the capacity figures for each trust. A hospital may appear to have a high emergency burden simply because it is a designated emergency hub or a large teaching hospital with significantly more resources. Without granular data on hospital size, bed stock, staffing levels and emergency designation, we cannot definitively interpret the ratio as a measure of strain.

The most plausible explanation is that the result reflects the system's compensatory behaviour and the political weight of cancer targets. Trusts that consistently manage high emergency volumes, particularly major teaching hospitals, tend to have stronger organisational infrastructure and deeper experience in rapid triage. When their emergency activity spikes, the political and financial consequences of a cancer breach compel them to protect cancer pathways at all costs. They achieve this by cancelling or postponing lower-priority elective procedures, such as orthopaedics or ophthalmology, to free up staff, theatres and beds.

The inclusion of the non-linear workforce squared term is also important. It captures organisational size and buffer capacity, helping explain why larger trusts with higher emergency volumes can still avoid a breach. The protective HR for the emergency ratio is therefore likely an artefact of these compensatory strategies: these trusts are just able to adapt aggressively enough to prevent an official breach, even while operating under considerable strain. In practice this means that maintaining compliance comes at the expense of other patient groups whose care is routinely delayed.

5.6 Ethical and Policy Dynamics

The systemic fragility identified in this research, quantified through Hazard Ratios and evident in acute bottlenecks, highlights the need for a fundamental shift in policy response. The NHS has traditionally relied on incentive-based mechanisms to enforce standards, but these systems carry significant risks and strengthen the case for a fully data-driven approach. As Li and Evans (2022) note, threshold-based targets can encourage gaming behaviour, where trusts prioritise easily achievable metrics or manipulate reporting to secure funding. This is not a hypothetical concern. The recent case of GP practices retaining “ghost patients” to obtain financial advantage (Guardian staff reporter, 2024) demonstrates how incentives can directly undermine data integrity.

Although Liu and Liu (2022) show that incentives can improve job performance and efficiency, and may help mitigate burnout such as that documented in the Greek oncology crisis, the predictive capacity of the CoxPH model offers a more reliable alternative to reward-based enforcement. The critical advantage of the CoxPH early warning system is that it allocates resources based on a risk calculation derived from objective, verifiable, third-party operational data (e.g. workforce figures) rather than subjective performance metrics that are easily manipulated. This approach reduces the ethical risks associated with gaming and ties resource allocation directly to the objective likelihood of system failure.

Under this framework, support is directed at trusts with quantifiable risk (HR greater than 1) and at those with the capacity to benefit from structural improvement (indicated by protective metrics such as efficiency ratios with HR below 1). The policy recommendations emerging from the model are not prescriptive mandates. Instead, they are data-informed scenarios showing what proactive support could look like when tailored to the specific risk profile of each trust.

Chapter 6

Conclusion

6.1 Summary of Key Findings

This study set out with a clear aim: to determine whether a predictive survival model could quantify the risk of an NHS trust breaching the 62 day cancer standard and, in doing so, offer a new way of understanding system strain. The long-standing failure to meet cancer waiting time targets showed that the current system, built on retrospective thresholds, reacts only after the damage is done. A different method was needed, one capable of forecasting risk rather than auditing failure.

Using a Cox Proportional Hazards model built on 410,378 observations, the analysis showed that trust-level breach risk can be quantified with meaningful accuracy. The descriptive phase confirmed that the 62 day breach reflects structural weakness. Productivity rose over time, but breaches rose with it, signalling a mismatch between demand and capacity. The 38 percent higher admitted breach rate underscored that treatment delivery, not diagnosis, is the true point of collapse. Seasonal patterns, outliers and structural inconsistencies further reinforced the fragility of the system.

The CoxPH model then ranked the variables shaping that fragility. Activity diversity was the most powerful risk factor. Referral conversion, daycase reliance and workforce buffers were protective. Together, these results demonstrated that cancer delays function as reliable indicators of operational pressure, and that predictive modelling can pinpoint where resilience is weakening long before failure occurs.

6.2 Main Conclusions

The central conclusion of this research is that the 62 day breach is a measurable expression of systemic overload. When an NHS trust misses this target, it is not simply a reporting failure but a signal that its high-capital, resource-intensive treatment capacity has been pushed past its limits. The CoxPH model quantified this dynamic, revealing a clear hierarchy of the forces that accelerate or slow down the risk of a breach.

Activity diversity emerged as the strongest accelerant. A wider service portfolio increases the hazard rate substantially, suggesting that trying to maintain too many service lines spreads expertise, equipment and specialist staff too thinly. This finding challenges the assumption that diversification inherently increases resilience. Instead, it introduces fragility into the cancer pathway.

Referral conversion and the daycase to overnight ratio acted as stabilisers by improving flow and reducing pressure on inpatient resources. Workforce strength provided a further buffer, echoing international evidence that staff shortages and burnout are early warning signs of collapse.

Together, these findings show that cancer delays are a valid proxy for system instability. The breach occurs precisely where physical and human capacity meet rising demand and fail to keep pace. The model’s ability to quantify this risk confirms that a predictive paradigm is both realistic and necessary. It offers a structured way to identify trusts approaching a tipping point and highlights which operational levers offer the strongest chance of preventing failure.

6.3 Recommendations

The recommendations here translate the model’s findings into practical actions for NHS leadership, policymakers and researchers. They focus on strengthening resilience rather than short-term performance.

Workforce and Wellbeing

Staffing capacity is the foundation of system stability. Trusts should prioritise recruitment pipelines, burnout mitigation programmes and protected oncology staffing. Workforce buffers should be aligned with historical demand peaks and local emergency pressures. Rather than relying on generic incentives, trusts should adopt evidence-based approaches that reinforce the protective factors identified in the model. This includes linking staffing investment to improvements in referral conversion and daycase throughput.

Targeted Capital and Capacity Investment

Investment should directly expand treatment capacity. Trusts with high hazard ratios should be prioritised for additional theatres, oncology bed space, chemotherapy chairs and radiotherapy access. The model highlights that breaches occur in treatment, not diagnosis, so capital expansion must be concentrated on post-diagnostic bottlenecks. National planners should also consider supporting regional consolidation of highly specialised services to safeguard cancer pathways at smaller or heavily diversified trusts.

Technology and Predictive Infrastructure

The next step is to embed predictive analytics into routine operations. A national survival-based risk dashboard would allow trusts to view breach risk in real time and act before thresholds are crossed. Integrating data on workforce, theatre utilisation, bed flow and referrals would create a continuous early warning system. Scenario modelling should be built into this platform so trusts can test the impact of operational changes and prepare for seasonal pressure.

Operational Planning and Pathway Design

Trusts should regularly assess their referral conversion rate and daycase utilisation, as both metrics have clear protective effects. Pathway redesign should focus on strengthening these ratios through improved triage, centralised scheduling and standardised clinical protocols. Activity diversification must also be monitored closely. Expanding service lines without a matched increase in workforce or capital is associated with a substantial rise in breach risk.

Policy and Governance

The NHS should shift from target-driven incentives to risk-based resource allocation. Threshold-based systems encourage data manipulation and short-term behaviour, whereas predictive indicators ground decisions in observable operational strain. Funding should be linked to measured hazard levels and the capacity for structural improvement. This ensures that support is directed to trusts where it will prevent system failure rather than reward performance that may not reflect underlying conditions.

6.4 Avenues for Future Development and Research

There are several clear opportunities to extend this work. The first is applying the predictive framework to other NHS services that experience similar bottlenecks, such as elective surgery, diagnostics, maternity care and emergency medicine. Each of these sectors faces capacity thresholds, and survival modelling could reveal their unique tipping points.

Future research should also integrate qualitative data. Understanding how operational leaders respond to rising pressure, or how clinical decision-making changes during periods of strain, would provide essential context for the hazard ratios identified in this study. Incorporating economic indicators such as inflation, wage pressures and drug costs would further enhance predictive accuracy.

Greater granularity is another major step. Access to bed occupancy data, theatre timetables, cancer-type specific pathways and staff rota information would allow the model to pinpoint exactly where delays originate within each trust. This would enable more precise policy interventions.

Finally, the methodological framework developed here is highly transferable. Any public service with finite capacity and variable demand can benefit from survival-based prediction. Social care systems, school capacity planning, emergency response networks and even local authority services can adapt this approach to anticipate strain before failure occurs. The value of this research lies not only in its application to cancer pathways but in the broader capacity it demonstrates for predictive governance.

References

Aggarwal, A., Choudhury, A., Fearnhead, N., Kearns, P., Kirby, A.M., Lawler, M., Quinlan, S., Palmieri, C., Roques, T., Simcock, R., Walter, F.M., Price, P. and Sullivan, R. (2023) 'The future of cancer care in the UK—time for a radical and sustainable National Cancer Plan', *Lancet Oncology*, 25(1). doi:https://doi.org/10.1016/s1470-2045(23)00511-9.

Aggarwal, A., Han, L., Sullivan, R., Haire, K., Sangar, V. and van der Meulen, J. (2023) 'Managing the cancer backlog: a national population-based study of patient mobility, waiting times and 'spare capacity' for cancer surgery', *The Lancet Regional Health – Europe*, 30, p. 100642.

Aggarwal, A., Simcock, R., Price, P., Rachet, B., Lyratzopoulos, G., Walker, K., Spencer, K., Roques, T. and Sullivan, R. (2024a) 'NHS cancer services and systems—ten pressure points a UK cancer control plan needs to address', *Lancet Oncology*, 25(8), pp. e363-e373. doi:10.1016/S1470-2045(24)00345-0.

Aggarwal, A., Walter, F.M., Sullivan, R. and van der Meulen, J. (2022) '"Shopping around" for treatment is not a solution to cancer backlog', *BMJ*, 379, p. e071967.

Asfoor, D. A., Tabche, C., Al-Zadjali, M., Mataria, A., Saikat, S., & Rawaf, S. (2024). Concept analysis of health system resilience. *Health Research Policy and Systems*, *22*, 43. https://doi.org/10.1186/s12961-024-01114-w

Bajwa J, Munir U, Nori A, Williams B. Artificial intelligence in healthcare: transforming the practice of medicine. Future Healthc J. 2021 Jul;8(2):e188-e194. doi: 10.7861/fhj.2021-0095. PMID: 34286183; PMCID: PMC8285156.

Bajwa, J., Munir, U., Nori, A. and Williams, B. (2021) 'Artificial intelligence in healthcare: transforming the practice of medicine'. *Future Healthcare Journal*, 8(2), pp. e188–e194.

Broadbent, P., Walsh, D., Katikireddi, S.V., Gallagher, C., Dundas, R. and McCartney, G. (2024) 'Is Austerity Responsible for the Stalled Mortality Trends Across Many High-Income Countries?

A Systematic Review'. *International Journal of Social Determinants of Health and Health Services*, 54(4), pp. 362–379.

Brookes, G. and Baker, P. (2023) 'Cancer services patient experience in England: quantitative and qualitative analyses of the National Cancer Patient Experience Survey', *BMJ Supportive & Palliative Care*, 13(2), pp. e1149-e1155.

Cima, J.D.F. and Almeida, A.F.S. (2024) 'Waiting times spillovers in a National Health Service hospital network: a little organizational diversity can go a long way', *Health Economics Review*, 14(87). doi:10.1186/s13561-024-00555-x.

Cooksley T, Clarke S, Dean J, Hawthorne K, James A, Tzortziou-Brown V et al. (2023). *NHS crisis: rebuilding the NHS needs urgent action.* BMJ, 380:p1. https://doi.org/10.1136/bmj.p1

Cox, S.R. and Daniel, C.L. (2022) 'Racial and Ethnic Disparities in Laryngeal Cancer Care', *Journal of Racial and Ethnic Health Disparities*, 9(3), pp. 800–811. doi:10.1007/s40615-021-01018-3.

Croghan, I.T., Chesak, S.S., Adusumalli, J., Fischer, K.M., Beck, E.W., Patel, S.R., Ghosh, K., Schroeder, D.R. and Bhagra, A. (2021) 'Stress, Resilience, and Coping of Healthcare Workers during the COVID-19 Pandemic'. *Journal of Primary Care & Community Health*, 12, pp. 1–9.

Dawkins, B., Renwick, C., Ensor, T., Shinkins, B., Jayne, D. and Meads, D. (2021) 'What factors affect patients' ability to access healthcare? An overview of systematic reviews'. *Tropical Medicine & International Health*, 26(10), pp. 1177–1188.

Denfeld, Q.E., Burger, D. and Lee, C.S. (2023) 'Survival analysis 101: an easy start guide to analysing time-to-event data', *European Journal of Cardiovascular Nursing*, 22(3), pp. 332–337. doi:10.1093/eurjcn/zvad023.

Dicuonzo, G., Donofrio, F., Fusco, A. and Shini, M. (2022). Healthcare system: Moving forward with artificial intelligence. Technovation, 120(1), p.102510. doi:https://doi.org/10.1016/j.technovation.2022.102510.

Dodhia, V., Parkes, J., Ghose, A. and Ng, W. (2023) 'The NHS SACT capacity crisis: Challenges and solutions', *Journal of Oncology Pharmacy Practice*, 0(0), p. 10781552231185050.

Esmaeilzadeh P. Challenges and strategies for wide-scale artificial intelligence (AI) deployment in healthcare practices: A perspective for healthcare organizations. Artif Intell Med. 2024 May;151:102861. doi: 10.1016/j.artmed.2024.102861. Epub 2024 Mar 3

Eurostat (2025). *Glossary:Early warning system - Statistics Explained - Eurostat*. [online] Europa.eu. Available at: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Early_warning_system [Accessed 10 Jul. 2025].

Eurostat (n.d.) *Glossary: Early warning system*, *Eurostat*. Available at: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Early_warning_system#:~:text=Enter_search_term-,Early_warning_system,a_subset_of_such_indicators (Accessed: 10 July 2025).

Fleming, P., Caffrey, L., Van Belle, S., Barry, S., Burke, S., Conway, J., Siersbaek, R. and Thomas, S. (2023) 'How International Health System Austerity Responses to the 2008 Financial Crisis Impacted Health System and Workforce Resilience - A Realist Review'. *International Journal of Health Policy and Management*, 12, p. 7420.

Fleron, A., Krishna, A. and Singhal, S. (2022) *The gathering storm: The transformative impact of inflation on the healthcare sector*.

Fox, L., Monroy-Iglesias, M.J., Aggarwal, A., Haire, K., Purushotham, A., Spicer, J., Papa, S., Rigg, A., Dolly, S. and Sullivan, R. (2022) 'Association between COVID-19 burden and delays to diagnosis and treatment of cancer patients in England', *Journal of Cancer Policy*, 31, p. 100316.

Gathani, T., Clayton, G., MacInnes, E. and Horgan, K. (2021) 'The COVID-19 pandemic and impact on breast cancer diagnoses: what happened in England in the first half of 2020', *British Journal of Cancer*, 124(4), pp. 710–712.

Ghanad, A. (2023). *An Overview of Quantitative Research Methods*. [online] ResearchGate. Available at:

https://www.researchgate.net/publication/373370007_An_Overview_of_Quantitative_Research_Methods.

Gibbs, N.K., Griffin, S., Gutacker, N., Villasenor-Lopez, A. and Walker, S. (2023) *A general modelling framework to estimate the impact of waiting times on health for patients undergoing elective procedures in the NHS in England*.

Giovanola, B., Tiribelli, S. Beyond bias and discrimination: redefining the AI ethics principle of fairness in healthcare machine-learning algorithms. AI & Soc 38, 549–563 (2023). https://doi.org/10.1007/s00146-022-01455-6

Guardian staff reporter (2024). *Number of 'ghost patients' registered with NHS GPs up two-thirds since 2018*. [online] the Guardian. Available at: https://www.theguardian.com/society/2024/jan/02/number-of-ghost-patients-registered-with-nhs-gps-up-two-thirds-since-2018 [Accessed 7 Nov. 2025].

Hartman, M., Martin, A.B., Whittle, L., Catlin, A. and National Health Expenditure Accounts Team (2024) 'National Health Care Spending In 2022: Growth Similar To Prepandemic Rates'. *Health Affairs*, 43(1), pp. 6–17.

Hayes, H., Meacock, R., Stokes, J. and Sutton, M. (2024) 'How do family doctors respond to reduced waiting times for cancer diagnosis in secondary care?', *The European Journal of Health Economics*, 25, pp. 813–828.

Indrayan, A. and Tripathi, C.B. (2022) 'Survival Analysis: Where, Why, What and How?', *Indian Pediatrics*, 59, pp. 74–79. doi:10.1007/s13312-022-2425-5.

Kirk-Wade, E., Harker, R. and Stiebahl, S. (2024) *NHS key statistics: England*.

Lawler, M., Aggarwal, A., Gralow, J., Sullivan, R. and Price, P. (2024) 'The UK needs to be a leader, not a lagger, in the global cancer effort', *The Lancet Oncology*, 25(10), pp. 1253–1254.

Li YH, Li YL, Wei MY, Li GY. Innovation and challenges of artificial intelligence technology in personalized healthcare. Sci Rep. 2024 Aug 16;14(1):18994. doi: 10.1038/s41598-024-70073-7. PMID: 39152194; PMCID: PMC11329630.

Li, X., Evans, J.M. (2022). *Incentivizing performance in health care: a rapid review, typology and qualitative study of unintended consequences.* BMC Health Serv Res, 22:690. https://doi.org/10.1186/s12913-022-08032-z

Liaqat A, Gallier S, Reeves K, Crothers H, Evison F, Schmidtke K, et al. (2022). *Examining organisational responses to performance-based financial incentive systems: a case study using NHS staff influenza vaccination rates from 2012/2013 to 2019/2020.* BMJ Qual Saf. 31(9):642–651.

Limiri, D. (2025) 'The Impact of Long Wait Times on Patient Health Outcomes: The Growing NHS Crisis', *Premier Journal of Public Health*. doi:https://doi.org/10.70389/PJPH.100020.

Liu, W., & Liu, Y. (2022). The Impact of Incentives on Job Performance, Business Cycle, and Population Health in Emerging Economies. *Frontiers in Public Health*, *9*, 778101. https://doi.org/10.3389/fpubh.2021.778101

Lyon, J.Y., Bogodistov, Y. and Moormann, J. (2021). AI-driven Optimization in Healthcare: the Diagnostic Process. European Journal of Management Issues, 29(4), pp.218–231. doi:https://doi.org/10.15421/192121.

Maleki Varnosfaderani, S. and Forouzanfar, M. (2024) 'The Role of AI in Hospitals and Clinics: Transforming Healthcare in the 21st Century'. *Bioengineering*, 11(4), p. 337.

Mennella, C., Maniscalco, U., Pietro, G.D. and Esposito, M. (2024). Ethical and Regulatory Challenges of AI Technologies in healthcare: a Narrative Review. Heliyon, [online] 10(4), pp.e26297–e26297. doi:https://doi.org/10.1016/j.heliyon.2024.e26297.

Monikapreethi S K, Preetha, J., Reddy, K. N., Ramya, S., Yuvaraj, S. and Murugan, S. (2024) "Survival Analysis with Cox Proportional Hazards Model in Predicting Patient Outcomes," *2024*

5th International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2024, pp. 1155-1161. doi: 10.1109/ICESC60852.2024.10689732.

National Cancer Institute. (2024). *Cancer Statistics.* [Online] Available at: https://www.cancer.gov/about-cancer/understanding/statistics

NHS England (2023) *Changes to cancer waiting times standards from 1 October 2023*, *NHS England*. Available at: https://www.england.nhs.uk/long-read/changes-to-cancer-waiting-times-standards-from-1-october-2023/ (Accessed: 11 July 2025).

OECD. (2020). *Waiting Times for Health Services: Next in Line*, OECD Health Policy Studies, OECD Publishing, Paris. https://doi.org/10.1787/242e3c8c-en

Ortiz, I. and Cummins, M. (2021) 'The Austerity Decade 2010-20'. *Social Policy & Society*, 20(1), pp. 142–157.

Pasricha, S. (2022). AI Ethics in Smart Healthcare. IEEE Consumer Electronics Magazine, 12(4), pp.1–7. doi:https://doi.org/10.1109/mce.2022.3220001.

Pittaka M, Sakellakis M, Metaxas V, Kardamakis D, Wagland R. (2022). *Burnout Syndrome among Doctors in Greek Oncology Departments.* Iran J Psychiatry. 17(2):162-176. https://doi.org/10.18502/ijps.v17i2.8906

Pozovna, I., Haluszko, K. and Polishchuk, A. (2023) 'MODELLING STRUCTURAL RELATIONS BETWEEN FINANCIAL, SOCIAL-ECONOMIC AND HEALTHCARE FACTORS'. *Health Economics and Management Review*, (4).

Quinn, L., Bird, P. & Lilford, R. (2024). *Effect of cancer waiting time standards in the English National Health Service: a threshold analysis.* BMC Health Serv Res, 24:929. https://doi.org/10.1186/s12913-024-11350-z

Rane, N., Choudhary, S. and Rane, J. (2024). Artificial Intelligence for Enhancing Resilience. Journal of Applied Artificial Intelligence, 5(2), pp.1–33. doi:https://doi.org/10.48185/jaai.v5i2.1053.

Sai, S., Gaur, A., Sai, R., Vinay Chamola, Mohsen Guizani and Joel (2024). Generative AI for Transformative Healthcare: A Comprehensive Study of Emerging Models, Applications, Case Studies and Limitations. IEEE Access, 12, pp.1–1. doi:https://doi.org/10.1. (n.d.)

Sands, S., Campbell, C., Shedd, L., Ferraro, C. and Mavrommatis, A. (2020) 'How small service failures drive customer defection: Introducing the concept of microfailures', *Business Horizons*, 63(4), pp. 573–584. doi:10.1016/j.bushor.2020.03.014.

Sartini, M., Carbone, A., Demartini, A., Giribone, L., Oliva, M., Spagnolo, A. M., Cremonesi, P., Canale, F., & Cristina, M. L. (2022). Overcrowding in Emergency Department: Causes, Consequences, and Solutions—A Narrative Review. *Healthcare*, *10*(9), 1625. https://doi.org/10.3390/healthcare10091625

Smallman, M. (2022). Multi Scale Ethics—Why We Need to Consider the Ethics of AI in Healthcare at Different Scales. Science and Engineering Ethics, 28(6). doi:https://doi.org/10.1007/s11948-022-00396-z..

Solbjør, M., Sand, K., Ervik, B. and Melby, L. (2021) 'Patient experiences of waiting times in standardised cancer patient pathways in Norway: a qualitative interview study', *BMC Health Services Research*, 21(1), p. 651.

Taimoor, N. and Rehman, S. (2022). Reliable and Resilient AI and IoT-Based Personalised Healthcare Services: A Survey. IEEE Access, 10, pp.535–563. doi:https://doi.org/10.1109/access.2021.3137364.

van der Schaar, M., Alaa, A.M., Floto, A., Gimson, A., Scholtes, S., Wood, A., McKinney, E., Jarrett, D., Lio, P. and Ercole, A. (2020). How artificial intelligence and machine learning can help healthcare systems respond to COVID-19. Machine Learning, 11

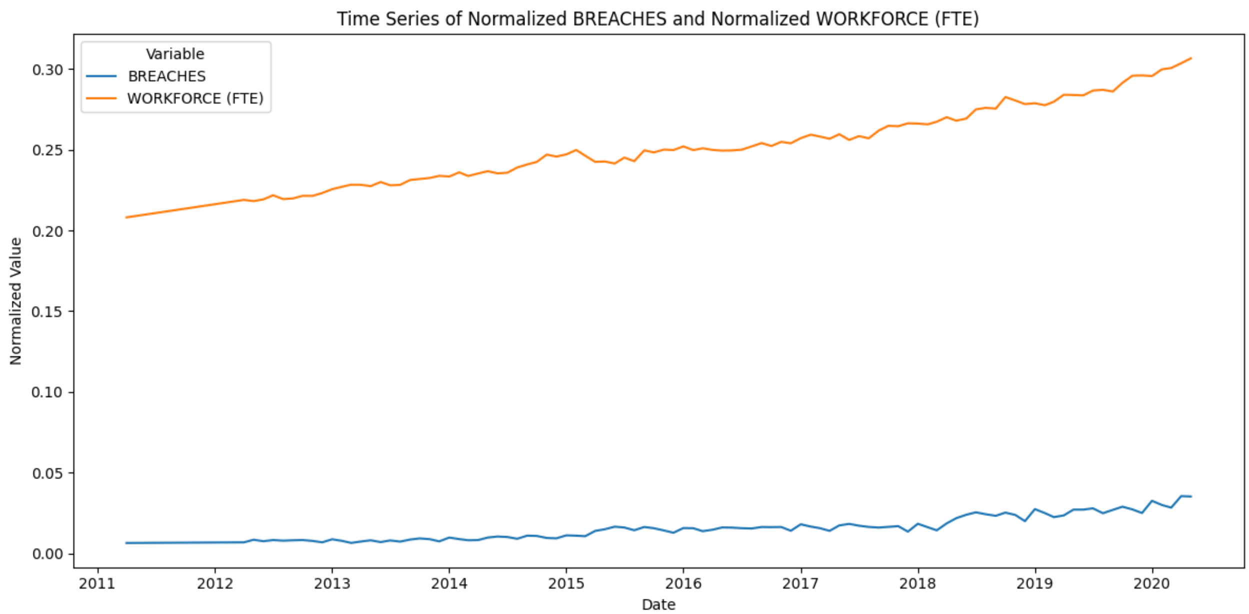
van Elk, S., Reinecke, J., Trenholm, S. and Ferlie, E. (2024) 'Constructing promissory futures to defer moral scrutiny: The dilemma of healthcare austerity'. *Human Relations*, 78(8), pp. 995–1029.

Varnosfaderani, S.M. and Forouzanfar, M. (2024). The Role of AI in Hospitals and Clinics: Transforming Healthcare in the 21st Century. Bioengineering, [online] 11(4), pp.1–38. Available at: https://www.mdpi.com/2306-5354/11/4/337.

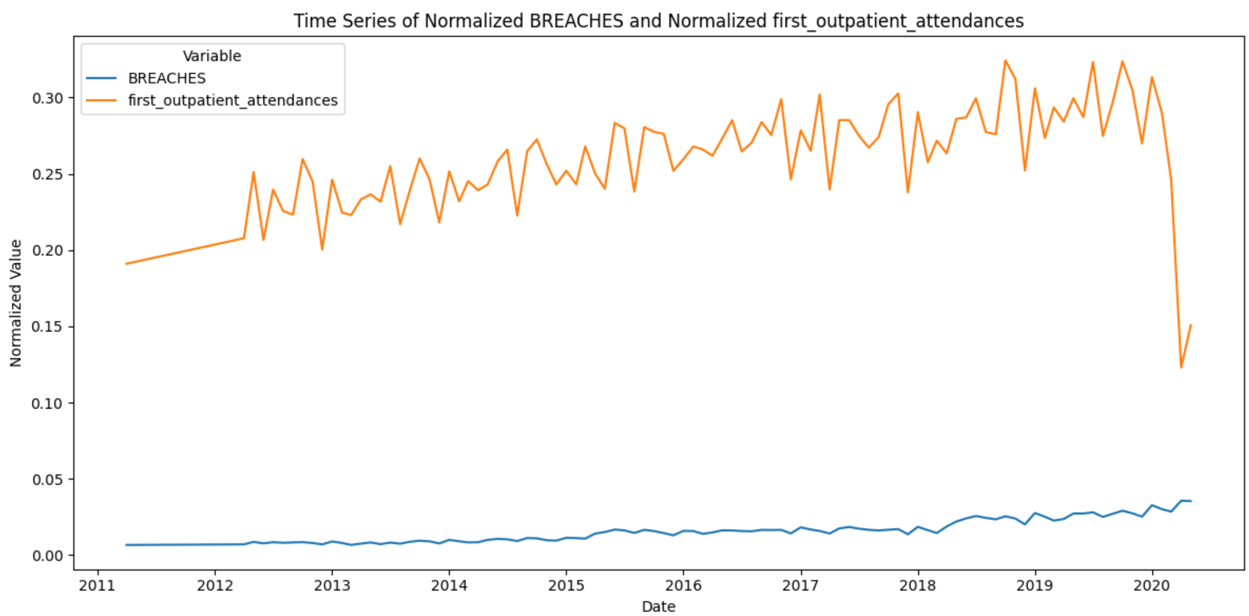
Warner, M. and Zaranko, B. (2024). *The past and Future of NHS Waiting Lists in England.* Institute for Fiscal Studies. [Online] Available at: https://ifs.org.uk/publications/past-and-future-nhs-waiting-lists-england

World Health Organization. (2025). *Cancer*. [Online] Available at: https://www.who.int/news-room/fact-sheets/detail/cancer

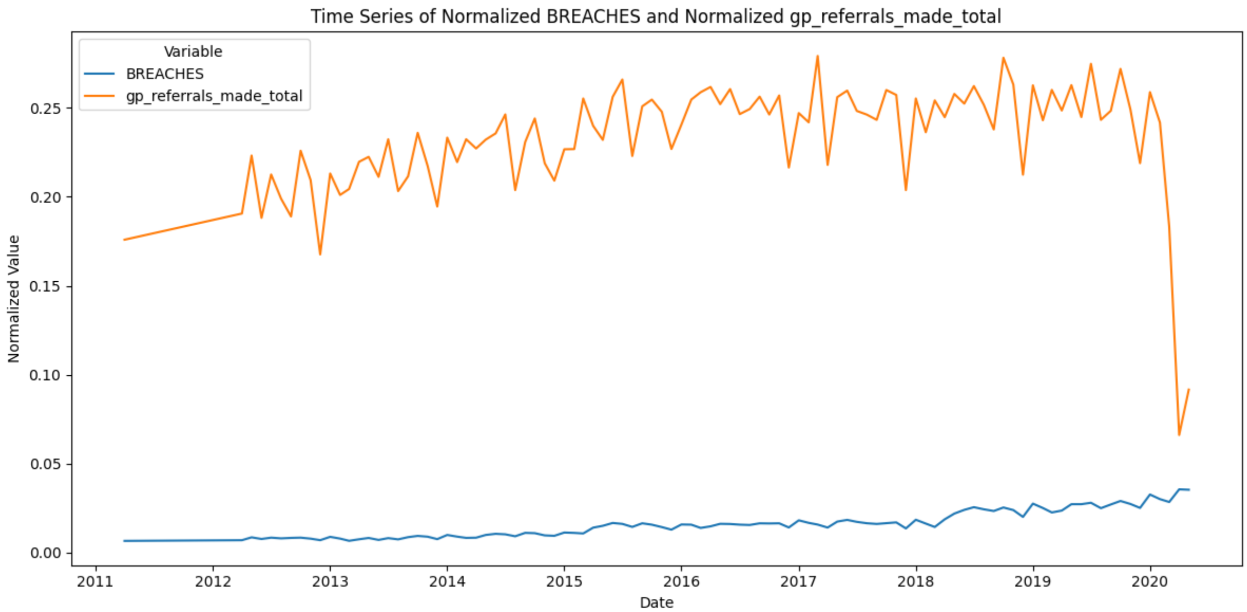
Appendices



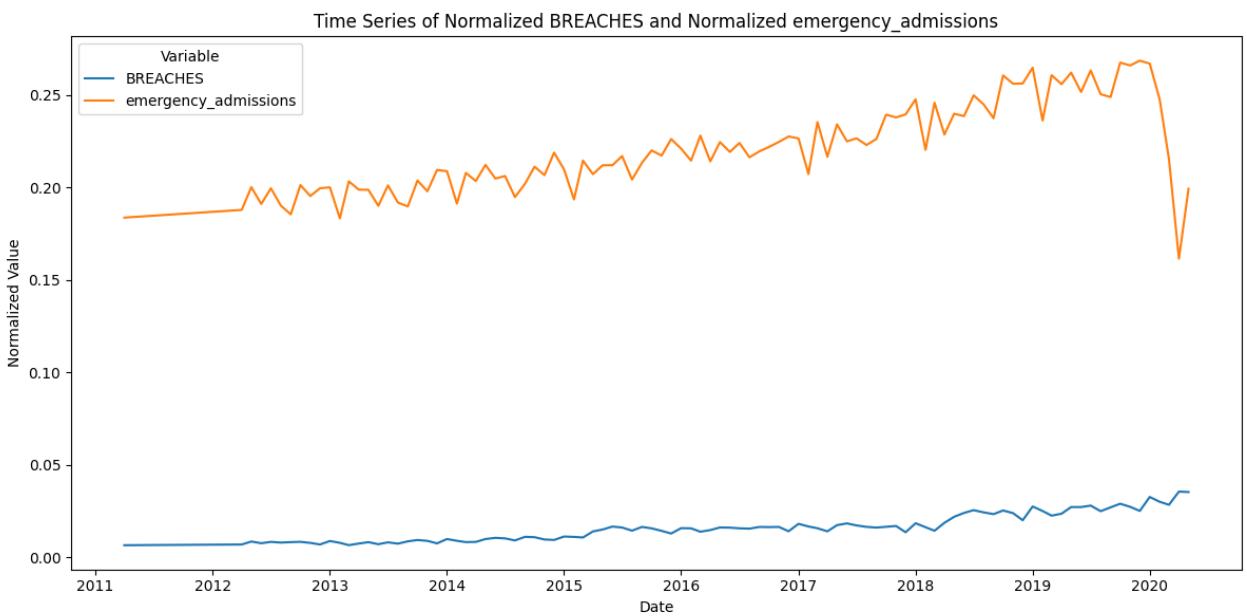
Appendix 1: Time series graph of Breaches and normalised Workforce (FTE)



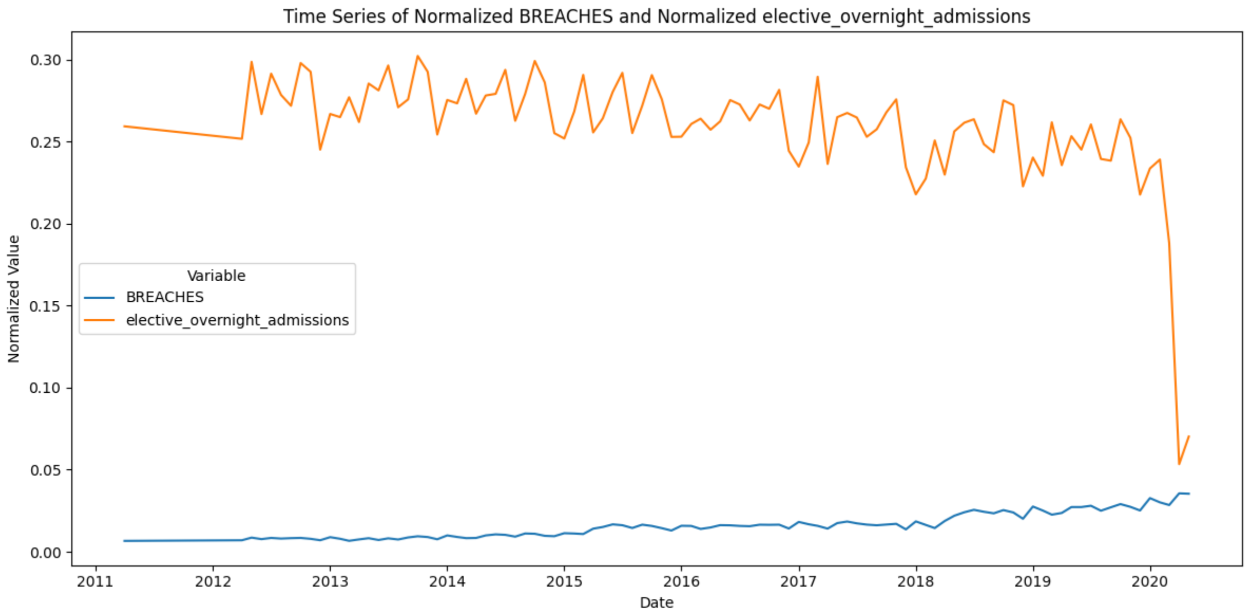
Appendix 2: Time series graph of Breaches and first_outpatint_attendances



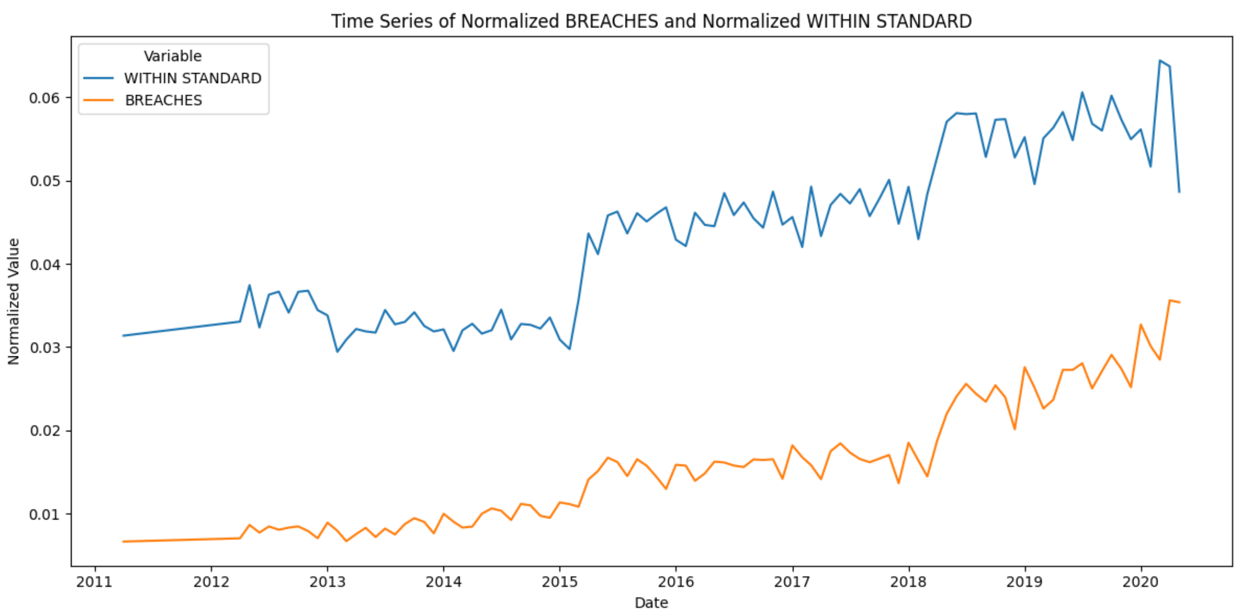
Appendix 3: Time Series graph of breaches and normalised gp_referrals_made_total



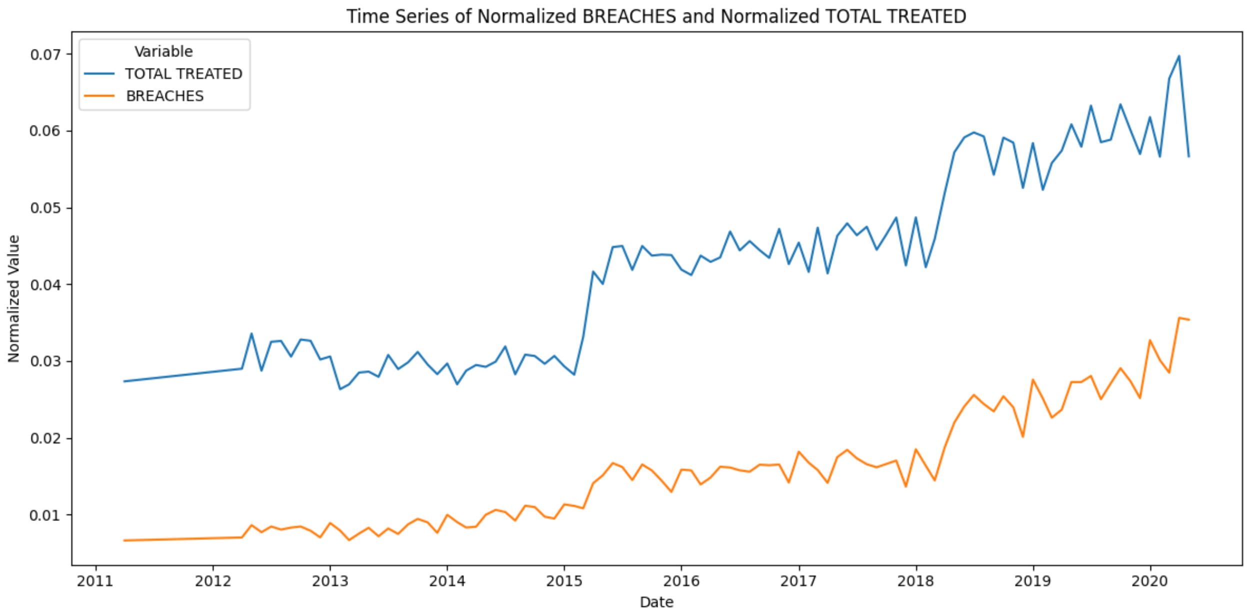
Appendix 4: Time series graph of breaches vs emergency_admissions.



Appendix 5: Time series graph of Breaches and normalised elective_overnight_admissions.



Appendix 6: Time series graph of breaches and normalised within the standard.



Appendix 7: Time series graph of breaches and normalised total treated.