

# ST344 Coursework

1719672

17/01/2020

## Executive summary

This report aims to answer 2 important questions about National Health Services in England. The first question is “Has demand on Coventry GPs (General practitioners) increased over the recent past?”, and the second is “Does the time between booking and appointment effect the missed appointment rates?”.

Firstly, we found that demand on Coventry GPs has no increased over the recent past. Rather, the number of registered patients and included practices has risen over the last 18 months, but the number of daily appointments has not risen proportionately.

We also found that patients are significantly more likely to not attend an appointment if it is in more than 1 days’ time. In particular, we found that attendance rates were highest for appointments the day after booking.

One problem which this report highlights is lack of funding, which is what is causing the decrease in appointments to registered patients. We recommend a funding increase of at least 5% for Coventry in order to keep up with the demand.

This report highlights the need to reduce amount of time patients need to wait before they are seen in order to reduce the economic burden of non-attendees. We are seeing a difference of at least 3% between being seen within 2 days and not, which is a lot of missed appointments when you carry that over to the millions of appointments which occur every year on the NHS. Again, additional funding may allow the practices to offer more appointments, and thus more on the same day and the next day.

## Introduction

This report aims to answer 2 important questions about National Health Services in England. The first question is “Has demand on Coventry GPs increased over the recent past?”, and the second is “Does the time between booking and appointment effect the missed appointment rates?”.

The data used in this report is from the NHS website (NHS 2020), namely the “Appointments in General Practice December 2019” report. This data set has information about GP appointments for practices all over England from July 2018 to December 2019, and a separate attached document also shows the number of registered patients and practices in each CCG (Clinical Commissioning Group).

## Refining of the research questions

Firstly, we must refine exactly what we mean by these questions, in relation to the data. Demand, from a business perspective, is defined as a need for something to be sold or supplied (Press 2020). Since this data set does not include any true demand data, such as including the number of available but not booked appointments, we must come up with a suitable proxy for demand. The options for different variables to consider are limited, but the variable which would seem at first thoughts to vary the most based on demand would be the time between booking an appointment, and the appointment itself.

However, it should be noted that this variable will actually depend on numerous factors, such as the appointment availability, patient availability, urgency of the appointment, and more. As such, as conclusions drawn from this data need to be carefully considered in this context.

Also, we must specify what we mean by the “recent past”. In this report, we will define it to be the last 18 months, as this is data which is readily available to us from the NHS.

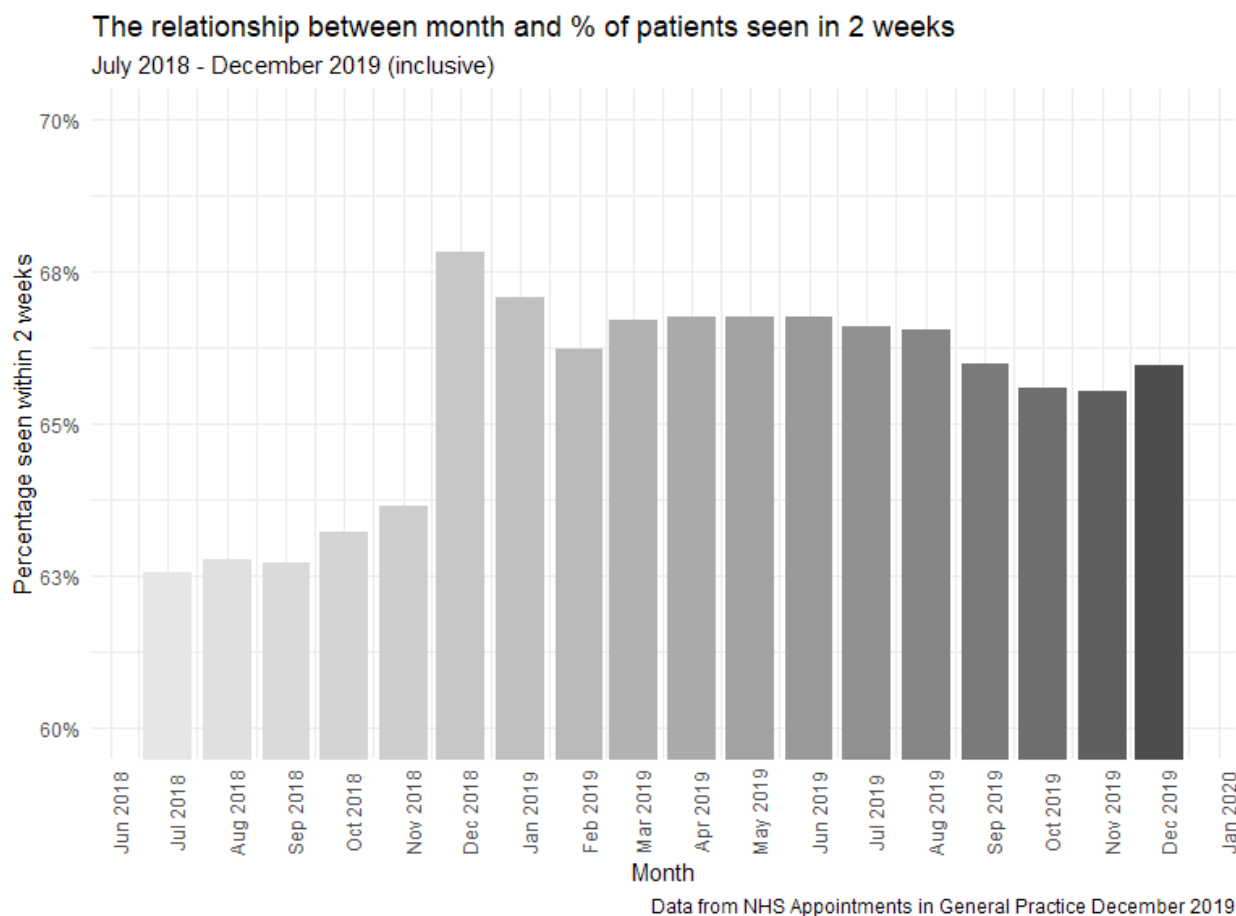
The second question is rather more explicit in what data will be exploring, namely the time between booking and the appointment, and attendance rates. This will be answered by some kind of statistical test to test significant difference over the different time periods. For example, we could show that there is significantly more missed appointments after 8 days compared to before 7 days. This would then indicate that there is a relationship.

## Analysis

### Question 1

The first step was to ensure that the data about registered patients and registered practices was correctly appended to the data about the individual appointments. This is because this data is vital if we want to be able to see the data in its full context.

The next step for this analysis was to get a visualization of that data so that we can begin to tackle it sensibly. We decided to create a bar plot with the percentage of people who booked an appointment within a week each month, from July 2018 - December 2019. We chose this comparison over the months of the year as it reduces the volatility seen if you plot the individual totals. This plot can be seen below:



What this appears to show is that the waiting times increased at the end of 2018, but the level of increase we have seen may just correspond purely to the number of registered patients/practices, so we must explore what effect this plays.

We decided to fit the model using the first 14 months and use this to predict the percentage for the 15th month and past. This will then be compared to the true value, to assess how demand is currently affecting Coventry and Rugby GPs. One limitation of this approach is that it ignores seasonality, namely that the model does not account for the fact that months across adjacent years should be somewhat correlated. Also, there is an implicit assumption that the relationship observed is linear, this may not actually be the case. In fact, the Q-Q plot for the linear model shows clusters of points suggesting that there could be multiple modes.

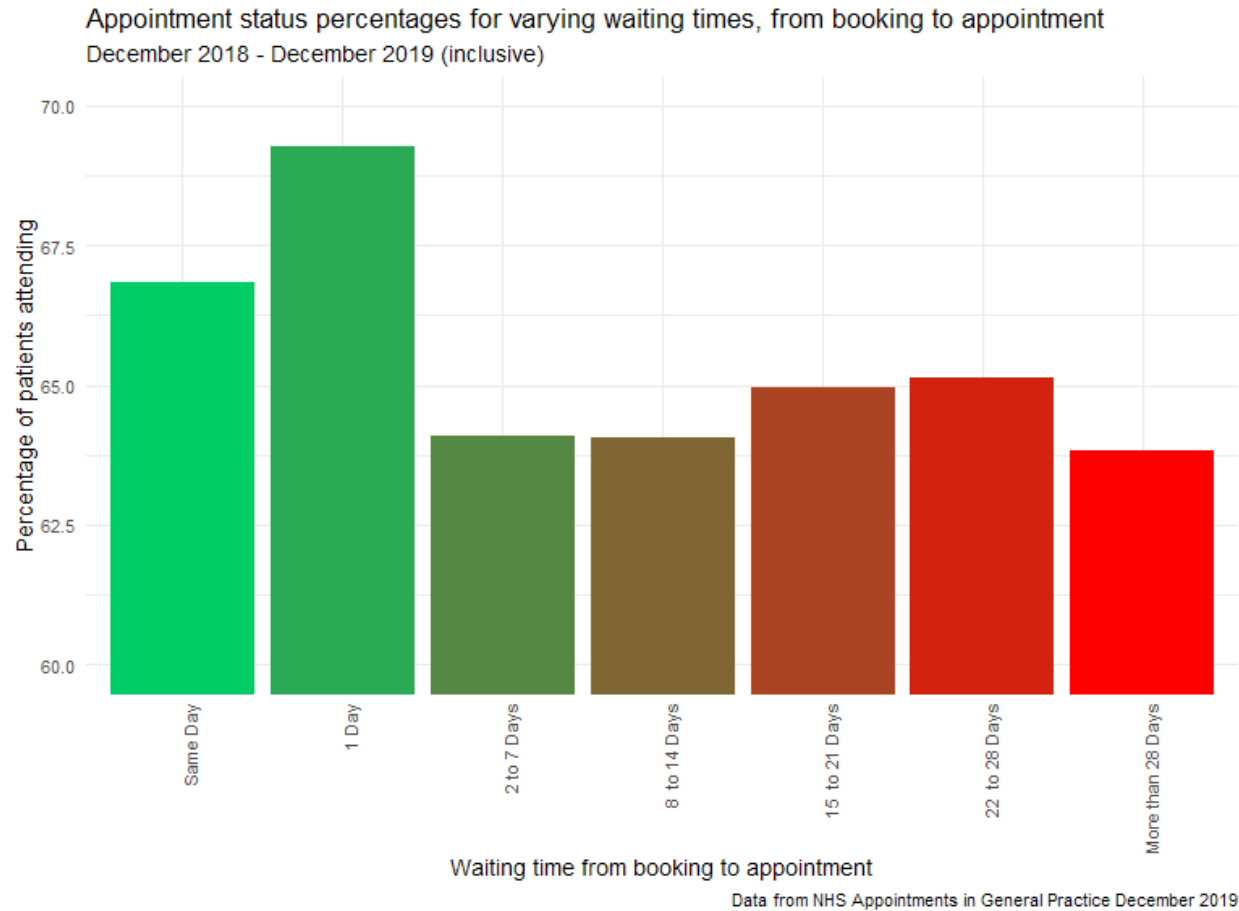
We found that in the final 4 months, the predicted values were all at least 1% over the real values. This suggests that the less than 2 week percentages in the final 4 months were lower than expected, which means more people waited more than weeks than expected, given the registered population and the number of practices.

## Question 2

For the second question, we first needed to clean up the data about the appointment statuses, as there are many “Unknown” and “Appt Status Not Provided” values. The “Unknown” entries correspond to appointments whereby the final status was never entered into the system, so we cannot use these appointments with regards to this question. “Appt Status Not Provided” also needs to be removed from the data, as again we don’t know if the patient attended.

We need to think carefully about how this removal of data could effect our results, as, for example, one could find that 80% of appointments with non-GPs are reported as “Unknown” due to a problem across the board. This may then skew the summary statistics of the data and how representative it truly is of the original population (the whole NHS). For this reason, it makes sense to explore the results both with this data in context, and without it.

Another problem with the data is that there is an issue with the data collection between June 2018 and November 2018 for certain practices, namely that the DNA (Did not attend) appointments were not being captured correctly. To account for this, these months will not be included in any analysis of the missed appointments. These affected appointments are all labelled “Appt Status Not Provided” and thus will be removed from the table anyway.



The plot seems to indicate that there is a noticeable difference between being booked in within 1 day (i.e. the same day, or the next day) and for being booked in for over 1 days time. Thus, we can regroup the waiting time variable to be TRUE or FALSE depending on whether they are seen within 1 day of booking. Then, we can use a Chi-squared test to see if the proportion of those who attend is significantly different depending on whether it is more than 1 day or not.

The result of this test is that the means of 64.4% attendance for over 1 day and 67.9% for 1 day or less are significantly different, to a 99.9% confidence level. The p-value reported from the test is less than  $2.2 \times 10^{-16}$  also, which shows the confidence level could be increased much closer to 100% also.

In order to increase confidence in this result, we thought it would be interesting to see if this trend which we are observing over 13 months really holds true in every month, or whether there appears to be any some months where this relationship is less or more strong even. There does appear to be some kind of cyclical trend throughout the year, which can be observed in the graphics file attached.

## Conclusions

One important note about the data set used in this report is that it is an experimental statistics publication, which means that it is still in the testing phase, and may suffer from issues such as poor coverage, poor data quality, and data undergoing evaluation. Also noted in the publication's support information is that there is not a national standard for data entry about activity, and the approach taken by management varies greatly between practices.

Many important data quality notes are limited in the supporting documents, but a few which are notable are that practices with an appointment rate of less than 1 appointment per registered patient per year are not

included, CCGs with 1 included practice are not included and prior to July 2019, practices with less than 1000 registered patients were also excluded.

With regards to the first question, we concluded that demand on Coventry’s GPs has not increased in the recent past, at least by how it is defined in this report. In order to answer this question in a more robust way, I would like to compare the waiting time statistics across multiple years in order to predict and thus account for the seasonality involved in health services. We found that under an increasing registered population, the number of appointments which are within 2 weeks was still decreasing from the expected values. This suggests to us that there may be a lack of funding which is stopping them from offering enough appointments, or maybe even that there is a lack of actual trained health professionals such as GPs.

The second question is a slightly more statistically rigorous result due to the Chi-squared test allow us to explicitly define that 2 variables are not independent of each other, to some degree of accuracy, 99.9% in this report. We can say with quite a high degree of certainty that this relationship exists, even with regards to the data quality issues. There were some missing values in the time between appointment and booking, which were quickly filtered out. Then, the main issue was the attendance figures. Due to an issue with data collection, as mentioned previously, this analysis only focuses on the last 12 months (from December 2019), which also slightly limits its scope of applications, but the strong relationships which we observed were satisfactory.

## References

NHS. 2020. “Appointments in General Practice December 2019.” *NHS*. NHS. <https://digital.nhs.uk/data-and-information/publications/statistical/appointments-in-general-practice/december-2019>.

Press, Cambridge University. 2020. “DEMAND: Meaning in the Cambridge English Dictionary.” *Cambridge Dictionary*. Cambridge University Press. <https://dictionary.cambridge.org/dictionary/english/demand>.