A&E delays: Why did patients wait longer?

Our econometric analysis

Econometric analysis was one of the three approaches we used to test our hypotheses.[1] We used it to see how closely the variation in A&E waiting times performance matched the variation in the potential drivers of this performance. This variation includes both the differences between hospitals at one point in time and the differences within a hospital over a period of time; this is known as panel data. We drew on a number of econometric models (essentially different combinations of statistical assumptions) to identify which factors (called variables) appear to explain type 1 A&E performance against the four-hour target and in turn the factors that appear to explain the deterioration in waiting times performance in Q3 2014/15. The data we used were at the trust (with a type 1 A&E department) and month level, covering 146 trusts between April 2011 and December 2014.

The advantage of employing econometric models is that we could explore the effect of a variable on A&E waiting times performance while controlling for the effects of other determinants of A&E waiting times performance. Without controlling for other determinants we are at risk of misattributing an effect to an irrelevant variable, *a*, that is merely acting as a proxy for a real determinant, *b* (eg because *a* and *b* are correlated).

Variables are considered statistically significant when there is a low probability that their observed association with A&E waiting times performance is driven by chance.[2] We were interested in those variables that are statistically significant **and** whose predicted effect on A&E waiting times performance is large. Additionally, we were interested in whether the values of these variables changed between Q3 2013/14 and Q3 2014/15. Where a factor shows both a change and a statistically significant effect, this is evidence that this factor contributed to the change in A&E waiting times performance.

Data

The variables we used can be divided into those relating to the inflow of patients to the A&E department and those relating to the capacity of the A&E department (and wider hospital) to respond to this inflow. Our choice of independent variables was based on our list of hypotheses, although data limitations prevented us from using econometric modelling to test the full set of our hypotheses.

The inflow variables we used related to the number and profile of attendances and admissions. These were:

---

[1] A complete description of our hypothesis is given in a separate document.

[2] Describing a relationship as being 'statistically significant at the 5% level' means that our model predicts there is only a 1 in 20 chance that the estimated relationship would have been observed if there were no true relationship between the variables.

- number of type 1 attendances and the conversion rate for those attendances into admissions

- proportion of those attendances by age group (based on Hospital Episode Statistics (HES) data[3])

- proportion of those attendances by referral source (based on HES data) [4]

- number of type 2 and 3 attendances and the conversion rate[5] for those attendances into admissions.

As we had more detailed data on admissions, we included a number of additional admission-specific variables:

- proportion of admissions with ambulatory care sensitive conditions (based on HES data)

- average Charlson index (measuring patient complexity) for a trust's admissions (based on HES data).

Finally, we included variables relating to the local area that may have affected the profile of patients attending A&E:

- social care expenditure per person (based on Department for Communities and Local Government and Office for National Statistics data).

The capacity variables we used related to the staff in A&E,[6] the occupancy level of the hospital and the general ability of the hospital to meet operational targets. These were:

- doctors, nurses and support staff in the A&E department (full-time equivalent (FTE), from the ESR dataset)

- percentage of doctors who are locums (from the ESR dataset)

---

[3] All references to HES data in this document refer to a 5% sample of the HES A&E dataset, stratified at the trust and week level, and linked to the HES Admitted Patient Care (APC) dataset.

[4] Though ambulance conveyance and emergency service referrals are both recorded in HES, we were unable to incorporate them together in the trust-level model due to collinearity between them. Each has a separate negative and significant coefficient, but including both leads to ambulance conveyance having an insignificant effect. Given this and the advantage of including other referral sources in the model, we opted to retain the emergency service referral source over ambulance conveyance.

[5] 'Conversion rates' are the proportion of A&E attenders who are admitted to hospital.

[6] We were only able to acquire staff data up to October 2014. To expand the scope of these data we extrapolated from past trends on a trust-by-trust basis.

- overnight general and acute bed occupancy of the hospital
- performance of the hospital against the referral-to-treatment (RTT) target for admitted elective activity.

Method

Over the course of our analysis we explored a number of specifications and models. Using ordinary least squares (OLS) as a base, we ran increasingly sophisticated models that relaxed some of the classical OLS assumptions and modelled the data-generating process more accurately. Here we describe our favoured models:

- fractional response model (FR)
- Arellano-Bond estimator (AB).

The FR model is a generalised linear model (GLM), similar to OLS. Instead of modelling A&E performance as a simple, linear sum of its determinants weighted by their marginal effect on A&E performance (as in OLS), FR specifies the relationship between A&E performance and its determinants as a non-linear mathematical function. The main advantage of FR over OLS is that the mathematical function is defined such that A&E performance will fall strictly between 0% and 100%. This means that extreme values of our independent variables do not lead us to predict impossible A&E performance (eg 150% of attendances meeting the target) and do not distort the effect of more moderate values.[7] In effect, we gain more precise estimates of the relationship between variables. For this reason it is our preferred method for analysing the effect on A&E waiting times performance of the variables we modelled.[8]

The AB estimator models the dynamics of A&E performance in a linear manner similar to standard OLS, using the generalised method of moments (GMM). With this we can include the previous period's A&E performance in our model, which allows us to estimate to what extent past A&E performance affects current A&E performance. The disadvantage of this is that, as past performance correlates with any current determinants that have not changed much in the short term, we lose accuracy in

[7] As the effect of an independent variable is not linear, a value 1,000 times larger than another need not have 1,000 times the effect on A&E performance.

[8] Because the marginal effect of each variable changes across the variable's range, we evaluate the marginal effect at the mean. We employ heteroscedastic-robust standard errors to account for differences in data accuracy between trusts. In technical terms, the model is a GLM with the dependent variable modelled as a member of the binomial family, and uses a logistic function as the link function.

these estimates. We note that including past performance will also pick up the effect of variables we have been unable to include in the model, such as culture.[9]

Using these models we estimated the effects of our explanatory variables on A&E waiting times performance. The size of the effects could be used to give an indication of the contribution of each variable to the decline in waiting times performance for Q3 2014/15 compared to the same period the previous year. We did this by multiplying these effects by the change in the underlying value of each variable between the two time periods.

Results

The results from our two preferred econometric models are presented in Table 1.

Identified associations with A&E waiting times performance were:

- A 10 percentage point increase in the **type 1 conversion rate** was associated with a 0.29 percentage point decline in A&E waiting times performance. This supports the view that it is admissions specifically, rather than attendances, that put pressure on A&E departments.
- An additional 1,000 **type 2 and 3 attendances** in a month were associated with a 0.29 percentage point decline in type 1 A&E waiting times performance. Two possible explanations can be given for this. First, the additional attendances generate extra work for A&E staff and reduce the time they could otherwise spend ensuring type 1 attendances meet the four-hour target. Alternatively, this finding may be picking up the effect that co-located type 3 A&E departments have on the profile of patients attending type 1 A&E departments, ie if more patients with minor illness or injury are being siphoned off to type 3 A&E departments, the illnesses/injuries of patients presenting to type 1 A&E departments will be more serious and challenging. The latter explanation is supported by the positive correlation of 0.25 between type 2 and 3 attendances and type 1 conversion rates.

[9] We employ heteroscedastic-robust standard errors to account for differences in data accuracy between trusts. In technical terms the model is based on the GLM and deals with the endogeneity of the lagged dependent variable by including lags and differences from within the model as instruments.

Table 1: Determinants of type 1 A&E department performance against the four-hour target

T1 A&E performance Determinant	Fractional response Marginal effect	Standard errors	Arellano-Bond estimator Marginal effect	Standard errors
T1 attendance	0.000187	(0.00019)	–0.000107	(0.00026)
T1 conversion rate	–0.0288***	(0.0086)	–0.00558	(0.013)
T2 and 3 attendances	–0.00292***	(0.00017)	–0.000753**	(0.00025)
T2 and 3 conversion rate	0.00505	(0.0055)	0.00517	(0.0044)
Proportion 0 to 20 years	0.0223	(0.016)	0.000997	(0.015)
Proportion 20 to 40 years	0.0492**	(0.019)	0.0195	(0.021)
Proportion 40 to 60 years				
Proportion 60 to 80 years	–0.0204	(0.022)	–0.0359	(0.024)
Proportion 80+ years	–0.102***	(0.024)	–0.0766**	(0.030)
Proportion self-referral				
Proportion GP referral	–0.0482***	(0.011)	–0.0123	(0.013)
Proportion emergency services referral	–0.0244***	(0.0040)	–0.00573	(0.0045)
Proportion other healthcare referral	–0.0731***	(0.012)	–0.0398*	(0.018)
Proportion other referral	–0.0236***	(0.0030)	–0.00745	(0.0041)
Mean Charlson index	–0.00528*	(0.0024)	–0.00632	(0.0046)
Proportion with ACSC	–0.0174	(0.0095)	–0.00581	(0.011)
Lagged social care spend	–0.0297***	(0.0031)	–0.00511	(0.0033)
Locum rate	0.00797	(0.013)	0.00858	(0.012)
Support staff	0.232*	(0.097)	0.0332	(0.11)
Nurses	–0.0914***	(0.014)	–0.0390*	(0.020)
Junior doctors	0.0868	(0.079)	–0.0499	(0.089)
Senior doctors	–0.0160	(0.039)	0.0258	(0.041)
Occupancy 85% to 90%	–0.00401***	(0.0012)	–0.00102	(0.0013)
Occupancy 90% to 95%	–0.0101***	(0.0013)	–0.00338*	(0.0014)
Occupancy 95% to 100%	–0.0160***	(0.0017)	–0.00654***	(0.0019)
RTT performance	0.0756***	(0.012)	0.0139	(0.014)
Q1	0	(.)		
Q2	0.00308*	(0.0012)	–0.00555***	(0.00090)
Q3	–0.0116***	(0.0013)	–0.0153***	(0.0011)
Q4	–0.0225***	(0.0016)	–0.0123***	(0.0015)
2011/12	0	(.)		
2012/13	–0.00921***	(0.0014)	–0.00703***	(0.0012)
2013/14	–0.00906***	(0.0014)	–0.00255*	(0.0011)
2014/15	–0.0337***	(0.0019)	–0.0162***	(0.0017)
London	0	(.)		
Midlands and East	–0.0174***	(0.0018)		
North	0.00160	(0.0013)		
South	–0.0123***	(0.0018)		
Lagged T1 performance			0.741***	(0.027)
Observations	5406		5405	
R^2	36%		64%	

t statistics in parentheses.
* p <0.05, ** p <0.01, *** p <0.001.

- Compared with the 40 to 60-year old reference group, the only two age groups associated with A&E waiting times performance are those aged **20 to 40 years** and those aged **over 80 years**. A 10 percentage point increase in the younger group as a percentage of total attendances was associated with a 0.49 percentage point increase in A&E waiting times performance. Conversely, a similar increase in the over 80-year age group was associated with a 1 percentage point decline.

- Compared with the reference group of self-referrers, every referral source was associated with a worse A&E waiting times performance. This is unsurprising as self-referrers can be expected to include the least ill patients. Referral from '**other healthcare provider**' was most associated with poor A&E waiting times performance – a 10 percentage point increase in the share of patients taking this route was associated with a 0.73 percentage point reduction in A&E performance.

- **Social care spend** was associated with a worse A&E waiting times performance. However, across all models we found that this association dissolved when trust-specific effects were included (as is the case in the AB model, among others not presented here). This strongly indicates that this variable picks up the effect of social care need (eg deprivation) between hospitals, rather than showing a direct negative consequence of social care expenditure.

- Of the staff variables, only the number of **nurses** was statistically significant at the 5% level. An increase in number of nurses of 100 was associated with a 0.9 percentage point fall in A&E waiting times performance. This result is unexpected although it should be noted that the effect size was very small (the average number of nurses was 86). This result could be driven by trusts struggling to maintain an appropriate staff mix while also struggling to meet the A&E waiting times target. Alternatively, it could be due to data problems; we were unable to differentiate between regular, bank or agency nursing staff.

- **Hospital occupancy rates** were strongly associated with A&E waiting times performance across all models. Moving from occupancy rates of under 85% to between 85% and 90% was associated with a 0.4 percentage point fall in A&E performance. Exceeding the 90% threshold was associated with a **further** 0.6 percentage point decline and exceeding 95% with a **further** 0.6% decline. That we found these results in spite of data quality concerns (occupancy is only measured once a quarter; occupancy rates cannot currently be broken down into wards that are more or less likely to influence A&E waiting times performance) leads us to suspect this is an underestimate of the true effect that hospital occupancy has on A&E waiting times performance.

- **RTT performance** was positively associated with A&E waiting times performance. This suggests that, on the whole, the relationship between the different operational performance metrics is complementary (good operational capacity leads to achievement across targets) rather than substitutive (when faced with multiple objectives, trusts focus on one target at the expense of another).
- Controls at the **time** and **regional levels** indicate that on average:
  - A&E performance is lower in the third and fourth financial quarters (1.2 percentage points and 2.3 percentage points respectively) compared with the first quarter
  - since 2011/12 A&E waiting times performance has fallen 3.4 percentage points nationally after controlling for trust-level variation in the above variables
  - performance is worse in the Midlands and the East (1.7 percentage points) and the South (1.2 percentage points) compared with London and the North.[10]

In general, these results held when we included the previous period's performance and changed the model to the AB specification; however, many became smaller in size and lost statistical significance. The **previous period's A&E waiting times performance** itself was highly significant, suggesting that the level of A&E waiting times performance tends to persist over time. A 1 percentage fall in performance one month was associated with a 0.7 percentage point fall the next month. This indicates that pressures that build up in the emergency pathway do not dissipate quickly.

By combining the marginal effects of our variables with their change over time, Table 2 shows the (percentage point) contribution of these variables to the waiting times performance decline.

[10] Note that looking at regional differences is beyond the scope of our project. Further analysis will be required to better understand the regional differences in A&E waiting times performance.

Table 2: Contribution of the determinants of A&E waiting times performance to explaining the decline between Q3 2013/14 and Q3 2014/15

Determinant	Change	FR contribution	AB contribution
T1 attendance (000s)	+0.64	+0.012	–0.007
T1 conversion rate	+0.0022	–0.0062***	–0.001
T2 and T3 attendances (000s)	+0.17	–0.05***	–0.017**
T2 and T3 conversion rate	+0.0015	+0.00076	+0.00078
Proportion 0 to 20	–0.0018	–0.004	+0.00018
Proportion 20 to 40	+0.00051	–0.016**	–0.0065
Proportion 60 to 80	+0.003	–0.006	–0.011
Proportion 80+	+0.005	–0.051***	–0.039**
Proportion GP referral	–0.0016	+0.0077***	+0.002
Proportion emergency services referral	+0.0061	–0.015***	-0.003
Proportion other healthcare referral	+0.0023	–0.017***	–0.009*
Proportion other referral	–0.015	+0.034***	+0.011
Mean Charlson index	–0.031	+0.016*	+0.019
Proportion with ACSC	–0.002	+0.0035	+0.0011
Lagged adult social care spend (£k/adult)	–0.01	+0.029***	+0.0049
Locum rate	+0.0013	+0.001	+0.0011
Support staff (000s)	+0.00012	+0.0029*	+0.00041
Nurses (000s)	+0.0046	–0.042***	–0.018*
Junior doctors (000s)	+0.00046	+0.004	–0.0023
Senior doctors (000s)	+0.0025	–0.004	+0.0064
Occupancy 85% to 90%	–0.036	+0.014***	+0.0037
Occupancy 90% to 95%	+0.0082	–0.0083***	–0.0028*
Occupancy 95% to 100%	+0.1	–0.17***	–0.068***
RTT performance (% point)	–0.026	–0.19***	–0.035
2014/15		–3.4***	–1.6***
Lagged T1 performance (% point)	–0.027		–2***
Total explained decline (% points)		–3.9	–3.8
T1 A&E performance (% points)	–4.6		

* p <0.05, ** p <0.01, *** p <0.001.

The first feature to note is that changes between Q3 2013/14 and Q3 2014/15 have been very modest. This accords with the results of our analysis (described in a separate document) – despite the widespread belief that certain factors, eg the number of sick or elderly patients attending A&E, are responsible for declining A&E waiting times performance, the limited change in these factors between Q3 2013/14 and Q3 2014/15 does not suggest they are responsible. This does not rule out their being the primary causes of some specific trusts' performance problems, but does suggest they are not the drivers behind the national picture. The general stability of the variables across the two periods also inhibits us from attributing much of the decline to specific causes.

Nevertheless, across both models the variables with the greatest capacity to explain the decline in A&E waiting times performance were those relating to the hospital – bed occupancy and RTT performance. These combine sizeable and significant effect sizes with substantial changes over time. There is currently little available national data on within-hospital processes, but our results indicate this area is a prime cause of the decline in A&E waiting times performance.

Additionally, in the AB model we found that persistence in the trend of A&E performance explains a significant part of the decline. This aligns with the explanation for the decline given by the Nuffield Trust, among others, that A&E departments exist in complicated and dynamic systems in the hospital (and more broadly in the emergency pathway). Problems in the flow at one point in the process cascade across the process in a chaotic manner and are not solvable in short timeframes.

Discussion

We tested the robustness of our findings by:

- Estimating different specifications of our model, eg reducing the number of variables to a 'bare bones' specification, including more variables such as staff turnover rates and age-specific conversion rates, and weighting the observations by A&E or trust size.
- Using multiple models/estimation techniques for our specifications. These include OLS, fixed effects and logit models. The purpose behind including these models in our research was to test whether the assumptions inherent in any individual model were driving a particular set of results. By varying both the estimation techniques and specifications we are able to identify those factors that have been consistently related to A&E performance across a range of analyses.
- Exploring the relationship between the determinants of A&E performance and individual-level waiting times from the HES A&E dataset (duration model analysis). Unfortunately, time constraints prevented us from completing this analysis; however, the results accorded with the analysis we have presented here.

Despite these tests of robustness, our approach has a number of limitations:

- It is highly unlikely that any of the variables modelled here have the same effect on A&E performance across all trusts. Individual features of specific trusts and specific months mean that we are only able to estimate the 'average' effect of these determinants. The monthly performance of any given A&E department against the four-hour target is determined by the combination of thousands of decisions across thousands of people; hence we would call it ‘noisy’. Additionally, we believe A&E performance is also affected by factors we cannot measure, such as the effect of staff culture. Reflecting this, the R^2 from our models ranged from 26% to 64%.[11] Our approach has been to explore what variables are associated with variations in performance, and not to comprehensively model every conceivable effect that influences A&E performance.

- Data availability and processing power have restricted our analysis in a number of ways. Variables derived from a 5% sample are inherently less accurate. We have tried to mitigate this by using a stratified sample. Lack of available data has prevented us from rigorously testing certain possible determinants of A&E performance, mainly surrounding alternative sources of emergency care (such as primary care), intra-hospital processes (such as patient flow and clinical culture) and post-discharge capacity (such as social and community care).

---

[11] R^2 measures the percentage of the variation in *y* (A&E performance in our models) explained by the predictions our model makes.