Decoding UK Hospital Wait Times: Systemic Challenges Beyond Resources*

How Infrastructure, Workforce, and Demand do not Explain Delays in Medical Care

Chenika Bukes

December 3, 2024

Hospital wait times in the United Kingdom are a critical indicator of healthcare efficiency, reflecting the interplay between resources and patient demand. This study investigates how hospital infrastructure, workforce availability, and patient attendance rates influence wait times for fifteen key medical procedures between 2015 and 2022. Regression modeling reveals no statistically significant predictors, with hospital bed availability, physician density, and attendance rates showing weak associations with wait times. These findings suggest that systemic inefficiencies and external disruptions, such as the COVID-19 pandemic, play a more significant role than individual resource metrics. This study underscores the need for a deeper understanding of healthcare delays and highlights the importance of data-driven strategies to improve resource allocation and patient outcomes.

Table of contents

1

1 Introduction

Timely access to healthcare is fundamental to achieving optimal health outcomes. In the United Kingdom, the National Health Service (NHS) faces mounting pressure to manage increasing patient demands amidst resource constraints. Hospital wait times, a key performance indicator, have become a focal point for policymakers and stakeholders alike (OECD 2024b). Extended wait times not only compromise patient care but also signal underlying inefficiencies in healthcare delivery systems.

This paper investigates how hospital infrastructure, workforce availability, and patient demand influence wait times for medical procedures across the UK. Specifically, it examines the roles of hospital beds per capita, physicians per 1,000 inhabitants, and Type 1 major care attendance rates in shaping delays from 2015 to 2022 for fifteen major fields of medicine presented in Table 3. The NHS consistently met the 18-week standard from 2008 to 2015, but has not since, thus the years 2015 on wards were selected for investigation. NHS waiting times continue to sharply increase making this a national healthcare crisis in the UK (Laura Salisbury 2023).

The estimand in this study is the is the median number of weeks spent waiting for treatment since referral averaged over fifteen key fields of medicine Table 3. The estimand will be modeled as a function of three predictors: hospital beds per 1,000 inhabitants, physicians per 1,000 inhabitants, and Type 1 major care attendance rates. This measure quantifies how changes in healthcare infrastructure and demand impact system-wide delays, offering insights into the relative contribution of each predictor to observed wait times.

The findings reveal that none of the predictors have a statistically significant association with wait times, highlighting the complexity of the relationships between healthcare resources, workforce availability, and patient demand. Declining hospital bed availability, despite being hypothesized to have a strong effect, shows only a weak positive relationship with wait times, suggesting that other systemic factors, such as bed utilization efficiency and discharge planning, may overshadow its direct impact. Similarly, higher physician density is associated with increased delays, potentially reflecting inefficiencies in resource allocation or increased demand generated by greater access. Attendance rates also show a weak positive association, but the substantial deviation from pre-COVID trends, particularly the sharp drop in 2020 and rebound in 2021, likely undermines the robustness of the findings.

Understanding the drivers of wait times is vital for improving healthcare access and efficiency (Fund 2024a). By highlighting the multifaceted nature of healthcare delays and the limitations of current models, this paper underscores the need for more granular data and advanced methodologies to uncover the true determinants of wait times. The remainder of this paper is structured as follows: Section 2 details the data and methodology, Section 3 presents the model used for analysis of the data, Section 4 presents the key results from the model, and Section 5 concludes with final analysis of the results and recommendations for future research.

2 Data

2.1 Overview

In this paper, the analysis will be carried out using the statistical programming language `R` (R Core Team 2023) and use the `tidyverse` (Wickham et al. 2019), `devtools` (Wickham, Hester, and Chang 2020) and `dplyr` (Wickham et al. 2021) packages. All figures in the report are generated using `ggplot2` (Wickham 2016) with `patchwork` (Pedersen 2024), and `gridExtra` (Auguie 2017).

This report integrates data from two key sources: the NHS and the Organisation for Economic Co-operation and Development (OECD). These datasets provide critical information to analyze trends in healthcare demand and resources in the United Kingdom.

The NHS Referral to Treatment (RTT) Waiting Times dataset (England 2024) is the primary source for waiting time statistics. These data measure the median time (in weeks) between

referral by a general practitioner (GP) and the start of treatment for elective care, capturing delays across 15 key medical procedures Table 3. This dataset offers procedure-specific insights that allow for cross-procedure and temporal comparisons. The data are aggregated and published monthly.

The NHS A&E Attendances and Emergency Admissions dataset (NHS 2024) captures the monthly demand for emergency services. It includes the total number of attendances at Accident & Emergency (A&E) departments and emergency admissions, alongside measures of wait times for admission. These statistics are collected at the provider organization level from NHS Trusts, Foundation Trusts, and independent sector organizations. The data are aggregated from monthly submissions.

Additional datasets are sourced from the OECD Health Statistics 2024 database (OECD 2024b), which provides standardized, internationally comparable data on healthcare systems. For the United Kingdom, the OECD compiles data on hospital resources hospital beds and physician counts from NHS Digital and Public Health Scotland (Scotland 2023). The OECD ensures consistency and comparability of this data through rigorous quality checks and methodological adjustments.

2.2 Measurement

Physicians per 1,000 (OECD 2024c). Physician counts are sourced from the OECD, relying on NHS Digital, Public Health Scotland, and the General Medical Council (GMC). The data reflect licensed physicians, encompassing both general practitioners and specialists. The metric is calculated as the total number of physicians per 1,000 population. The dataset includes both headcount and rolecount metrics, with post-2009 data transitioning to headcount for greater accuracy. Physician counts include GP retainers and full-time equivalents, providing an accurate picture of the available workforce. Historical data adjustments account for changes in collection methodologies. The data are collected and reported monthly through NHS administrative systems, with automated validation checks ensuring accuracy and consistency. The population data is taken from population estimates as those registered with the NHS.

Beds per 1,000 Population (OECD 2024a): This metric, provided by the OECD, uses data from NHS Digital, Public Health Scotland, and national agencies across the UK. It tracks the availability of inpatient beds. It provides annual averages of beds available overnight in public hospitals. Includes acute care and psychiatric beds but excludes private sector facilities for consistency. Data are for financial years and represent publicly funded healthcare infrastructure. The number of beds data is gathered through automated administrative streams from NHS hospitals and care facilities. The population data is taken from population estimates as those registered with the NHS.

Type 1 Major Care Attendances (NHS 2024): These figures represent percentages indexed to the baseline year 2011, providing a normalized measure of demand changes over time. The NHS employs administrative records to capture real-time data, validated through internal

processes to mitigate potential inconsistencies or reporting delays. However, some unreported data at the trust level (e.g., support facilities and non-inpatient services) are excluded, which may lead to minor underestimations. Type 1 Major Care is defined as major emergency departments that provide a consultant-led 24-hour service with full facilities for resuscitating patients (Fund 2024b).

Waiting Times for Key Medical Procedures (England 2024): Waiting time data are sourced directly from NHS England's Referral to Treatment (RTT) waiting times statistics, as published on NHS England's statistics portal. These statistics measure the time elapsed between a patient's referral by a general practitioner (GP) and the initiation of treatment, providing insights into NHS performance. Waiting times are calculated as the median duration (in weeks) for patients to begin their procedures, ensuring consistency across treatment categories. The dataset includes procedure-specific insights for key medical interventions such as general surgery, urology, and trauma & orthopedics, among others, enabling cross-procedure comparisons. Data are collected monthly through NHS administrative systems and validated using automated checks and standardized reporting protocols to ensure accuracy and consistency.

Overall: All datasets rely on administrative data validated by their respective agencies, minimizing reporting errors. While historical methodological changes (e.g., hospital reporting standards) necessitated adjustments, these do not affect the study's time period from 2015 to 2022. Similar datasets from the World Health Organization (WHO) and Eurostat health statistics could offer supplementary insights but were not utilized due to a lack of UK-specific detail.

2.3 Methodology

The data cleaning method involved many steps and is described in Section .1.2. All data cleaning was done using `arrow` (Richardson et al. 2024), `tidyverse` (Wickham et al. 2019), `readxl` (Wickham and Bryan 2023), and `stringr` (Wickham 2023). Below is the cleaned data which will be used for this paper. Table 4 provides the data including individual treatment area percentage change in waiting times from 2015 in table format.

Table 1: NHS Hospital Statistics from the cleaned healthcare data for 2015-2022

Year	beds_per_1000	physicians_per_1000	attendance	Total
2015	2.61	2.77	0	9.34
2016	2.57	2.78	3	9.88
2017	2.54	2.81	3	10.10
2018	2.50	2.84	6	10.22
2019	2.45	2.95	6	10.37
2020	2.43	3.03	-19	9.68
2021	2.43	3.19	9	10.21
2022	2.45	3.19	9	12.07

Table 1 presents United Kingdom hospital and healthcare statistics from the cleaned dataset, consisting of the 3 predictor variables and estimand and 8 observations spanning the years 2015 to 2022. The predictors include year, beds per 1,000 inhabitants, physicians per 1,000 inhabitants, and Type 1 major care attendance rates. The outcome variable `Total` captures wait times (median number of weeks per year) averaged over key treatment areas such as cardiology, ENT, gynaecology, plastic surgery and neurosurgery. All metrics are standardized based on population or healthcare utilization data for each year. The dataset including individual treatment waiting times in weeks is provided in Table 4 and the full list of treatments from which `Total` is computed is provided in Table 3.
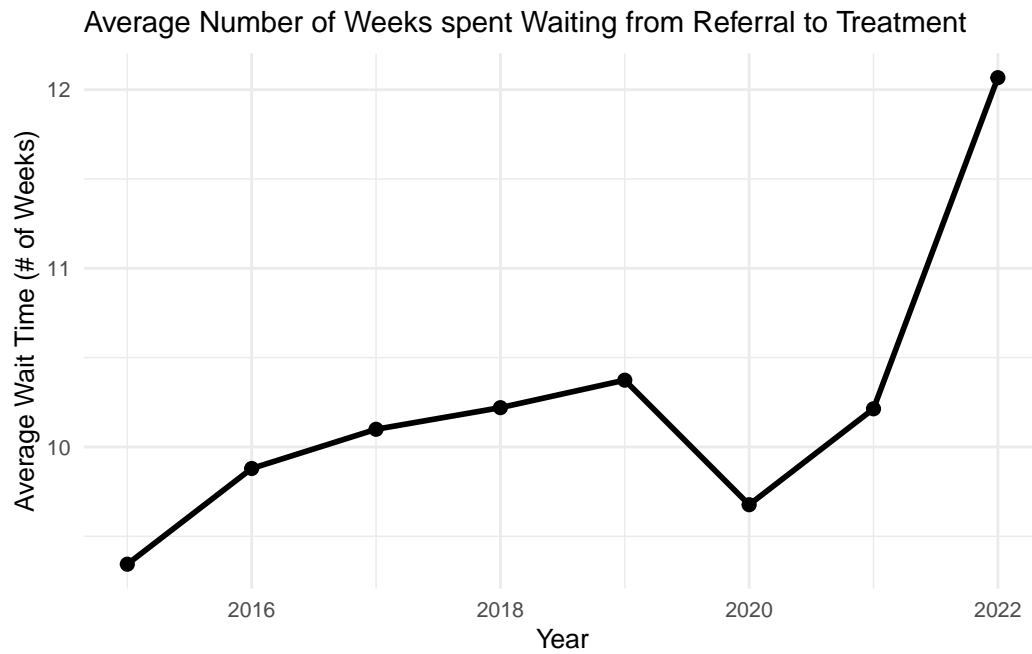
2.4 Outcome variables

Figure 1

Figure 1 shows the percentage increase in total wait times for key medical procedures in the UK relative to 2015, revealing a steady rise through 2019, followed by fluctuations during the COVID-19 pandemic. The moderate growth between 2015 and 2019 suggests manageable strain on healthcare resources, but the dip in 2020 reflects reduced elective procedures as hospitals prioritized emergency care. This was followed by a sharp increase in 2021 and a steep surge in 2022, likely due to the backlog created during the pandemic, ongoing demand pressures, and resource constraints. Overall, wait times have risen close to 30% since 2015, further emphasizing the need for targeted interventions to improve healthcare accessibility and efficiency.

2.5 Predictor variables
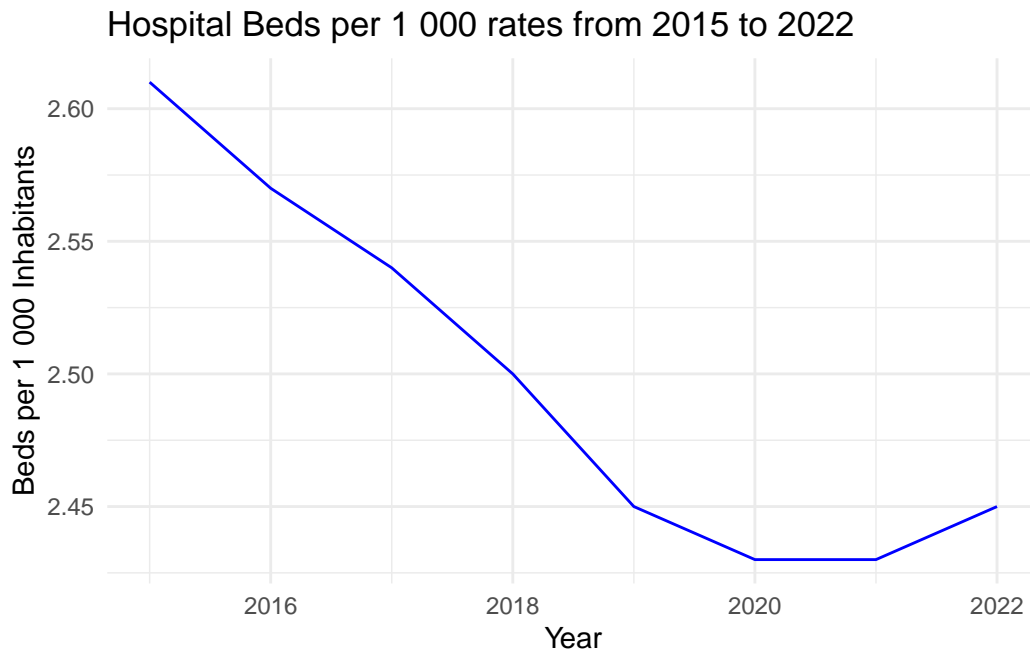
2.5.1 Hospital Beds per 1,000

Figure 2: United Kingdom's Hospital Beds per 1000 Inhabitants from 2015 to 2022

Figure 2 shows a steady decline in hospital beds per 1,000 inhabitants from 2015 to 2020, decreasing from approximately 2.6 to 2.45 beds. This consistent downward trend indicates a reduction in bed capacity relative to population growth, reflecting potential healthcare system pressures such as resource reallocation, efficiency measures, or underinvestment in infrastructure. However, a slight rebound is observed in 2021 and 2022, with the number of beds per 1,000 inhabitants stabilizing at just above 2.45. The trend underscores the importance of evaluating the impact of long-term capacity changes on patient care and exploring strategies to ensure sufficient resources to meet population needs.

2.5.2 Physicians per 1,000
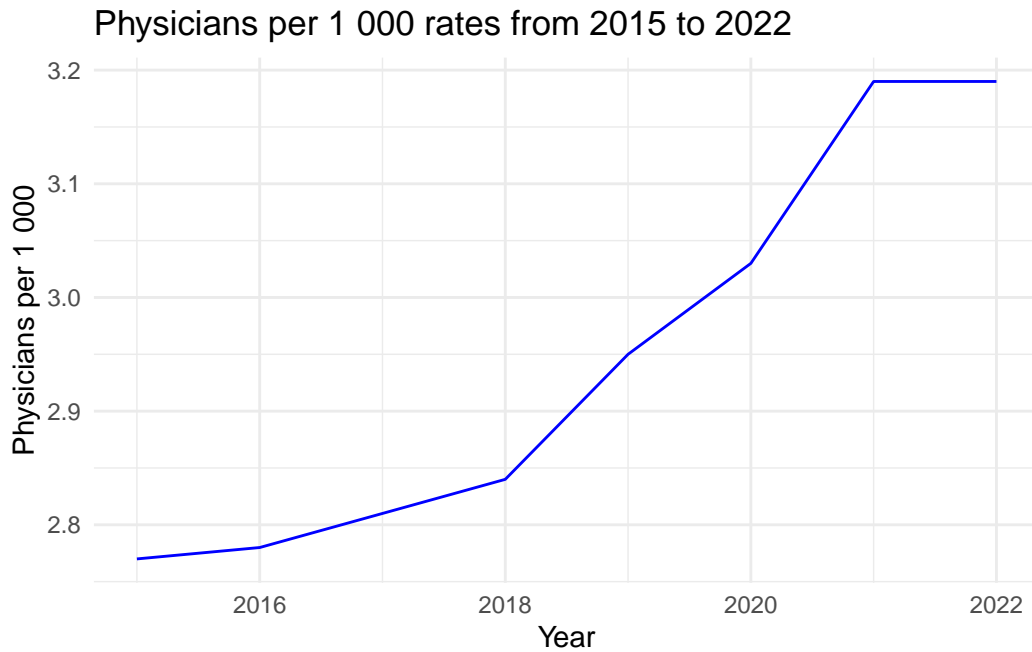
Physicians per 1 000 rates from 2015 to 2022

Figure 3: United Kingdom's Physician rate per 1000 Inhabitants from 2015 to 2022

Figure 3 shows a steady increase in the number of physicians per 1,000 inhabitants from 2015 to 2022, rising from approximately 2.8 to 3.2. This consistent growth reflects an improvement in physician availability relative to population size, likely driven by increased investment in healthcare workforce development, recruitment efforts, and possibly adjustments in healthcare policies. The marked acceleration between 2019 and 2021 may reflect efforts to bolster healthcare capacity during the COVID-19 pandemic. While this trend highlights progress in enhancing workforce availability, further analysis is needed to assess whether these increases adequately meet rising healthcare demands and address regional disparities in physician distribution.

2.5.3 Demand for Services

Type 1 Major Care Attendance in the United Kingdom (2015–2022)
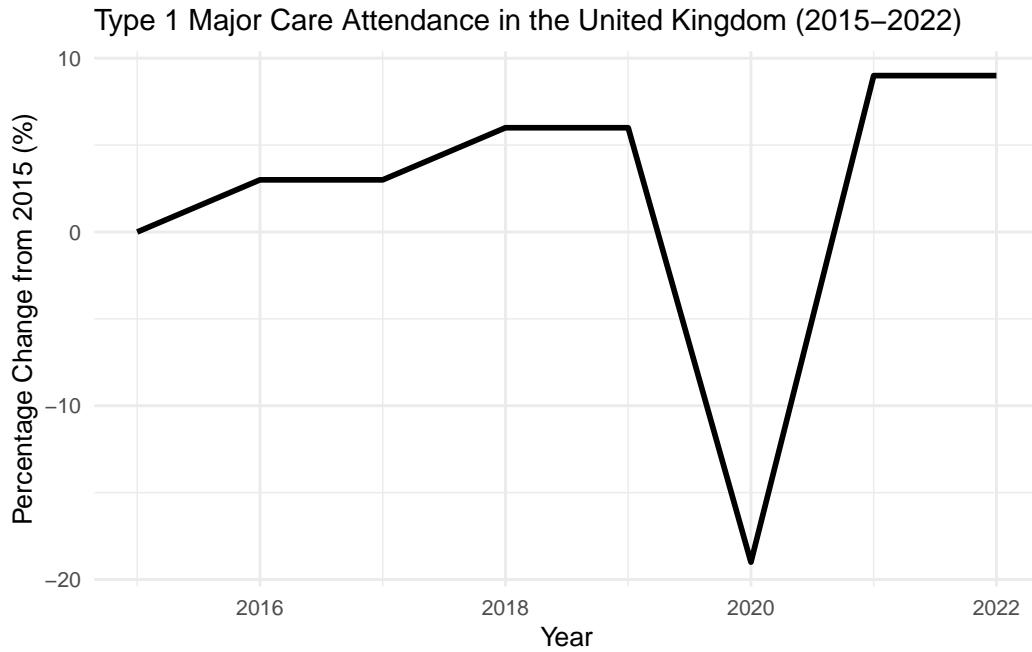
Figure 4: Type 1 Major Care Attendance in the United Kingdom from 2015 to 2022

Figure 4 illustrates the percentage change in Type 1 Major Care Attendance (A&E Attendances) in the UK from 2015 to 2022. Between 2015 and 2019, A&E attendances showed a modest, steady increase, reflecting a growing demand for emergency services. However, there was a sharp decline in 2020, likely attributable to the COVID-19 pandemic and related restrictions, which disrupted normal healthcare-seeking behavior. Attendance levels rebounded significantly in 2021, surpassing pre-pandemic levels, and remained stable through 2022. This trend highlights both the resilience of demand for emergency services and the need for strategies to manage surges in attendance effectively, such as enhancing access to primary care and preventive healthcare services.

3 Model

After conducting exploratory analysis on the dataset, linear trends were observed in each of the predictors and the outcome variable for years 2015 to 2019. Years 2020 to 2022 exhibited significant deviations from the previous years' linear trend, likely due to the Covid19 pandemic. Thus years 2015 to 2019 were thought to be the best to model. Linear relationships were observed between healthcare system factors (physicians per 1,000 people, hospital beds per 1,000 people, and Type 1 attendance rates) and wait times for medical procedures. These variables indicate potential predictive power, suggesting a linear relationship. To further investigate and quantify how these factors influence wait times, we implemented a multivariate linear regression model.

The goal of this linear regression model is to estimate the coefficients $\beta_0, \beta_1, \beta_3$ such that the model fits the data well and provides insights into how each predictor contributes to wait times for medical procedures. Additionally, this model will enable predictions of wait times under different healthcare conditions.

The statistical significance of each β_k coefficient will be assessed using a t-test, testing whether the coefficient is significantly different from zero. If the p-value for a coefficient is less than a chosen significance level, we can conclude that the corresponding predictor has a significant effect on wait times.

The model was run in `R` (R Core Team 2023) using the `modelsummary` package (Arel-Bundock 2022).

Model Instantiation

The final model is displayed here:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \epsilon_i$$

- Y_i: The median number of weeks averaged over the 15 specific medical procedures in year i.
- X_{1i}, X_{2i}, X_{3i}: The independent variables representing physicians per 1,000 people, beds per 1,000 people, and Type 1 attendance, respectively, for procedure j in year i.
- β_0: The constant term, representing the expected wait time when all predictors are zero.
- $\beta_1, \beta_2, \beta_3$: The slope coefficients for the predictors, representing the estimated change in wait time for a one-unit increase in the respective predictor.
- ϵ_i: The error term, representing the deviation of the actual wait time from the predicted wait time based on the regression equation.

The goal of the linear regression model is to estimate the values of β_1, β_2, and β_3 such that the model fits the data well and to predict the expected value of the average total waiting time for procedures for different values of physician rates, hospital bed rates, and type 1 major care attendance. The statistical significance of 1 can be assessed using a t-test, which tests

whether the estimated coefficient is significantly different from zero. If the p-value of the t-test for each coefficient is less than a chosen significance level, we can conclude that there is a significant relationship between the average percentage change in wait time and the healthcare predictors.

Model Assumptions

The linear regression model assumes:

- Linearity: The relationship between the predictors and response variable is linear.
- Independence: Observations are independent of each other.
- Homoscedasticity: The variance of the error term (ϵ) is constant across all levels of the predictors.
- Normality: The residuals of the model are approximately normally distributed.

These are checked in Section .2.1.

Model justification

This model framework allows us to evaluate the impact of key healthcare system factors on wait times. The selected predictors *beds_per_*1000, *physicians_per_*1000, and *attendance* are directly linked to resource availability and demand, making them relevant for understanding delays in medical care. The results from this model can provide actionable insights to inform policies aimed at optimizing resource allocation, improving healthcare infrastructure, and managing patient flow to reduce wait times. This analysis will help identify key leverage points for optimizing resource allocation and reducing wait times across various medical procedures.

Hypotheses

Each predictor variable is expected to contribute uniquely:

- Beds per 1,000 inhabitants (X_1): A lower bed-to-population ratio may increase wait times by limiting inpatient care capacity.
- Physicians per 1,000 inhabitants (X_2): Greater physician availability is hypothesized to decrease wait times.
- Attendance Rates (X_3): Increased attendance could indicate growing demand on the healthcare system, potentially leading to longer wait times.

Evaluation and Validation

To ensure the robustness of the findings, the model was validated through the following steps:

1. Residual Diagnostics: residual plots were assessed for patterns suggesting non-linearity or heteroscedasticity. No significant violations of assumptions were detected Section .2.1.
2. Significance Testing: t-tests were conducted to assess whether each coefficient (β_0, β_1, β_2, β_3) is statistically significant at the 0.05 level Section 4.
3. Goodness of Fit: The R-squared statistic was calculated to measure how well the model explains the variability in wait times Section 4.
4. Alternative Models: a variant containing data from 2015 to 2022 was tested in Section .2.2 to confirm whether the anomalies present in the Covid-19 pandemic years of 2020-2021 was a probable cause for poor fit. This model had much smaller R^2 and adjusted R^2 values and a much larger RMSE. The observation size differences between the two models were not significant enough to account for the much poorer fit, thus the smaller model was kept for final analysis.

Table 2: Linear model explaining wait times (median number of weeks per year) based on healthcare system predictors (2015-2019)

	Multivariate Linear Model
(Intercept)	73.255
	(0.355)
beds_per_1000	−16.826
	(0.387)
physicians_per_1000	−7.196
	(0.423)
attendance	−0.075
	(0.714)
Num.Obs.	5
R2	0.966
R2 Adj.	0.866
AIC	−3.1
BIC	−5.0
Log.Lik.	6.527
RMSE	0.07

4 Results

The linear regression model, summarized in Table 2, evaluates the impact of healthcare system factors on wait times from 2015–2019, revealing that hospital beds per 1,000 inhabitants ($\beta_1 = -16.826$) and physicians per 1,000 inhabitants ($\beta_2 = -7.196$) are negatively associated with wait times, indicating that increased resources may reduce delays however the findings are statistically insignificant ($p > 0.05$), highlighting the limitations imposed by the small sample size ($n = 5$). Attendance rates ($\beta_3 = -0.075$) exhibit a negligible effect, with a low p-value (0.714) further suggesting minimal impact on wait times when controlling for other factors. The model explains 96.6% of the variability in wait times ($R^2 = 0.966$), but the adjusted R^2 of 0.866 points to potential overfitting, as the inclusion of weak predictors may inflate the apparent fit. A small root mean squared error (RMSE) of approximately 0.07 indicates that the model's predictions closely align with observed values, yet this performance may not generalize beyond the dataset. Despite the excellent overall fit, the lack of significant predictors underscores the limitations of the analysis, with the small sample size undermining the precision of coefficient estimates and the robustness of conclusions. Furthermore, the variability in predictor effects suggests the potential presence of omitted variables or non-linear interactions.
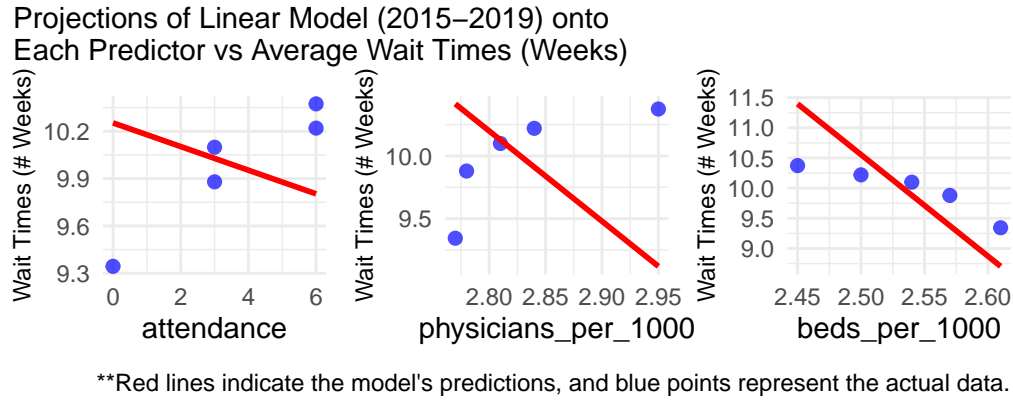
Projections of Linear Model (2015–2019) onto
Each Predictor vs Average Wait Times (Weeks)

**Red lines indicate the model's predictions, and blue points represent the actual data.

Figure 5

The projection plots in Figure 5 illustrate the linear model's predictions for average wait times in relation to the three predictors: attendance, physicians per 1,000, and beds per 1,000. Red lines represent the predicted relationships, while blue points reflect observed data. For attendance, the model suggests a negative relationship, but considerable variability and notable deviations hint at unmodeled interactions or confounding factors. For physicians per 1,000, the negative relationship is clearer, with a steeper downward slope, though deviations at lower densities reveal limitations in the model's assumptions. The strongest relationship is seen for beds per 1,000, where a consistent downward slope indicates a significant association between increased bed availability and reduced wait times, with observed data aligning closely to this trend.

However, the small sample size (n = 5) limits the reliability and generalizability of these findings. While the 2015–2019 data avoids the confounding effects of COVID-19, the limited dataset hampers definitive conclusions, and predictor coefficients remain non-significant. Variability in predictions for attendance and physicians per 1,000 suggests potential missing interactions or non-linear dynamics. Overall, the model better captures the relationship between predictors and wait times than the one analyzed in Section .2.2.

15

5 Discussion

5.1 Overview

This study investigates the relationships between healthcare infrastructure, workforce availability, and patient attendance rates with average wait times for 15 medical sectors in the United Kingdom from 2015 to 2022. Our analysis employs a linear regression model but finds no statistically significant relationships between the predictors and wait times. While hospital beds per 1,000 inhabitants showed a weak positive association with wait times, the relationship was not significant, suggesting that bed capacity alone may not directly mitigate delays. Similarly, physician density and attendance rates demonstrated positive but insignificant associations with wait times, reflecting the complex interplay of factors influencing healthcare delays. These findings indicate that systemic inefficiencies, resource allocation challenges, and unmeasured variables may play a more significant role than initially hypothesized. The limited explanatory power of the model and the presence of anomalies in the data post-2019 highlight the need for more detailed data and refined analyses to uncover the true drivers of healthcare delays.

5.1.1 Relationship between Wait Times and Beds per 1,000

The weak positive association between hospital beds per 1,000 inhabitants and wait times was unexpected and not statistically significant. While prior research emphasizes the importance of bed availability in managing healthcare demands, this study suggests that increasing bed numbers alone may not guarantee reduced delays. The minimal effect size indicates that factors such as bed utilization efficiency, patient flow management, and the quality of care delivery could overshadow the raw number of beds available. For instance, hospitals with higher bed capacity but poor discharge planning or limited staffing may still experience significant delays. These findings underscore the importance of complementing infrastructure investments with systemic reforms, such as optimizing bed turnover rates, enhancing care coordination, and adopting alternative care models like outpatient procedures or home-based care.

5.1.2 Relationship between Wait Times and Physicians per 1,000

The negative but statistically insignificant relationship between physician density and wait times challenges the expectation that greater workforce availability reduces delays. This result may reflect systemic inefficiencies or increased demand driven by greater access to healthcare services. For example, higher physician availability can lead to more diagnostic and treatment referrals, overwhelming other parts of the healthcare system, such as specialized care or operating rooms. Additionally, unmeasured factors like disparities in physician specialization, workflow inefficiencies, or misalignment between physician density and patient needs could further obscure the relationship. These findings highlight that increasing physician numbers

alone is insufficient to reduce wait times without addressing systemic bottlenecks. Interventions such as improving care pathways, enhancing interdisciplinary collaboration, and aligning physician availability with infrastructure and support staff could better leverage workforce capacity to reduce delays.

5.1.3 Relationship between Wait Times and Type 1 Major Attendance

The weak negative association between attendance rates and wait times, though not significant, aligns with broader concerns about rising patient loads straining healthcare systems. The relationship is likely weak due to the significant deviation from pre-Covid trend in 2020, with a sudden drop by 25% from 2019 levels. This drop was recovered in 2021, but the impact on the linear model constructed by a limited 8 observation data set cannot be ignored. As attendance increases, hospitals may prioritize emergency cases, diverting resources from elective and non-urgent procedures, potentially exacerbating delays. Factors such as limited access to primary care, changes in population demographics, and healthcare-seeking behaviors likely drive the observed trends. Although the relationship was not statistically significant, addressing the root causes of rising attendance remains critical to reducing systemic strain. Potential strategies include expanding primary care access, implementing community triaging systems, and improving public health education to promote appropriate healthcare utilization. These measures could help balance patient loads and reduce delays for non-emergency procedures.

5.2 Limitations

The primary limitation of this study is the small sample size ($n = 5$ for the focused 2015–2019 model), which weakens the statistical power and limits the generalizability of the findings. The dataset also includes anomalies introduced by the COVID-19 pandemic, with wait times deviating sharply from the pre-pandemic trend in 2020 and 2021. These anomalies introduce inconsistent variance in the data, further reducing the reliability of the adapted linear model in Section .2.2, which assumes consistent patterns across the dataset. As illustrated in Figure 1, wait times followed a consistent linear increase from 2015 to 2019, but 2020 deviated sharply below trend, while 2021 exhibited a sharp increase above trend. These anomalies suggest inconsistent variance in the data and indicate that a linear model may not be appropriate for post-2019 data. Moreover, the dataset lacks information on key confounding variables such as regional healthcare policies, socioeconomic disparities, or detailed patient demographics, which could significantly influence wait times. Unmeasured systemic inefficiencies, such as staffing misalignment, operating room availability, or referral practices, may also account for the insignificant results. Furthermore, the linear model does not account for potential interaction effects or non-linear relationships between predictors, which could offer deeper insights into the dynamics of healthcare delays.

5.3 Future Work

To address the limitations of this study, future research should prioritize collecting monthly data for the years 2015 to 2019 to explore more granular trends before the anomalies introduced by COVID-19. Section 4 shows the potential for this further investigation with a model consisting of the yearly data from 2015 to 2019 exhibiting a much stronger R^2 value of 0.966 than the model in Section .2.2. This model is small (5 observations), thus a more extensive model involving the monthly data would be necessary. Analyzing these pre-pandemic patterns could provide a clearer understanding of the role that metrics such as beds per 1,000, physician density, and attendance rates played during a relatively stable period. Additionally, expanding the dataset to include broader healthcare metrics, such as regional funding levels, technological investments, and patient demographics, could yield a more comprehensive analysis. Future studies should also explore non-linear models or interaction effects to capture the intricate relationships between predictors. Simulation models and scenario analyses could be used to test policy interventions, such as optimizing resource allocation or improving workflow efficiency, to identify actionable strategies for reducing wait times and improving healthcare delivery.

Appendix

.1 Analysis Dataset

.1.1 Treatments

Table 3: A summary table of the treatments

Treatments
Cardiology
Cardiothoracic
Dermatology
ENT
Gastroenterology
General_S
General_M
Geriatric
Gynaecology
Neurology
Neurosurgery
Ophthalmology
Oral
Plastic
Rheumatology
Thoracic
Trauma
Urology

.1.2 Data Cleaning

The data cleaning process undertaken in this project involved several steps to ensure the dataset was accurate, standardized, and ready for analysis. First, raw data on healthcare waiting times, physicians, hospital beds, and emergency admissions from the OECD and NHS were loaded. For the waiting times data, all Excel files were processed by standardizing column names, extracting relevant fields (treatment function and average waiting time), and appending metadata such as the year and month derived from file names. Treatment functions were harmonized across years, with those from April 2021 onward mapped to their pre-2021 equivalents using a predefined renaming scheme. Data with irrelevant or missing treatment functions (e.g., "Other") were excluded.

Yearly averages for each treatment function were calculated by grouping the data by year and function and computing the mean waiting time. The data was then pivoted to make years the rows and treatment functions the columns, providing a wide-format structure suitable for further analysis.

All datasets were merged on the year variable to produce a consolidated dataset. Percent changes in attendance rates from 2015 levels were computed to highlight trends over this study's time as opposed to the original percent changes from 2011 data. Lastly, columns with multiple-word treatment names were shortened to their first word, with duplicates (e.g., "General Surgery" and "General Medicine") resolved by appending unique suffixes.

.1.3 Individual Treatment Wait Times

Table 4: A summary table of the cleaned healthcare data displayed in parts

Year	beds_per_1000	physicians_per_1000	attendance	Cardiology
2015	2.61	2.77	0	7.34
2016	2.57	2.78	3	7.94
2017	2.54	2.81	3	8.02
2018	2.50	2.84	6	8.39
2019	2.45	2.95	6	8.91
2020	2.43	3.03	-19	10.60
2021	2.43	3.19	9	10.24
2022	2.45	3.19	9	11.29

Table 5: Part 2: Hospital statistics

Cardiothoracic	Dermatology	ENT	Gastroenterology	General_S	General_M
6.86	7.44	11.54	4.35	3.10	9.32
7.00	6.82	12.76	4.16	2.81	9.79
7.06	6.60	13.49	4.17	2.77	10.08
7.73	6.58	13.38	4.04	2.11	9.49
8.21	6.82	15.24	4.24	1.70	9.88
7.73	5.71	15.78	4.40	1.74	10.14
8.19	5.90	15.97	4.32	4.56	12.60
9.05	6.77	18.17	4.34	1.16	14.38

Table 6: Part 3: Hospital statistics

Geriatric	Gynaecology	Neurology	Neurosurgery	Ophthalmology	Oral
1.37	8.35	2.93	10.26	11.46	12.49
0.84	9.13	3.71	11.31	11.78	13.82
1.21	9.57	5.10	11.24	12.56	14.54
1.89	9.77	4.49	11.53	12.95	15.48
1.01	10.49	4.95	11.83	12.93	16.56
0.80	10.93	5.03	9.97	14.55	20.51
0.80	13.06	6.21	13.36	11.30	18.29
1.46	14.68	5.88	13.00	11.76	20.09

Table 7: Part 4: Hospital statistics

Plastic	Rheumatology	Thoracic	Trauma	Urology	Total
7.23	3.60	3.08	12.41	8.01	9.34
7.44	3.55	3.07	13.58	8.43	9.88
6.73	3.62	3.03	14.10	8.62	10.10
6.72	3.81	2.92	14.19	8.28	10.22
6.98	3.50	3.15	14.73	8.99	10.37
5.76	3.53	2.90	20.36	7.96	9.68
6.55	4.45	3.80	22.00	9.05	10.21
8.33	4.97	4.13	22.77	10.32	12.07

.2 Model Details

.2.1 Original Model Residuals
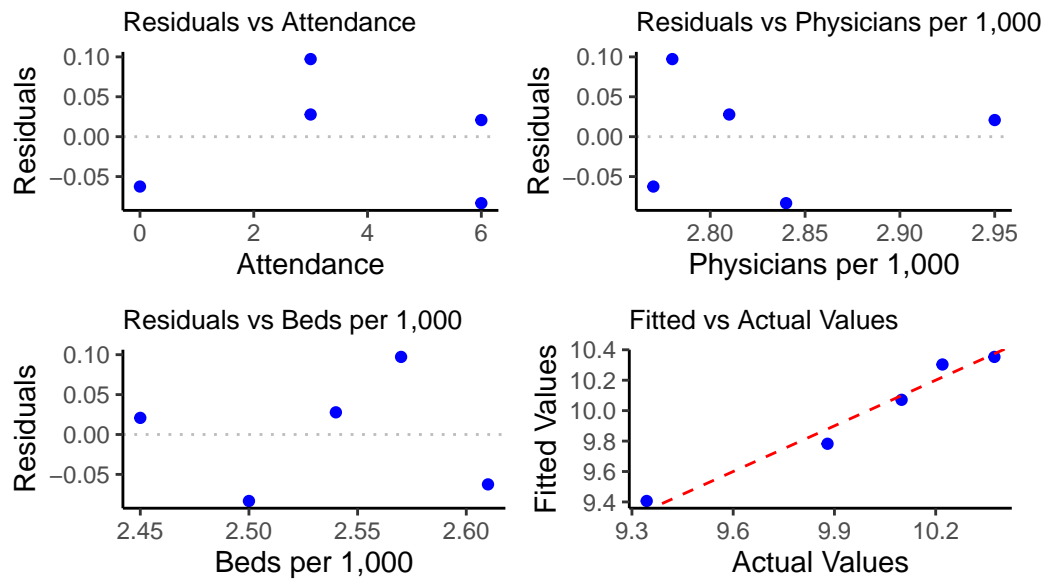
Diagnostic Plots for Linear Model (2015–2019)

Figure 6

Figure 6 residuals show no clear patterns against attendance, physicians per 1,000, or beds per 1,000, supporting the assumption of linearity. There are no violations against constant variance, unlike in Figure 7. The fitted values reasonably align with the actual values. However, the limited data points (5 observations) reduce the robustness of these conclusions.

Table 8: Linear model explaining change in wait times based on healthcare system predictors

	Frequentist Linear Model
(Intercept)	4.126
	(0.885)
beds_per_1000	−0.556
	(0.945)
physicians_per_1000	2.517
	(0.447)
attendance	0.040
	(0.259)
Num.Obs.	8
R2	0.593
R2 Adj.	0.287
AIC	21.1
BIC	21.5
Log.Lik.	−5.571
RMSE	0.49

.2.2 Adapted Model

The linear regression model, presented in Table 8, explores the relationships between healthcare system factors and the percentage change in wait times, though none of the variables demonstrate statistically significant effects, as indicated by p-values greater than 0.05. The number of hospital beds per 1,000 inhabitants shows a small negative effect (β_1=-0.556), suggesting a minimal decrease in wait times with higher bed availability, though this effect lacks statistical significance. Physician availability per 1,000 (β_2=2.517) indicates a larger positive association, yet it is also not statistically significant. Attendance rates (β_3 =0.40) exhibit the smallest positive effect, suggesting a weak association with wait times. The model achieves an R^2 value of 0.599, indicating that approximately 60% of the variability in wait times is explained by the model, though the adjusted R^2 value of 0.299 reflects limited predictive power. With a sample size of n=8, the findings should be interpreted cautiously, as the lack of significance and low adjusted R^2 highlight potential limitations in the model's robustness and generalizability.
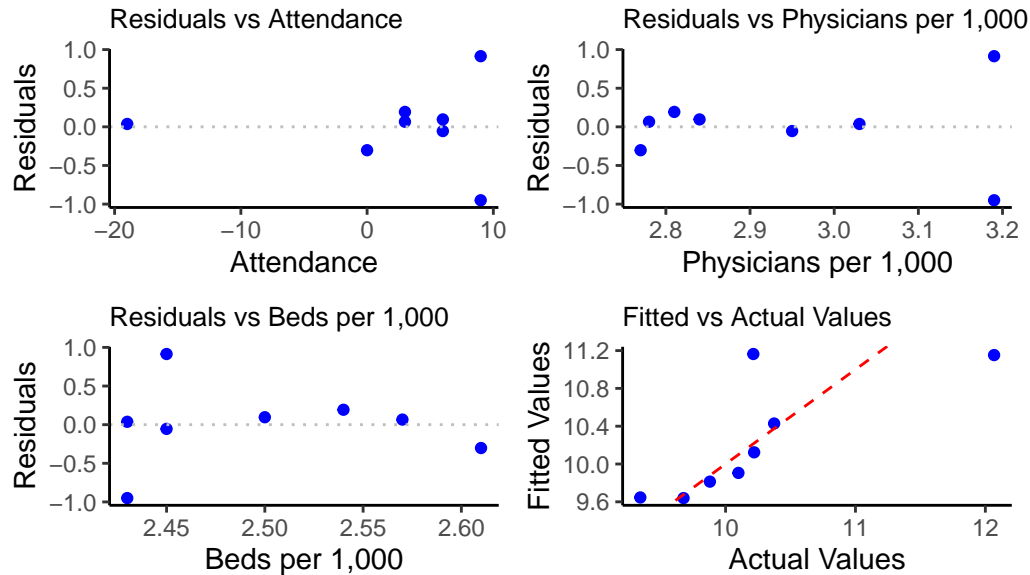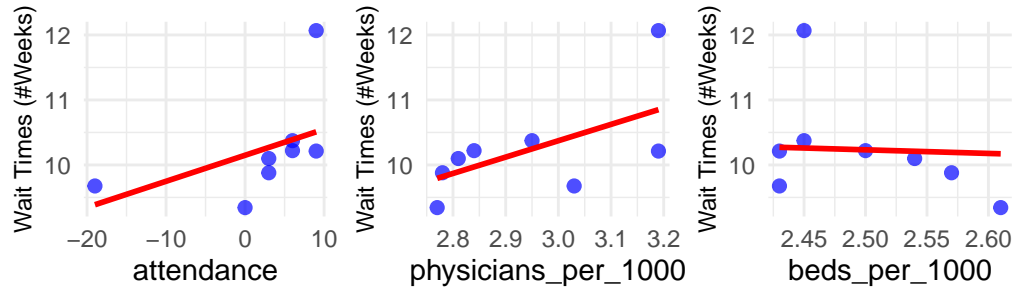
Figure 7

Figure 7 presents the residual plots which indicate that the linear model generally fits the data well, with residuals scattered randomly around zero for all predictors, suggesting no major violations of linearity or independence assumptions. However, in each residual plot there is a violation of constant variance where there are noticeable outliers and potential influential points, particularly in the plots for attendance and physicians per 1,000, which may affect the model's robustness. The fitted vs. actual values plot shows a good overall alignment with most points near the predicted line, but two significant deviations highlight discrepancies that could stem from these outliers. Given the small sample size, these findings should be interpreted with caution.

Projections of Linear Model onto
Each Predictor vs Average Wait Times Compared to 2015

**Red lines indicate the model's predictions, and blue points represent the actual data.

Figure 8

The projection plots in Figure 8 illustrate the linear model's predictions for the percentage change in wait times (`Total`) compared to the three predictors: `attendance`, `physicians_per_1000`, and `beds_per_1000`. The red lines depict the predicted relationships, while the blue points represent the observed data values. The `physicians_per_1000` predictor shows a clear positive relationship with wait times, as indicated by the upward-sloping red line, and most data points align reasonably well with the model's trend. Attendance also demonstrates a positive trend, but the variability in the observed data points around the red line suggests a weaker fit compared to attendance. In contrast, `beds_per_1000` shows minimal association with wait times, with the red line nearly flat and the data points scattered without a clear trend. For each plot of the predictor vs outcome variable there seems to be at least two data points which deviate significantly from the remaining points. These were hypothesized to be due to the outliers from trend witnessed in Figure 1 stemming from the Covid-19 pandemic years.

.3 Idealized Methodology

Limitations of Current NHS and OECD Data

The NHS and OECD datasets utilized in this study provide valuable insights into hospital wait times through aggregate statistics such as median waiting times, hospital beds per capita, and physician density. These metrics are essential for maintaining patient confidentiality and offering a broad overview of healthcare system performance. However, these datasets exhibit limitations. They fail to capture detailed patient demographics, regional healthcare disparities, or systemic inefficiencies that may drive the increasing RTT (Referral to Treatment) wait times. Moreover, they omit subjective factors such as patient satisfaction, perceived urgency, and the emotional toll of delayed care—elements that could provide a more comprehensive understanding of the healthcare experience.

Given this study's findings of no significant relationship between the predictors—hospital beds per 1,000, physicians per 1,000, and Type 1 Major Care Attendance Rates—and wait times, there is reason to explore more nuanced factors impacting delays. The absence of demographic data limits the ability to identify disparities in healthcare access and outcomes across different population groups. Furthermore, the lack of published records capturing patient perceptions of NHS wait times restricts the ability to evaluate how these delays are experienced by those affected.

Addressing these gaps would require integrating more granular data on demographic variables, regional healthcare infrastructure, and patient-reported outcomes. Such data would enable a deeper analysis of the underlying drivers of wait times and support the development of targeted, effective interventions to improve healthcare delivery and patient experiences during the waiting period.

Overview

To address the limitations of the current data, an idealized survey methodology is proposed to gather more granular and patient-centered data on hospital wait times and their systemic determinants. This methodology incorporates advanced sampling techniques, targeted recruitment strategies, and robust data validation measures to ensure a comprehensive and representative dataset. By focusing on both quantitative and qualitative aspects of healthcare delays, this approach aims to provide a deeper understanding of the factors influencing wait times and inform more effective policy interventions.

Sampling Method

A stratified multi-stage sampling design will be employed to ensure representative data collection across the UK. This method captures the diversity of experiences in different regions, healthcare facilities, and patient populations.

1. Stratification: The population will be divided into strata based on:

   - Region: Urban, suburban, and rural areas.

- Demographics: Age, income, and ethnicity.

For instance, one stratum might consist of elderly rural residents requiring orthopedic care, while another might focus on young urban adults accessing emergency services.

2. Primary Sampling Units (PSUs): Regions within each stratum will be selected as PSUs using probability proportional to size (PPS). For example, a densely populated area like London will have a higher chance of selection compared to a sparsely populated rural region.

3. Secondary Sampling Units (SSUs): Within each PSU, healthcare facilities (e.g., hospitals and clinics) will be chosen as SSUs. Facilities will be selected based on factors such as patient volume and type of services offered.

4. Patient Selection: Patients within selected facilities will be randomly sampled, ensuring a balanced representation of various demographic and clinical profiles. Quotas will be used to ensure sufficient representation of underserved groups, such as low-income or minority patients.

Recruitment Strategy

A multi-modal recruitment approach will maximize participation across diverse populations. Participants will be invited through:

- *In-hospital Invitations*: Patients visiting selected facilities will be invited to participate, either during or after their appointments.

- *Mail and Email*: Recruitment materials will be sent to patients identified through administrative records, with links to online surveys or options for telephone interviews.

- *Community Outreach*: Partnerships with local organizations, especially in underserved areas, will help reach populations that might otherwise be excluded.

- *Incentives*: Participants will receive small incentives, such as gift cards, to encourage participation and improve response rates.

Data Collection and Variables

The survey will collect both quantitative and qualitative data:

1. Demographic Information:

   - Age, gender, income, education level, and ethnicity.
   - Geographic location (postcode-level precision).

2. Healthcare Access and Experiences:

   - Type and urgency of medical procedures.

- Time elapsed between referral and treatment.

- Patient satisfaction with care received.

3. Systemic and External Factors:

   - Perceptions of healthcare resource adequacy (e.g., staffing, equipment).

   - Preferences for care delivery models (e.g., in-person vs. virtual consultations).

4. Perceived Impact of Delays:

   - Emotional, physical, and financial effects of waiting for treatment.

Data Validation

The collected data will undergo validation to ensure accuracy and reliability:

- Consistency Checks: Responses will be cross-validated with administrative records to ensure coherence (e.g., matching reported wait times with NHS data).

- Non-Response Bias Analysis: Patterns of non-response will be analyzed, and statistical adjustments will be applied to account for underrepresented groups.

- Pilot Testing: The survey will be tested on a small sample to refine questions, eliminate ambiguities, and improve response rates.

Simulation and Future Data Collection

To complement the survey data, simulations will be conducted to model the impact of hypothetical interventions (e.g., increasing bed availability, optimizing physician allocation) on wait times. Additionally, future data collection efforts will incorporate continuous monitoring through partnerships with NHS facilities, enabling real-time updates and trend analysis.

Conclusion

This idealized methodology addresses the critical gaps in existing data by integrating more nuanced information which cannot be collected by automated NHS portals, but rather surveying. The survey captures patient perspectives, systemic factors, and granular regional insights. By leveraging stratified sampling, a variety of recruitment strategies, and validation techniques, this approach can provide a more nuanced understanding of healthcare delays and patient perspectives in the UK. The findings will guide targeted interventions to improve healthcare access and efficiency, ensuring that resources are allocated where they are needed most.

The survey is found at this link: https://forms.gle/QJBeSUjxUejnSi3A9.

Copy of the Survey

NHS Patient Wait Time Experience Survey:

This survey aims to gather insights into patients' experiences with NHS referral-to-treatment (RTT) wait times, their perceptions of healthcare resources, and the impact of delays on their well-being. Your responses will help inform strategies to improve healthcare access and efficiency. Participation is voluntary, and all responses will remain confidential.

Contact Information: For questions regarding this survey or the methodology used, please reach out to:

- **Chenika Bukes**, University of Toronto. Email: chenika.bukes@mail.utoronto.ca

Section 1: Demographics

1. **What is your age?**

- 18-25

- 25-34

- 35-44

- 45-54

- 55-64

- 65+

2. **What sex were you assigned at birth, on your original birth certificate?**

- Male

- Female

3. **How do you currently describe yourself (mark all that apply)?**

- Female

- Male

- Transgender

- I use a different term [free-text]

4. **What is your annual household income?**

- Under £20,000

- £20,000-£39,999

- £40,000-£59,999

- £60,000-£79,999

- £80,000+

- Prefer not to say

5. **What is your ethnicity?**

- White

- Black or Black British

- Asian or Asian British

- Mixed/Multiple ethnic groups

- Other

- Prefer not to say

6. **What is your post code?**

Short answer text

Section 2: Healthcare Access

7. **Have you had to wait longer than 18 weeks from referral for a medical procedure in the past 12 months?**

- Yes

- No

8. **If yes, how long did you wait to be consulted after referral?**

- 18-24 Weeks

- 24-30 Weeks

- 30-36 Weeks

- 36-42 Weeks

- 42-48 Weeks

- 48-52 Weeks

- 52+ Weeks

9. **What type of medical procedure did you wait for?**

- General Surgery

- Orthopedics

- Cardiology

- ENT (Ear, Nose, Throat)

- Neurology

- Thoracic

- Rheumatology

- Plastic Surgery

- Geriatric

- Gynecology

- Opthamology

- Oral Surgery

- Cardiothoracic

- Neurosurgery

- Dermatology

- Gastroenterology

10. **If you waited for longer than 18 weeks for an appointment, were you contacted during the waiting period by an NHS representative for a check-up on your related health problem?**

- Yes

- No

Section 3: Impact of Wait Times

11. **Did the wait time affect your health or well-being?**

- Yes, significantly

- Yes, moderately

- No

12. **If yes, how were you affected?**

- Physical health deterioration

- Emotional distress or anxiety

- Financial burden

- Other...

13. **How do you feel about virtual consultations?**

- Very satisfied

- Somewhat satisfied

- Impartial

- Unsatisfied

```
- Depends on the medical procedure in discussion
```

```
- Other...
```

Section 4: Perceptions of Healthcare System

14. **How would you rate the adequacy of healthcare resources in your area?**

```
- Excellent
```

```
- Good
```

```
- Average
```

```
- Poor
```

15. **What do you think would most improve healthcare wait times?**

```
- More hospital beds
```

```
- More healthcare workers
```

```
- Better resource allocation
```

```
- Improved scheduling systems
```

```
- Policy changes
```

Section 5: Feedback and Suggestions

16. **What changes would you like to see in the NHS to reduce referral to treatment waiting times?**

17. **Do you have any suggestions to improve this survey?**

References

Arel-Bundock, Vincent. 2022. "modelsummary: Data and Model Summaries in R." *Journal of Statistical Software* 103 (1): 1–23. https://doi.org/10.18637/jss.v103.i01.

Auguie, Baptiste. 2017. *gridExtra: Miscellaneous Functions for "Grid" Graphics.* https://CRAN.R-project.org/package=gridExtra.

England, NHS. 2024. *Referral to Treatment (RTT) Waiting Times.* https://www.england.nhs.uk/statistics/statistical-work-areas/rtt-waiting-times/.

Fund, The King's. 2024a. "Key Facts and Figures about the NHS." *Policy, Finance, and Performance.* https://www.kingsfund.org.uk/insight-and-analysis/data-and-charts/key-facts-figures-nhs.

———. 2024b. "What's Going on with a&e Waiting Times?" *Journal of Open Source Software.* https://www.kingsfund.org.uk/insight-and-analysis/long-reads/whats-going-on-with-ae-waiting-times. .

Laura Salisbury, Jocelyn Catty, Lisa Baraitser. 2023. *A Waiting Crisis?* Lancet. https://pmc.ncbi.nlm.nih.gov/articles/PMC7614272/pdf/EMS170587.pdf.

NHS. 2024. *A&E Attendances and Emergency Admissions.* https://www.england.nhs.uk/statistics/statistical-work-areas/ae-waiting-times-and-activity/.

OECD. 2024a. *Hospital beds by sector.* https://data-explorer.oecd.org/vis?lc=en&tm=beds&pg=0&snb=11&df%5Bds%5D=dsDisseminateFinalDMZ&df%5Bid%5D=DSD_HEALTH_REAC_HOSP%40DF_BEDS_SECT&df%5Bag%5D=OECD.ELS.HD&df%5Bvs%5D=1.0&dq=..10P3HB.._T....&pd=2015%2C&to%5BTIME_PERIOD%5D=false.

———. 2024b. *OECD Data Explorer: Health.* https://data-explorer.oecd.org/?lc=en&pg=0&fs%5B0%5D=Topic%2C0%7CHealth%23HEA%23&fc=Topic&snb=80&bp=true.

———. 2024c. *Physicians.* https://data-explorer.oecd.org/vis?lc=en&tm=physicians&pg=0&snb=15&df%5Bds%5D=dsDisseminateFinalDMZ&df%5Bid%5D=DSD_HEALTH_EMP_REAC%40DF_PHYS&df%5Bag%5D=OECD.ELS.HD&df%5Bvs%5D=1.0&dq=.......P.&pd=2015%2C&to%5BTIME_PERIOD%5D=false.

Pedersen, Thomas Lin. 2024. *Patchwork: The Composer of Plots.* https://CRAN.R-project.org/package=patchwork.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Richardson, Neal, Ian Cook, Nic Crane, Dewey Dunnington, Romain François, Jonathan Keane, Dragoș Moldovan-Grünfeld, Jeroen Ooms, Jacob Wujciak-Jens, and Apache Arrow. 2024. *Arrow: Integration to 'Apache' 'Arrow'.* https://CRAN.R-project.org/package=arrow.

Scotland, NHS. 2023. *NHSScotland workforce data warehouse.* TURAS Data Intelligence. https://turasdata.nes.nhs.scot/about-our-data-and-reports/methods/nhsscotland-workforce-data-warehouse/.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis.* Springer-Verlag New York. https://ggplot2.tidyverse.org.

———. 2023. *Stringr: Simple, Consistent Wrappers for Common String Operations.* https:

//CRAN.R-project.org/package=stringr.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, and Jennifer Bryan. 2023. *Readxl: Read Excel Files.* https://CRAN.R-project.org/package=readxl.