# Decoding UK Hospital Wait Times: Systemic Challenges Beyond Resources*

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

Chenika Bukes

November 28, 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 Introduction** **2**

**2 Data** **3**

2.1 Overview . . . 3
2.2 Measurement . . . 4
2.3 Methodology . . . 6
2.4 Outcome variables . . . 7
2.5 Predictor variables . . . 8
2.5.1 Hospital Beds per 1,000 . . . 8
2.5.2 Physicians per 1,000 . . . 9

*Code and data are available at: https://github.com/chenikabukes/UnitedKingdomHealthcare.

2.5.3 Demand for Services . . . 10

**3 Model** **11**

3.1 Model set-up . . . 11

3.1.1 Model Assumptions . . . 12

3.1.2 Model justification . . . 12

3.1.3 Hypotheses . . . 12

3.1.4 Evaluation and Validation . . . 12

**4 Results** **14**

**5 Discussion** **16**

5.1 Overview . . . 16

5.1.1 Relationship between Wait Times and Beds per 1,000 . . . 16

5.1.2 Relationship between Wait Times and Physicians per 1,000 . . . 16

5.1.3 Relationship between Wait Times and Type 1 Major Attendance . . . 17

5.2 Limitations . . . 17

5.3 Future Work . . . 18

**Appendix** **19**

.1 Full Analysis Dataset . . . 19

.1.1 Treatments . . . 19

.1.2 Data Cleaning . . . 19

.1.3 Individual Treatment Wait Times . . . 20

.2 Model Details . . . 22

.2.1 Original Model Residuals . . . 22

.2.2 Adapted Model . . . 23

**References** **26**

# 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 average percentage change from 2015 in wait times for 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 a snapshot of the first five rows of the key data columns 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 | 0.00 |
| 2016 | 2.57 | 2.78 | 3 | 5.73 |
| 2017 | 2.54 | 2.81 | 3 | 8.08 |
| 2018 | 2.50 | 2.84 | 6 | 9.85 |
| 2019 | 2.45 | 2.95 | 6 | 11.03 |
| 2020 | 2.43 | 3.03 | -19 | 3.56 |
| 2021 | 2.43 | 3.19 | 9 | 9.31 |
| 2022 | 2.45 | 3.19 | 9 | 29.15 |

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 (in percent change from 2015 levels) for 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 percetentage change from 2015 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

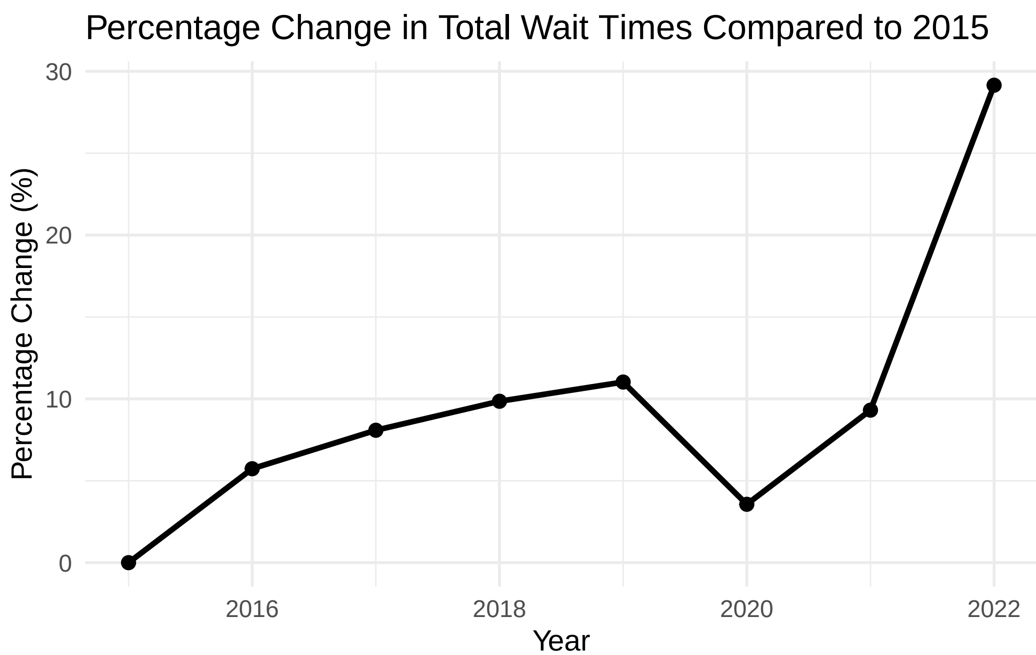


Figure 1

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. These trends underscore the urgent need for targeted healthcare interventions, including expanded capacity, streamlined workflows, and policies to address the post-pandemic backlog and increasing demand. 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

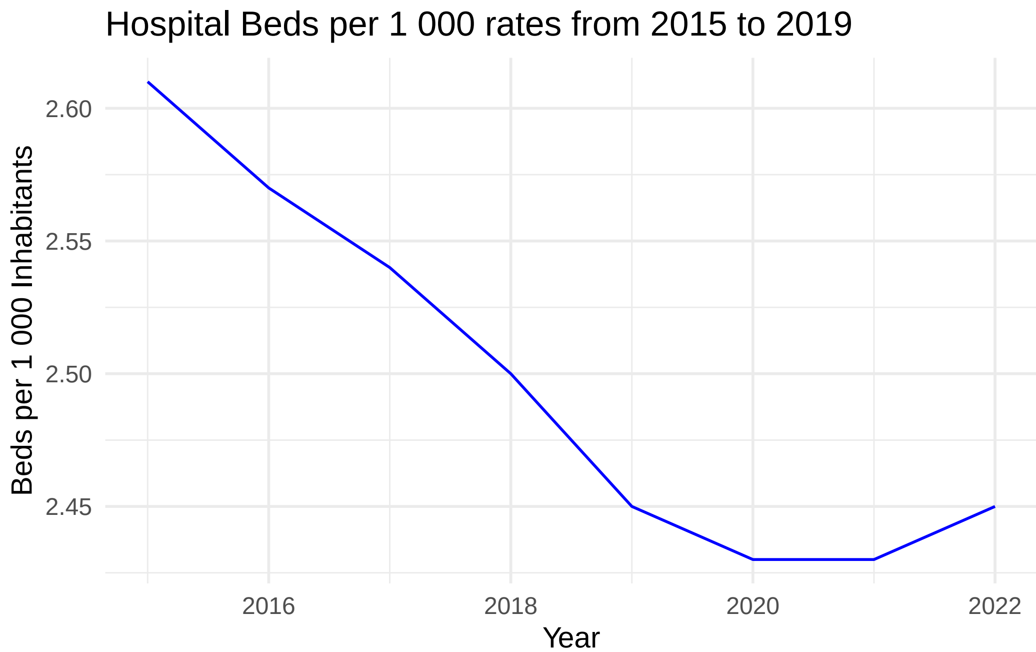


Figure 2: United Kingdom's Hospital Beds per 1000 Inhabitants from 2015 to 2019

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. This recent stabilization could suggest efforts to address capacity constraints or temporary adjustments in response to the increased demand for hospital resources during the COVID-19 pandemic. 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

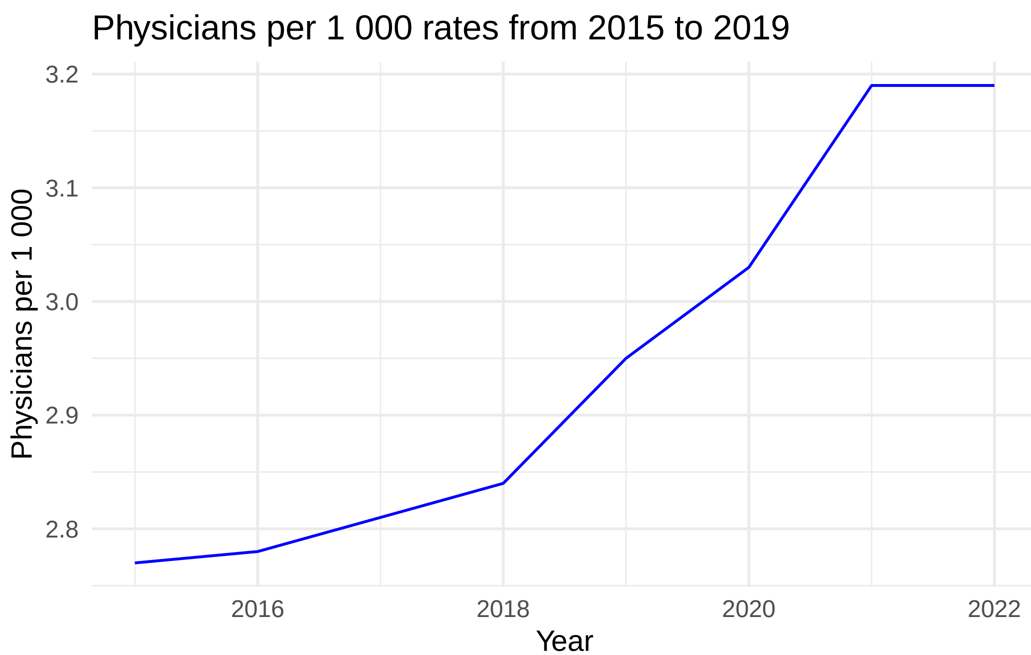


Figure 3: United Kingdom's Physician rate per 1000 Inhabitants from 2015 to 2019

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

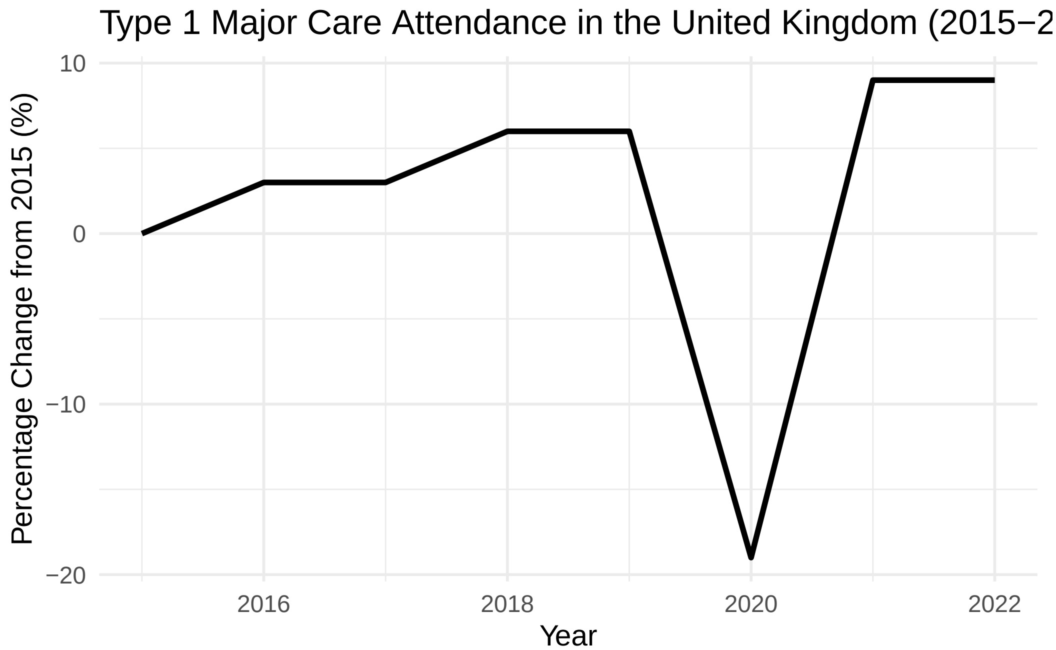


Figure 4: Type 1 Major Care Attendance in the United Kingdom from 2015 to 2019

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, we observed relationships between healthcare system factors (e.g., 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 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 $\beta_k$ 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).

## 3.1 Model set-up

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 average percentage change in wait time for 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 $i$.
- $\beta_0$: 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.
- $\epsilon_i$: 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 $\beta_1$, $\beta_2$, and $\beta_3$ 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.

### 3.1.1 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 ($\epsilon$) is constant across all levels of the predictors.
- Normality: The residuals of the model are approximately normally distributed.

### 3.1.2 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.

### 3.1.3 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.

### 3.1.4 Evaluation and Validation

Evaluation and Validation To ensure the robustness of our findings, the model was validated through the following steps:

1. Residual Diagnostics: We assessed residual plots for patterns suggesting non-linearity or heteroscedasticity. No significant violations of assumptions were detected Section .2.1.

2. Significance Testing: We conducted t-tests to assess whether each coefficient ($\beta_0$, $\beta_1$, $\beta_2$, $\beta_3$) 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 only from 2015 to 2019 was tested to confirm whether the anomalies present in the Covid-19 pandemic years of 2020-2021 was a probable cause for poor fit Section .2.2. In both models, the p-values of the coefficients were insignificant so the model containing the larger data set (original model from 2015-2022) was used for final analysis.

Table 2: Linear model explaining percentage increase in wait times based on healthcare system predictors

| | Frequentist Linear Model |
|---|---|
| (Intercept) | −97.108 |
| | (0.751) |
| beds_per_1000 | 4.714 |
| | (0.956) |
| physicians_per_1000 | 31.829 |
| | (0.374) |
| attendance | 0.396 |
| | (0.290) |
| Num.Obs. | 8 |
| R2 | 0.599 |
| R2 Adj. | 0.299 |
| AIC | 59.0 |
| BIC | 59.4 |
| Log.Lik. | −24.507 |
| RMSE | 5.18 |

# 4 Results

The linear regression model, presented in Table 2, 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 positive effect ($\beta_1$=4.71), suggesting a minimal increase in wait times with higher bed availability, though this effect lacks statistical significance. Physician availability per 1,000 ($\beta_2$=31.83) indicates a larger positive association, yet it is also not statistically significant. Attendance rates ($\beta_3$ =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.

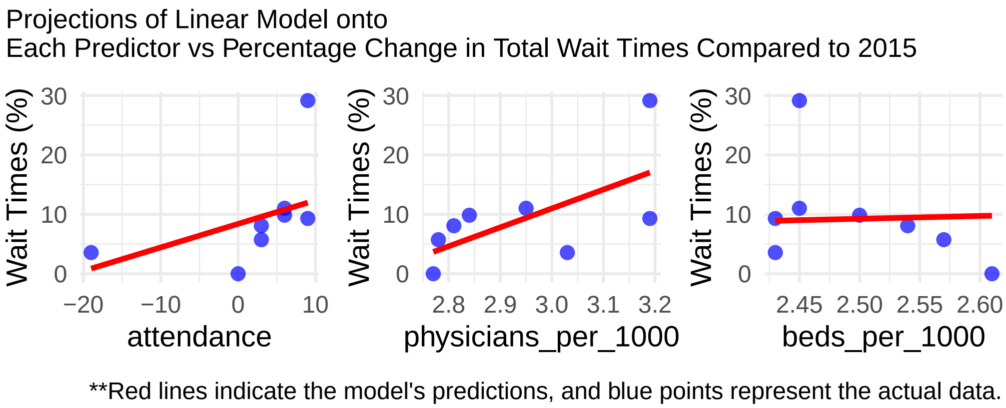


Figure 5

The projection plots in Table 2 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. Section .2.2 developed and analysed the significance of an adapted model for years 2015-2019 before the pandemic to avoid such outlier effects in capturing the trend. The overall model significance was much greater and linear assumption violations in the residual plots were non-existent, however individual coefficients were still non-significant and the dataset on which this model was constructed was even smaller with 5 observations.

# 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 positive 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 positive 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, which includes only eight observations spanning 2015 to 2022. This limits the generalizability of findings and weakens the statistical power needed to detect significant relationships. Additionally, the dataset is influenced by the effects of COVID-19, with wait times deviating significantly from the pre-pandemic trend. 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 .2.2 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 0.875. This model is small (5 observations), thus a more extensive model involving the monthly data would be beneficial. 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 Full 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 key 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 various sources, such as 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. Special corrections were applied, such as manually imputing specific values for ENT in 2019 and 2020.

All datasets were merged on the year variable to produce a consolidated dataset. Percent changes in waiting times from 2015 levels were computed for each treatment function to highlight trends over time. 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 | Cardiothoracic |
|---|---|---|---|---|---|
| 2015 | 2.61 | 2.77 | 0 | 0.00 | 0.00 |
| 2016 | 2.57 | 2.78 | 3 | 8.12 | 2.12 |
| 2017 | 2.54 | 2.81 | 3 | 9.23 | 2.93 |
| 2018 | 2.50 | 2.84 | 6 | 14.26 | 12.65 |
| 2019 | 2.45 | 2.95 | 6 | 21.45 | 19.62 |
| 2020 | 2.43 | 3.03 | -19 | 44.37 | 12.74 |
| 2021 | 2.43 | 3.19 | 9 | 39.53 | 19.41 |
| 2022 | 2.45 | 3.19 | 9 | 53.77 | 31.96 |

Table 5: Part 2: Hospital statistics

| Cardiothoracic | Dermatology | ENT | Gastroenterology | General_S | General_M | Geriatric |
|---|---|---|---|---|---|---|
| 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 2.12 | -8.32 | 10.53 | -4.27 | -9.28 | 5.01 | -39.07 |
| 2.93 | -11.25 | 16.86 | -4.11 | -10.65 | 8.20 | -11.98 |
| 12.65 | -11.51 | 15.92 | -7.07 | -31.78 | 1.88 | 37.53 |
| 19.62 | -8.30 | 29.53 | -2.39 | -45.14 | 6.01 | -26.88 |
| 12.74 | -23.31 | 31.83 | 1.36 | -43.68 | 8.84 | -41.59 |
| 19.41 | -20.64 | 38.39 | -0.69 | 47.38 | 35.18 | -41.46 |
| 31.96 | -9.02 | 57.39 | -0.14 | -62.55 | 54.25 | 6.40 |

Table 6: Part 3: Hospital statistics

| Geriatric | Gynaecology | Neurology | Neurosurgery | Ophthalmology | Oral | Plastic |
|---|---|---|---|---|---|---|
| 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| -39.07 | 9.33 | 26.63 | 10.18 | 2.77 | 10.62 | 2.85 |
| -11.98 | 14.70 | 73.96 | 9.52 | 9.61 | 16.38 | -6.88 |
| 37.53 | 17.03 | 53.13 | 12.37 | 13.02 | 23.91 | -7.06 |
| -26.88 | 25.63 | 68.87 | 15.20 | 12.82 | 32.55 | -3.46 |
| -41.59 | 30.95 | 71.58 | -2.83 | 26.98 | 64.20 | -20.40 |
| -41.46 | 56.49 | 111.64 | 30.18 | -1.35 | 46.43 | -9.39 |
| 6.40 | 75.82 | 100.57 | 26.68 | 2.60 | 60.82 | 15.14 |

Table 7: Part 4: Hospital statistics

| Plastic | Rheumatology | Thoracic | Trauma | Urology | Total |
|---|---|---|---|---|---|
| 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 2.85 | -1.21 | -0.21 | 9.50 | 5.34 | 5.73 |
| -6.88 | 0.73 | -1.49 | 13.66 | 7.69 | 8.08 |
| -7.06 | 6.00 | -4.93 | 14.40 | 3.37 | 9.85 |
| -3.46 | -2.65 | 2.51 | 18.70 | 12.31 | 11.03 |
| -20.40 | -1.73 | -5.84 | 64.11 | -0.61 | 3.56 |
| -9.39 | 23.80 | 23.37 | 77.36 | 13.02 | 9.31 |
| 15.14 | 38.32 | 34.21 | 83.52 | 28.91 | 29.15 |

## .2 Model Details

### .2.1 Original Model Residuals

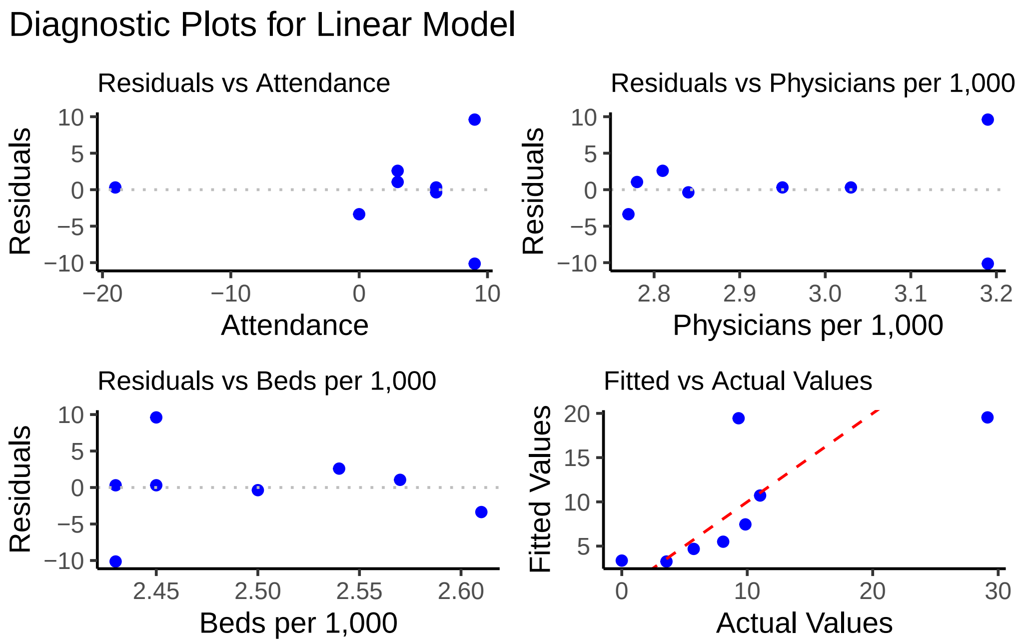


Figure 6

Figure 6 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.

Table 8: Linear model explaining percentage increase in wait times based on healthcare system predictors (2015-2019)

| | Frequentist Linear Model |
|---|---|
| (Intercept) | 679.074 |
| | (0.342) |
| beds_per_1000 | −176.991 |
| | (0.337) |
| physicians_per_1000 | −78.186 |
| | (0.362) |
| attendance | −0.660 |
| | (0.716) |
| Num.Obs. | 5 |
| R2 | 0.978 |
| R2 Adj. | 0.911 |
| AIC | 18.8 |
| BIC | 16.8 |
| Log.Lik. | −4.382 |
| RMSE | 0.58 |

### .2.2 Adapted Model

The linear regression model, summarized in Table 8, examines the impact of healthcare system factors on wait times from 2015–2019. Hospital beds per 1,000 inhabitants ($\beta_1 = -176.99$) and physicians per 1,000 ($\beta_2 = -78.19$) both show negative associations with wait times, suggesting increased resources reduce delays, though neither effect is statistically significant ($p > 0.05$). Attendance rates ($\beta_3 = -0.66$) exhibit a minimal effect. The model explains 97.8% of the variability in wait times ($R^2$=0.978), though the adjusted $R^2$ of 0.911 reflects possible overfitting due to the small sample size ($n = 5$). While the model fits well, the lack of significant predictors highlights the need for further data to better capture drivers of wait times. The small RMSE of 0.58 indicates the linear trend is never far off from the actual wait times.

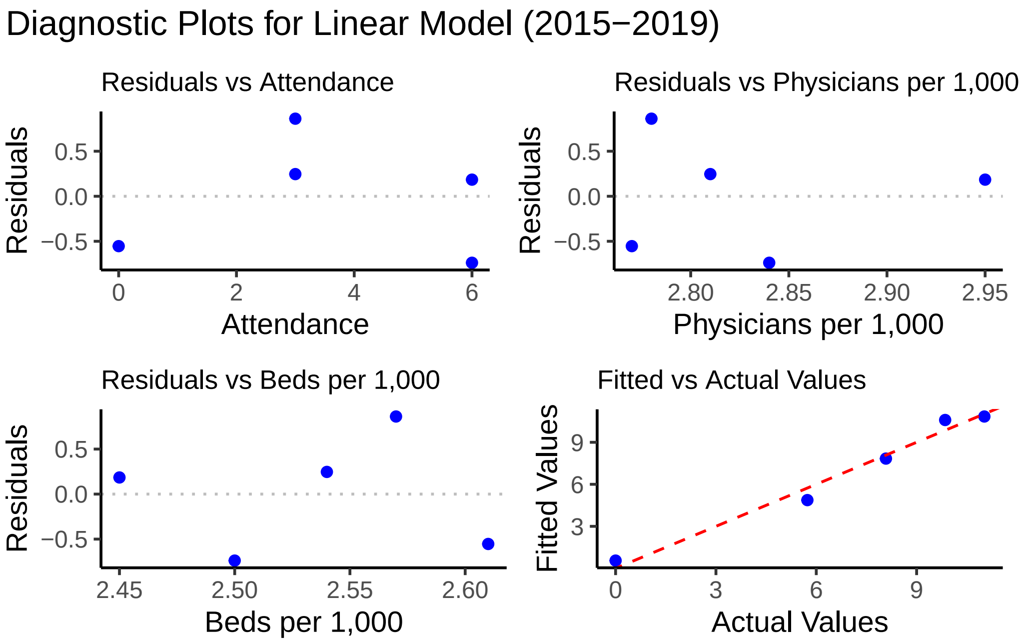


Figure 7

Figure 7 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 6. The fitted values reasonably align with the actual values. However, the limited data points (5 observations) reduce the robustness of these conclusions.

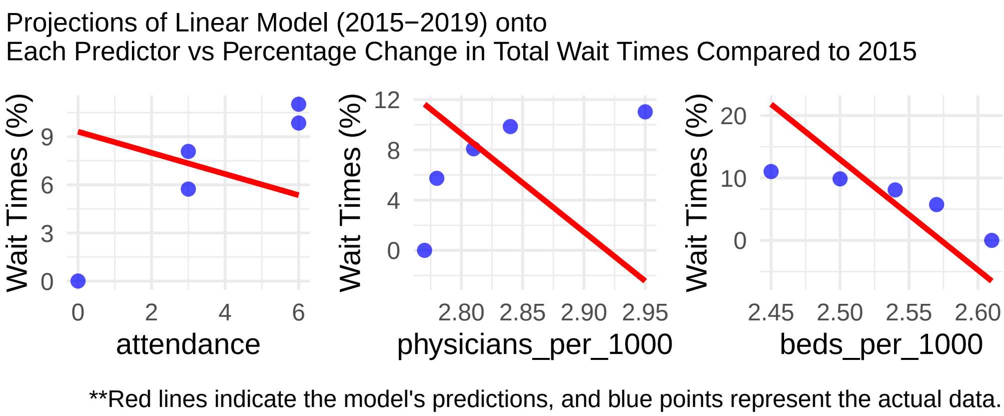


Figure 8

Overall, the linear model projections in Figure 8 effectively captures general trends, particularly for hospital bed availability. However, the variability around the red lines, particularly for attendance and physicians per 1,000, indicates potential model limitations or missing interactions between predictors. The small sample size (n = 5) further limits the reliability of these insights, necessitating caution when generalizing these findings. Additional data or non-linear modeling may better address these discrepancies and provide a more comprehensive understanding of the relationships between these predictors and wait times.

## 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.