Part of

UNIVERSITY of York

CHE Centre For Health Economics

Do hospitals exhibit efficiency differences justifying local price variation? A retrospective study of lengths of stay for renal transplant in NHS England

Martin Chalkley, Rowena Jacobs, Adriana Castelli, Panos Kasteridis, Irene Sanchez, María José Aragón

CHE Research Paper 200

Do hospitals exhibit efficiency differences justifying local price variation? A retrospective study of lengths of stay for renal transplant in NHS England

[a] **Martin Chalkley**
[b] **Rowena Jacobs**
[c] **Adriana Castelli**
[d] **Panos Kasteridis**
[e] **Irene Sanchez**
[f] **María José Aragón**

[a] Centre for Health Economics, University of York
[b] Centre for Health Economics, University of York
[c] Centre for Health Economics, University of York
[d] Centre for Health Economics, University of York
[e] Centre for Health Economics, University of York & Research Institute for Evaluation and Public Policies, Universitat Internacional de Catalunya (UIC), Barcelona, Catalonia, Spain
[f] Centre for Health Economics, University of York & HCD Economics; Knutsford, UK

November 2025

Background to series
CHE Discussion Papers (DPs) began publication in 1983 as a means of making current research material more widely available to health economists and other potential users.

The CHE Research Paper series takes over that function and provides access to current research output via web-based publication. The RPs can take various forms and results and ideas do not necessarily represent the final position and may include work in progress not subject to peer review at the time of publication.

CHE RP 200 reports on a scoping exercise to assess whether available administrative data sources are sufficiently robust to serve as an evidence base for payment reform in the NHS. It serves to identify what gaps or additional data are required.

Acknowledgements
This research is funded by the National Institute for Health and Care Research (NIHR) Policy Research Programme, conducted through the NIHR Policy Research Unit in the Economics of Social and Health Care, PR-PRU-103/0001. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

We are grateful to Anne Mason for providing us with the code to create some of the provider level variables. We are grateful to the comments provided on previous versions of this report by clinicians and by members of NHS England Specialised Commissioning Team.

This project was undertaken using the Data Safe Haven, which is an ISO 27001 certified environment for handling sensitive data, and is provided by the University of York. We are grateful for support from the University of York Data Safe Haven team and the Research Computing team.

Further copies
Only the latest electronic copy of our reports should be cited. Copies of this paper are freely available to download from the CHE website www.york.ac.uk/che/outputs/. CHE Research Papers can be reused and distributed for non-commercial purposes provided the source is properly cited. Please use doi.org/10.15124/yao-7nn2-e683 in your citation.

Centre for Health Economics
Alcuin College
University of York
York,
YO10 5DD, UK
www.york.ac.uk/che

Abstract

Objectives To examine variation in lengths of stay between hospitals supplying renal transplant services in England, accounting for hospital and patient characteristics to establish whether there are statistically significant differences that would justify paying different local prices.

Design Analysis of retrospective administrative data using multi-level models including hospital specific and patient specific characteristics. Graphical presentation of adjusted provider-level lengths of stay and associated confidence intervals.

Setting Renal transplants undertaken in the financial year 2019/20 across 19 providers in England.

Population 3,444 patients who were subject to renal transplants for which 58% were male and whose average age was 51 years.

Data Hospital Episode Statistics (Admitted Patient Care) and Civil Registration Deaths data for England 2019/2020.

Main Outcome Measures Estimated hospital level lengths of stay for patients and associated standard errors, after accounting for a suite of hospital level and patient level variables.

Results After controlling for hospital characteristics and patient characteristics there is little statistically significant variation in length of stays between hospitals. Only 3 hospitals exhibit lengths of stay that are significantly longer or shorter than the overall average. Hence, most observed variability can be attributed to factors that are external to a hospital or outside of its control.

Conclusions There is little support from these data for the use of different locality specific prices with a view to compensating for (or providing incentives to align) differences is efficiency across locations in respect of renal transplant services. Any discretion in terms of price setting will need to be supported by data supplementary to the available administrative data analysed in this study.

1 INTRODUCTION

The NHS in England is in the process of transitioning from a system in which the prices for many hospital services were set nationally, to a system where prices are determined locally within the context of Integrated Care Systems. Accompanying this is a movement towards payments that comprise both prices for specific services and a lump sum payment; this system is termed blended payment[1].

A key point of departure in this transition, and the concern of this paper, is the emphasis on taking account of local conditions in setting prices for health care with a particular focus on quality and efficiency in delivery.

We report an analysis undertaken on renal transplantation and services that are subject to a centralised purchasing and where there is a longstanding concern to align prices to local conditions. We set out to establish whether available data could support the use of local prices set to reflect not actual costs, but costs that would prevail if a hospital were as efficient as possible. The focus is on the variation in lengths of stay since this measure is frequently used as a proxy for the resources that are consumed during treatment and thus can be viewed as an indicator of efficiency. If there are negligible differences in the health outcome following care which is the case for renal transplants, then shorter lengths of stay indicate a lower cost of delivering a service and it has been suggested that local prices should take account of efficiency differences, in the simplest case by setting a price that reflects an efficient provision of service and thereby giving an incentive for hospitals to make efficiency savings

A key challenge is to account for factors other than efficiency that might lead to variation in lengths of stay. Since such factors might operate at the level of either the hospital or the patients it treats we required an empirical approach that controlled for differences in both domains. This naturally suggests the adoption of a multi-level model (Bryk and Raudenbush, 2002).

Our findings are potentially important for the development of local approaches to pricing of the kind the NHS is embarking on and indicate that existing data may be of limited use in guiding the setting of local prices.

2 INSTITUTIONAL SETTING

Most hospital services in the NHS in England have previously been subject to a system of national pricing in which the purchasers of services were organisations called clinical commissioning groups (CCGs) and contracted with hospitals on the basis of the National Tariff Payment System. Whilst that tariff system remains in place, its most recent version sets out a different approach predicated upon a conglomeration of purchasers and providers into Integrated Care Systems and in which local discretion can be applied much more generally – see https://www.england.nhs.uk/wp-content/uploads/2020/11/22-23-National-tariff-payment-system.pdf.

There are however some specialised health care services that were and currently remain subject to centralised purchasing with discretion to set prices according to circumstances specific to the locality or a particular hospital. Renal transplants which are the subject of this research were one such service. The policy goal is to integrate specialised services into the new local-led purchasing arrangements but whilst renal transplants are considered a suitable candidate for such a move they

[1] Details of the new arrangements are set out in documents linked from https://www.england.nhs.uk/pay-syst/national-tariff/national-tariff-payment-system/#National-Tariff-Payment-System

are not yet considered ready for transition – see https://www.england.nhs.uk/wp-content/uploads/2022/05/PAR1440-specialised-commissioning-roadmap-addendum-may-2022.pdf.

Renal transplant refers to kidney transplantation in the renal replacement therapy for patients with chronic kidney disease stage 5 who are considered to be medically suitable (NHS England, 2015). The suitability for transplantation is established by assessing the potential benefits of improved quality of life and longer survival versus the risks of major surgery and chronic immunosuppression.

Renal Transplant is paid for based on the concept of Healthcare Resource Group (HRG) classification which is a cornerstone of the National Tariff Payment System and establishes complete packages of treatment that are considered homogenous in terms of their resource requirements. There are several HRGs for renal transplant, each classified according to the 3 stages of the patient's transplantation, which also describes the pathway of the service. These are:

1. Preparation for transplant.

Includes: first referral to transplant surgeon to assess suitability, surgical work-up for transplant, maintenance on transplant list, living donor suitability and multidisciplinary transplant review, work-up of potential living donor.

HRGs: LA12A (kidney pre-transplantation work-up of recipient 19+), LA10Z (live donor screening), LA11Z: kidney pre-transplantation work – live donor).

2. Transplant episode and post-discharge.

Includes: preoperative checks and tests, kidney transplant procedure, any required readmissions to theatre while patient still in hospital, all post-operative inpatient care, stent removal, elective removal of PD catheter, up to 90 days post-transplant drugs, immunosuppressive drugs for initial inpatient episode.

HRGs: LA01A (kidney transplant from cadaver non-heart beating donor 19+), LA02A (kidney transplant from cadaver heart beating donor 19+), LA03A (kidney transplant from live donor 19+), LB46Z (live donation of kidney).

3. Post-transplant outpatients.

Includes: all post-transplant outpatient activity related to both recipient and donor, antibody monitoring, annual review.

HRGs: LA13A (examination for post-transplantation of kidney – recipient is 19+), LA14Z (examination for post-transplantation of kidney of live donor.

HRGs do not capture deceased donor organ donation and cost of retrieval (responsibility of NHS Blood and Transplant), antibody-incompatible transplant, and outpatient attendances for assessing suitability for transplant.

The renal transplant HRGs have national non-mandatory prices that can be used as a starting point for local negotiations (NHS England and NHS Improvement, 2020c). There are 23 centres performing adult kidney transplantation in the UK (19 of which are in England), which performed over 3,000 transplants in 2019/20, most of them from deceased donors (NHS Blood and Transplant, 2020).

3 METHODS

3.1 Data

Our main source of data are the Hospital Episode Statistics (HES), Admitted Patient Care (APC), for the financial year 2019/20. For more details see the HES Data Dictionary: Admitted Patient Care (Health and Social Care Information Centre, 2018) and the activity report of this dataset for the year we use (NHS Digital, 2020). HES-APC records all admissions to NHS hospitals in England, its unit of reporting is the episode (period under the care of one consultant), including information regarding patient's age, gender, diagnoses and procedures. We group together consecutive episodes in the same hospital into a spell, i.e. our unit of analysis is the period admitted to a hospital (admission to discharge).

Renal Transplant patients can be identified in HES Admitted Patient Care (APC) using HRGs: LA01A (kidney transplant from cadaver non-heart beating donor 19+), LA02A (kidney transplant from cadaver heart beating donor 19+), LA03A (kidney transplant from live donor 19+), LB46Z (live donation of kidney).

3.1.1 Individual Level Variables

In HES-APC we observe both patient characteristics and some clinical variables: patient characteristics include age, gender and deprivation, and clinical variables include type of treatment (elective / non-elective), diagnoses, procedures and whether the patient died in hospital.

We linked HES-APC data to Civil Registration data to identify whether the patient was alive a year after his/her hospital admission.

3.1.2 Provider Level Variables

In terms of provider (hospital) level variables, we include variables to identify specific hospital characteristics (e.g. teaching status) and type of hospital activity (e.g. degree of specialisation) (Street et al., 2012a).

Teaching or Foundation Trust status can be inferred from the name of each Trust: i.e. if the name includes ‘teaching’ (‘foundation trust’), the teaching (foundation trust) status indicator is equal to one. Teaching hospitals tend to be bigger hospitals and their funding is different; on top of the funds they receive related to the care they provide (as all hospitals do), they also receive funding for their teaching/research activities, therefore it is possible that the two types of activity cross-subsidise each other. NHS hospitals are divided into Foundation Trusts (FTs) and non-Foundation Trusts (NFTs), FTs are not-for-profit public organisations with greater managerial and financial autonomy from direct central government control (Department of Health, 2003).

In order to measure hospital size, we use the average number of overnight beds available over the four quarters in the financial year. Available number of beds is reported each quarter by NHS England.[2]

Volume of activity is measured as the number of patients treated in a given financial year, and it is calculated using the HES-APC data, counting the number of different patients in each provider.[3]

The degree of specialisation of each hospital trust is measured using a Gini coefficient based on the activity in different HRG chapters (Daidone and D’Amico, 2009), (Street et al., 2012b).

3.2 Empirical implementation

We adopt an approach that takes into account that providers treat different patients and

[2] See https://www.england.nhs.uk/statistics/statistical-work-areas/bed-availability-and-occupancy/bed-data-overnight/ [accessed 19 May 2021].

[3] Note that the sum of the number of patients across providers will not match the total number of patients in HES as patients can be admitted to more than one hospital during the financial year. Also, the number of patients in a given provider can be different to the number of admissions, as the same patient can be admitted more than once.

make choices given a set of exogenous factors, therefore choices are correlated with one another. Multilevel models, described below, can be useful in this context as they consider patients nested within providers, they allow for explanatory variables at both levels, i.e. patient and provider,[4] and account for variation due to exogenous factors. Multilevel models are used to analyse data with a nested structure and where each level of data has its own variability. Ignoring the fact that there are multiple sources of variability might lead to incorrect conclusions (Snijders and Bosker, 2012).

Our two-level model has patients i nested in hospitals j, uses a patient outcome (which can be a measure of quality or of efficiency) as dependent variable and as explanatory variables patient and hospital characteristics. It takes the form:

$$Y_{ij} = \beta_0 + \sum_{m=1}^{N} \beta_m X_{ij}^m + \sum_{m=N+1}^{M} \beta_m X_j^m + \mu_j + e_{ij}$$

Where Y_{ij} is length of stay for patient i in hospital j, β_0 is the common intercept, β_m are the coefficients for the N patient (X_{ij}^m) and the $M - N$ hospital (X_j^m) level variables, μ_j is the hospital specific intercept and e_{ij} is the residual. Residuals are assumed to be normal.

The variation across hospitals can be reported in different ways. The Intraclass Correlation Coefficient (ICC) measures how much of the observed variation in the dependent variable is due to the hospital, with the ICC defined as (Snijders and Bosker, 2012):

$$ICC = \frac{Var(\mu_j)}{Var(Y_{ij})} = \frac{Var(\mu_{0j})}{Var(\mu_{0j}) + Var(e_{ij})}$$

The ICC shows us the extent to which the Y_{ij} from individuals in the same group are more alike as compared with individuals from other groups.

We can also plot the group level residuals and see if they are different from the average. These plots are known as 'caterpillar plots' and show a confidence interval for the group specific residual, which can be used to identify differences from the overall average (β_0).

Whether the observed variation between hospitals (after controlling for other factors) represents variation in "performance" depends on the quality of the data available and the model specification. When analysing health outcomes, it is important to consider the definition of the relevant population and the outcome, the explanatory variables considered and the quality of the data used (Goldstein and Spiegelhalter, 1996).

Assuming we have controlled for the right explanatory variables, if a hospital specific intercept, μ_j, is significantly different from the average, we can say that that hospital's local price should be different

[4] For example, a Teaching hospital is likely to have a different staff workload than other hospitals as part of the time of some staff is used to teach, leaving less time to treat patients. This will have effects on both costs and quality. We believe it is reasonable to expect costs to be higher as more staff is required to have a given number of patient care hours (because part of the staff's time is used to teach). In terms of quality it could go either way: quality of treatment could be higher due to the longer time spent with each patient (the doctor must not only deliver the treatment but also has to explain it to the students, which means spending more time with the patient) and could be lower if the teaching "distract" attention from the patient.

from the national average based on its performance on outcome Y_{ij}. However, if we find no differences between providers, there would be no justification for variation in local prices.

We used MLwiN v3.05 (Rasbash et al., 2020a) to estimate multilevel models[5] within Stata 16 (StataCorp., 2019), using the command runmlwin (Leckie and Charlton, 2013).

Explanatory variables we used were patient-based (e.g. age, sex, comorbidities) and hospital-based (e.g. size, teaching status) characteristics.

4 RESULTS

The unit of analysis for Renal Transplant is the HES-APC spell. We only consider spells which include episodes with HRGs LA01A (Kidney Transplant, 19 years and over, from Cadaver Non Heart-Beating Donor), LA02A (Kidney Transplant, 19 years and over, from Cadaver Heart-Beating Donor), LA03A (Kidney Transplant, 19 years and over, from Live Donor), LB46Z (Live Donation of Kidney).

In the appendix we show results separating the renal transplant HRGs into three groups, to distinguish between transplantations from live (LA03A) and cadaver (LA01A and LA02A) donors and live donation of kidney (LB46Z). We also present results for elective and non-elective transplant procedures separately.

The renal transplantation sample consists of 3,444 patients. Descriptive statistics are presented in Table 1. Renal Transplantation admissions last on average 8.68 days, the median is seven days and the 99th percentile is 40 days (the distribution is skewed). Around 1.6 percent of patients dies within a year of their hospital admission.

[5] For details on how to use this software see Rasbash, J., Steele, F., Browne, W. J. and Goldstein, H. (2020b). A User's Guide to MLwiN, v3.05. Centre for Multilevel Modelling, University of Bristol.

Table 1. Descriptive Statistics – Renal Transplantation

Variable	Mean	P1	P99
Dependent variable			
Renal transplantation length of stay	8.68	2	40
Demographic			
1-year mortality	0.016	0	1
Age in years (at admission)	51.02	21	76
Male	0.58	0	1
Index of multiple deprivation			
quintile 1 – Most Deprived (ref)			
quintile 2	0.20	0	1
quintile 3	0.20	0	1
quintile 4	0.19	0	1
quintile 5	0.16	0	1
Clinical			
Elective	0.47	0	1
Number of procedures	5.07	2	22
Number of diagnoses	8.06	1	20
Adverse event	0.03	0	1
Post-operation complications	0.01	0	1
Died in Hospital	0.003	0	0
Hospital level variables			
CQC score is good	0.68	0	1
Average number of beds 2019/20	1443.48	931.96	2629.99
Teaching Trust	0.09	0	1
Foundation Trust	0.59	0	1
Number patients in 1000s	131.05	77.07	214.28
Specialisation Index	0.22	0.10	0.40
Survival rate in top quartile	0.35	0	1
Survival rate in lower quartile	0.14	0	1
Offer decline rate lower quartile	0.06	0	1
Number of Observations	3,444		

P1, P99 are 1%, 99% percentiles, respectively.

Figure 1 shows the distribution of length of stay in each provider. All admissions from a provider are arranged in a column, the mean of each provider is represented with a diamond and the overall average with a horizontal line. Providers are ordered alphabetically based on their code, a list of codes and names can be found in the Appendix.

Figure 1. Length of Stay. Renal Transplantation

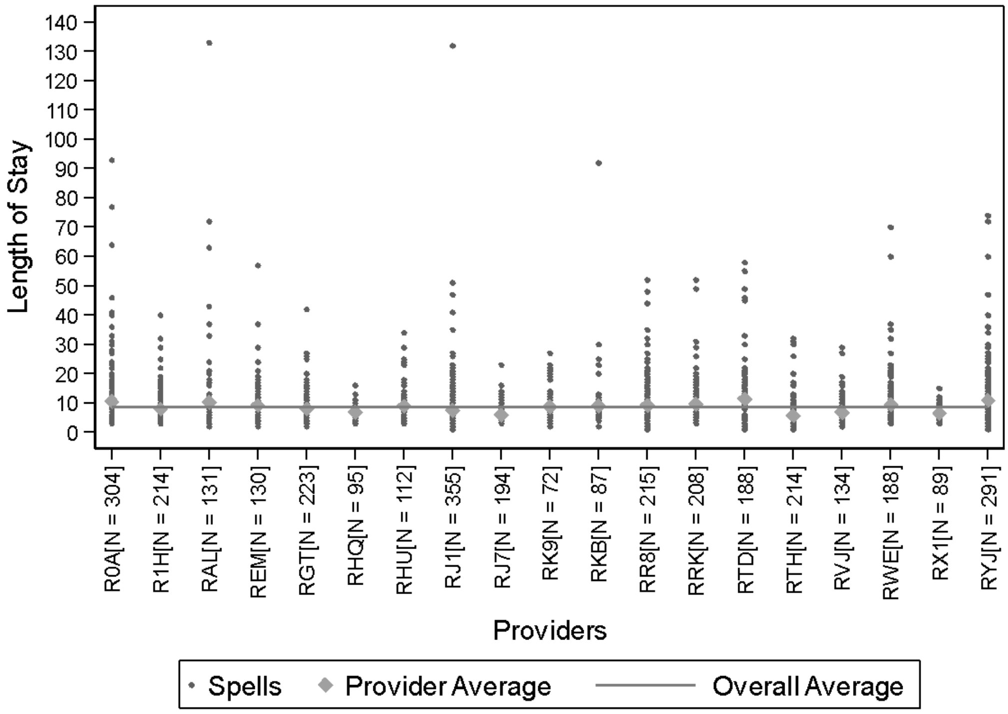


Table 2 presents estimates from the multilevel model. Residuals were not normal distributed after running the model using the dependent variable, length of stay, in its original form, so we use its natural logarithm, which does generate normal residuals (as the model assumes). The coefficients on a log-linear model can be interpreted as semi-elasticities, i.e. a coefficient of 0.01 indicates that a unit change in the explanatory variables is associated with a 1% change in the dependent variable.

We find no significant associations between patients' characteristics and the duration of renal transplantation admissions.

Elective admissions are negatively associated with the duration of renal transplantation admissions. The interpretation of the results relating to clinical variables must be considered with caution because they are likely to be endogenous, for example, the number of procedures may be a choice variable, but they could also indicate a more complex caseload.

We also include in the regression hospital level variables that are unlikely to be under hospitals' control in the short run. They are not statistically significant. Neither are the variables that reflect the historical survival rate of the providers' patients or the provider's offer decline rate.

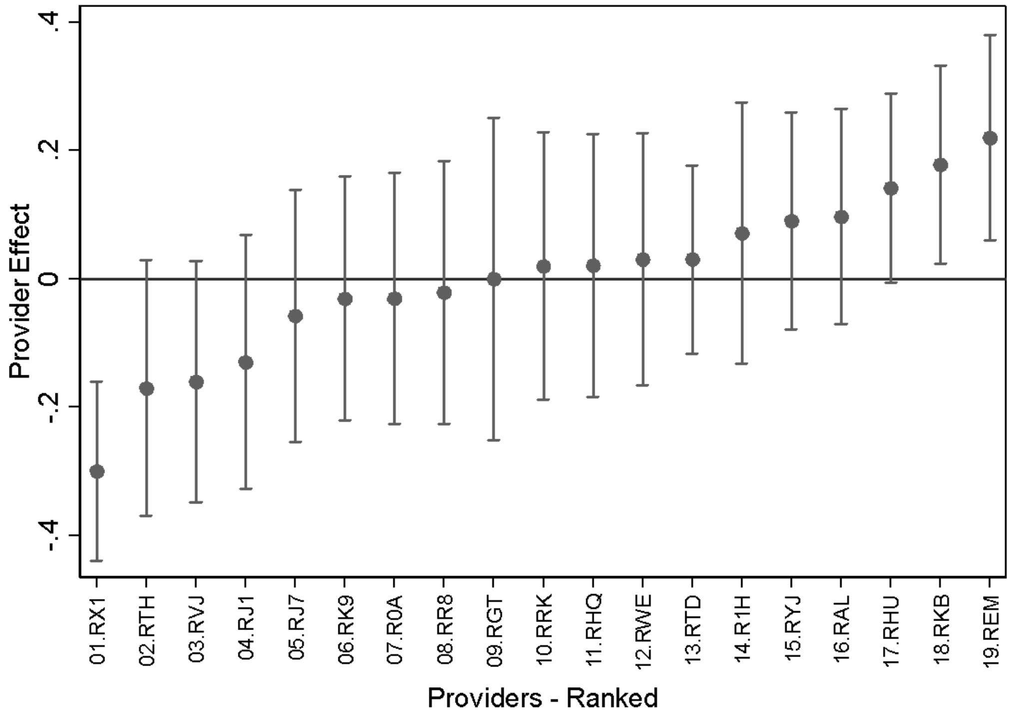
The intraclass correlation is 0.092 [0.016/(0.016+0.158)], which implies that about 9.2% of the variation in renal transplantation days is explained by between providers' heterogeneity.

Table 2. Estimates for renal transplantation

Dependent Variable: ln(Length of Stay)	Mean	P-value
Demographic		
1-year mortality	0.007	0.893
Age in years (at admission)	0.001	0.197
Male	0.002	0.867
Index of multiple deprivation		
quintile 1 – Most Deprived (ref)		
quintile 2	-0.017	0.402
quintile 3	-0.013	0.527
quintile 4	-0.017	0.417
quintile 5	-0.004	0.877
Clinical		
Elective	-0.234	<0.001
Number of procedures	0.056	<0.001
Number of diagnoses	0.052	<0.001
Adverse event	0.005	0.912
Post-operation complications	0.094	0.144
Died in Hospital	-0.476	<0.001
Hospital level variables		
CQC score is good	0.073	0.502
Av num. of beds 2017/18 (in 100s)	0.044	0.113
Teaching Trust	-0.045	0.699
Foundation Trust	0.057	0.473
Number patients in 1000s	-0.005	0.129
Specialisation Index	0.039	0.927
Survival rate in top quartile	-0.048	0.636
Survival rate in lower quartile	0.238	0.037
Offer decline rate lower quartile	-0.118	0.491
Between providers variation	0.016	0.004
Between patients variation	0.158	<0.001
Intraclass correlation	0.092	
Number of Observations	3,444	

The relative efficiency across providers is shown graphically in Figure 2. There are 19 providers performing renal transplantations. Only three providers are significantly different from the average, one has a shorter than expected (given the observed patient and provider characteristics) length of stay and two have longer than expected stays.

Figure 2. Relative efficiency across providers – Renal transplantation days



5 DISCUSSION

This research has shown that after controlling for patient and institutional factor there is relatively little variation in performance of NHS England hospitals in terms of the time that patients spend in hospital for renal transplants. Hence, there are few differences in resources used once other factors have been accounted for.

This is important because it establishes that there is not an evidential basis, given current data, for paying different prices on account of differences in speed of treatment. Furthermore, since we were unable to establish any significant variation in outcomes – similar resources lead to similar health benefits – there is also no evidence of significant differences in efficiency.

This is a relevant finding for policy in this area because one rationale for allowing variation in prices across hospitals is to seek to influence or account for efficiency. Either inefficient hospitals are paid according to their observed costs – hence supporting those inefficiencies or they are paid lower prices to establish incentives for them to make efficiency savings. This lies at the heart of the NHS reforms that encourage more local price negotiation with a view to incentives – under what is called the aligned payment incentive approach of the NTPS -- https://www.england.nhs.uk/wp-content/uploads/2021/12/22-23NT_Dt-Guidance-on-aligned-payment-and-incentive-approach.pdf.

Our results indicate that existing administrative data regarding performance does not offer strong evidential support for sustaining these differential efficiency payments.

There are important limitations to our analysis, and which point the way for strengthening evidence in the future.

As with any retrospective exercise we are making statements that relate to an earlier period. In this case treatments undertaken before the effects of Covid19 impacted on the health care system and prior to the changes in funding that are currently underway. Therefore, updating the analysis to account for stable post Covid19 data would be beneficial.

We rely on the appropriateness of the model and the accuracy of the data. In respect of the former, we have chosen a framework that is broadly supported as being appropriate for establishing differences after controlling for multiple levels of effects. In respect of the latter, the same concern would arise if data were to be used to inform policy choices – so data veracity is a concern for the proposed approach as much for our analysis.

It is also important to note that because data used in this study do not support this payment reform function, it does not imply that the function cannot be supported using better data. One use of our research is to establish priorities in terms of data collection that will perform the required function.

REFERENCES

Bryk, Stephen W. Raudenbush, Anthony S. (2002). Hierarchical linear models : applications and data analysis methods (2. ed., [3. Dr.] ed.). Thousand Oaks, CA [u.a.]: Sage Publications. ISBN 978-0-7619-1904-9.

Daidone, S. and D'Amico, F. J. J. o. P. A. (2009). Technical efficiency, specialization and ownership form: evidences from a pooling of Italian hospitals. 32, 203.

Department of Health (2003). Health and Social Care (Community Care and Standards) Act. (Ed, Health, D. o.). London.

Goldstein, H. and Spiegelhalter, D. J. (1996). League Tables and Their Limitations: Statistical Issues in Comparisons of Institutional Performance. *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 159, 385-443.

Health and Social Care Information Centre (2017a). Adult Critical Care (ACC) Hospital Episodes Statistics (HES) Data Dictionary.

Health and Social Care Information Centre (2017b). Hospital Adult Critical Care Activity 2015-16.

Health and Social Care Information Centre (2018). HES Data Dictionary: Admitted Patient Care.

Leckie, G. and Charlton, C. (2013). runmlwin: A Program to Run the MLwiN Multilevel Modeling Software from within Stata. *Journal of Statistical Software; Vol 1, Issue 11 (2013)*.

NHS Blood and Transplant (2020). Annual Report on Kidney Transplantation 2019/20.

NHS Digital (2020). Hospital Admitted Patient Care Activity, 2019-20.

NHS England (2015). National Tariff Payment System: A consultation notice. In *Annex 7a: Specified acute services for local pricing.*

NHS England and NHS Improvement (2016). 2017/18 and 2018/19 National Tariff Payment System.

NHS England and NHS Improvement (2020a). Developing the payment system for 2021/22: Engagement on national tariff and related contracting policies for 2021/22.

NHS England and NHS Improvement (2020b). Integrating care: Next steps to building strong and effective integrated care systems across England.

NHS England and NHS Improvement (2020c). National tariff payment system documents, annexes and supporting documents.

Rasbash, J., Steele, F., Browne, W. J. and Goldstein, H. (2020a). MLwiN version 3.05. Centre for Multilevel Modelling, University of Bristol.

Rasbash, J., Steele, F., Browne, W. J. and Goldstein, H. (2020b). A User's Guide to MLwiN, v3.05. Centre for Multilevel Modelling, University of Bristol.

Snijders, T. A. B. and Bosker, R. J. (2012). *Multilevel Analysis.* SAGE Publications.

Street, A., Kobel, C., Renaud, T., Thuilliez, J., Group, O.B. of the E., 2012. How Well Do Diagnosis-Related Groups Explain Variations in Costs or Length of Stay Among Patients and Across Hospitals? Methods for Analysing Routine Patient Data. Health Economics 21, 6–18. https://doi.org/10.1002/hec.2837

StataCorp. (2019). Stata Statistical Software: Release 16. College Station, TX: StataCorp LLC.

APPENDIX. List of Renal Transplant Providers

Code	Name
R0A	Manchester University NHS Foundation Trust
R1H	Barts Health NHS Trust
RAL	Royal Free London NHS Foundation Trust
REM	Liverpool University Hospitals NHS Foundation Trust
RGT	Cambridge University Hospitals NHS Foundation Trust
RHQ	Sheffield Teaching Hospitals NHS Foundation Trust
RHU	Portsmouth Hospitals University NHS Trust
RJ1	Guy's and St Thomas' NHS Foundation Trust
RJ7	St George's University Hospitals NHS Foundation Trust
RK9	University Hospitals Plymouth NHS Trust
RKB	University Hospitals Coventry and Warwickshire NHS Trust
RR8	Leeds Teaching Hospitals NHS Trust
RRK	University Hospitals Birmingham NHS Foundation Trust
RTD	The Newcastle upon Tyne Hospitals NHS Foundation Trust
RTH	Oxford University Hospitals NHS Foundation Trust
RVJ	North Bristol NHS Trust
RWE	University Hospitals of Leicester NHS Trust
RX1	Nottingham University Hospitals NHS Trust
RYJ	Imperial College Healthcare NHS Trust

APPENDIX. Results by type of transplant.

In this appendix we reproduce the results presented in Table 2 and Table 3, but separating transplantations from live (LA03A) and cadaver (LA01A and LA02A) donors and live donation of kidney (LB46Z).

Table 3 shows descriptive statistics by type of transplant (equivalent to Table 2). Admissions for transplantations from cadaver donors are longer, have higher one-year mortality and are for older patients than those for transplantations from live donor.

Table 3. Descriptive Statistics – by type of transplant

	Transplantations from live donor (LA03A) N = 754			Transplantations from cadaver donor (LA01A and LA02A) N = 1,932			Live donation of kidney (LB46Z) N = 758		
Variable	**Mean**	**P1**	**P99**	**Mean**	**P1**	**P99**	**Mean**	**P1**	**P99**
Dependent variable									
Length of stay	8.82	4	41	10.52	4	46	3.86	1	11
Demographic									
1-year mortality	0.009	0	0	0.024	0	1	0.001	0	0
Age in years	47.43	20	76	53.52	22	76	48.22	22	74
Male	0.59	0	1	0.61	0	1	0.48	0	1
IMD (reference: Q1 – Most Deprived)									
quintile 2	0.19	0	1	0.21	0	1	0.17	0	1
quintile 3	0.20	0	1	0.21	0	1	0.18	0	1
quintile 4	0.21	0	1	0.17	0	1	0.22	0	1
quintile 5	0.19	0	1	0.13	0	1	0.20	0	1
Clinical									
Elective	0.90	0	1	0.10	0	1	0.97	0	1
Num procedures	5.26	2	21	5.74	2	24	3.20	2	9
Num diagnoses	8.34	2	20	9.76	2	20	3.43	1	12
Adverse event	0.04	0	1	0.03	0	1	0.01	0	0
Post-op complic.	0.01	0	0	0.02	0	1	0.01	0	0
Died in Hospital[1]	-			0.006	0	0	-		
Hospital level variables									
CQC score is good	0.68	0	1	0.66	0	1	0.72	0	1
Av. Num beds [100s]	14.41	9.32	26.30	14.61	9.32	26.30	14.01	9.32	26.30
Teaching Trust	0.08	0	1	0.10	0	1	0.08	0	1
Foundation Trust	0.61	0	1	0.58	0	1	0.61	0	1
N patients [1000s]	131.42	77.07	214.28	131.90	77.07	214.28	128.52	77.07	214.28
Specialisation Index	0.22	0.10	0.40	0.22	0.10	0.40	0.23	0.10	0.40
Survival rate in top quartile	0.33	0	1	0.35	0	1	0.35	0	1
Survival rate in lower quartile	0.16	0	1	0.13	0	1	0.16	0	1
Offer decline rate lower quartile	0.05	0	1	0.07	0	1	0.06	0	1

[1] There are no patients/donors who died in hospital during a live donor transplantation admission.

Table 4 shows the regression results for length of stay by type of transplant.

Table 4. Regression Results – by type of transplant

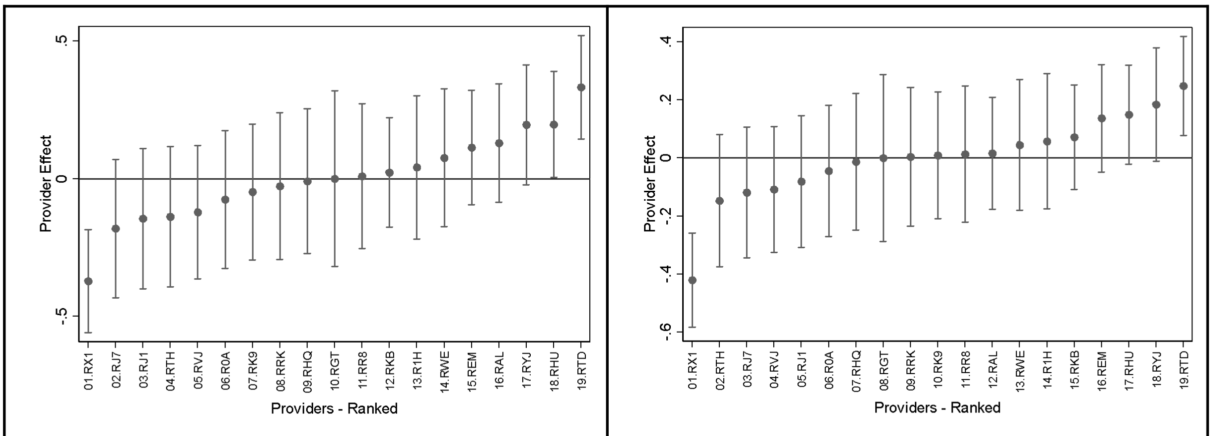
Dependent Variable: ln(Length of Stay)	**Transplantations from live donor (LA03A) N = 754**		**Transplantations from cadaver donor (LA01A and LA02A) N = 1,932**		**Live donation of kidney (LB46Z) N = 758**	
Variable	**Mean**	**P-value**	**Mean**	**P-value**	**Mean**	**P-value**
Demographic						
1-year mortality	-0.016	0.887	0.030	0.578	0.117	0.724
Age in years (at admission)	0.001	0.401	0.002	<0.001	0.001	0.306
Male	-0.015	0.483	-0.021	0.211	-0.034	0.164
Index of multiple deprivation quintile 1 – Most Deprived (ref)						
quintile 2	-0.049	0.144	-0.029	0.238	-0.056	0.149
quintile 3	-0.053	0.114	-0.016	0.527	-0.062	0.115
quintile 4	-0.035	0.301	-0.006	0.824	-0.073	0.054
quintile 5	-0.051	0.148	-0.017	0.551	-0.066	0.088
Clinical						
Elective	0.010	0.806	0.005	0.874	0.110	0.182
Number of procedures	0.070	<0.001	0.055	<0.001	0.053	<0.001
Number of diagnoses	0.018	<0.001	0.028	<0.001	0.042	<0.001
Adverse event	-0.055	0.351	0.106	0.029	0.079	0.638
Post-operation complications	0.158	0.216	0.150	0.025	0.314	0.070
Died in Hospital		(omitted)	-0.352	-0.357		(omitted)
Hospital level variables						
CQC score is good	0.056	0.691	0.077	0.539	0.160	0.354
Av num. of beds (in 100s)	3.804	0.291	3.924	0.221	-1.281	0.773
Teaching Trust	-0.123	0.416	-0.031	0.815	-0.149	0.425
Foundation Trust	-0.013	0.901	0.022	0.810	0.003	0.979
Number patients in 1000s	-0.004	0.347	-0.005	0.214	0.003	0.560

Dependent Variable: ln(Length of Stay)	**Transplantations from live donor (LA03A) N = 754**		**Transplantations from cadaver donor (LA01A and LA02A) N = 1,932**		**Live donation of kidney (LB46Z) N = 758**	
Variable	**Mean**	**P-value**	**Mean**	**P-value**	**Mean**	**P-value**
Specialisation Index	0.048	0.93	0.180	0.713	-0.588	0.382
Survival rate in top quartile	0.000	1.000	-0.003	0.981	-0.067	0.679
Survival rate in lower quartile	0.113	0.439	0.267	0.042	0.160	0.378
Offer decline rate lower quartile	-0.174	0.435	-0.179	0.361	0.027	0.922
Between providers variation	0.026	0.005	0.021	0.004	0.021	0.002
Between patients variation	0.079	<0.001	0.129	<0.001	0.107	<0.001
Intraclass correlation	0.248		0.140		0.277	

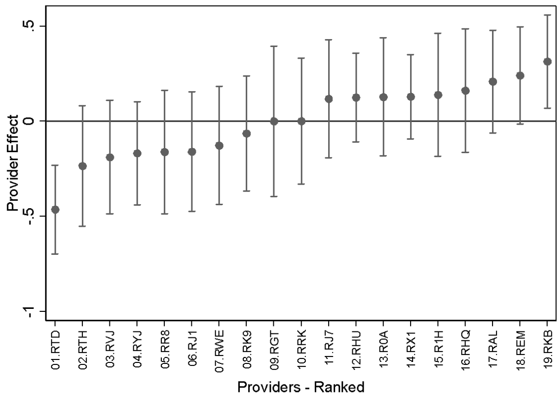
Note: All variables in the table are included in all models, but some of them were omitted because there was no variation, e.g. if no patient dies in hospital, the variable "died in hospital" is omitted from the regression.

Figure 3 shows the provider effects for the three regressions reported in Table 7. We observe only two or three outliers in the different types of transplant.

Figure 3. Relative efficiency across providers – by type of transplant



Transplantations from live donor (LA03A)	Transplantations from cadaver donor (LA01A and LA02A)



Live donation of kidney (LB46Z)

APPENDIX. Results by type of admission.

In this appendix we reproduce the results presented in Table 2 and Table 3, but separating transplantations that are elective from those that are not.

We start with the descriptive statistics, equivalent to Table 2, by type of admission, where we see that non-elective admissions are longer, have higher one-year mortality, have older patients and record more procedures and diagnoses than elective ones.

Table 5. Descriptive Statistics – by type of admission

	Elective N = 1,605			Non-Elective N = 1,839		
Variable	**Mean**	**P1**	**P99**	**Mean**	**P1**	**P99**
Dependent variables						
Length of stay	6.77	1	30	10.35	4	45
Demographic						
1-year mortality	0.007	0	0	0.023	0	1
Age in years	48.66	20	75	53.08	22	76
Male	0.54	0	1	0.61	0	1
IMD (reference: Q1 – Most Deprived)						
quintile 2	0.18	0	1	0.21	0	1
quintile 3	0.19	0	1	0.21	0	1
quintile 4	0.21	0	1	0.17	0	1
quintile 5	0.19	0	1	0.13	0	1
Clinical						
Waiting Time (N = 1,418)	35.27	1	276	N/A		
Num procedures	4.25	2	15	5.79	2	24
Num diagnoses	6.28	1	20	9.61	2	20
Adverse event	0.02	0	1	0.03	0	1
Post-op complication	0.01	0	0	0.02	0	1
Died in Hospital	0.001	0	0	0.005	0	0
Hospital level variables						

CQC score is good	0.73	0	1	0.63	0	1
Av. number beds	14.24	9.32	26.30	14.60	9.32	26.30
Teaching Trust	0.09	0	1	0.09	0	1
Foundation Trust	0.64	0	1	0.55	0	1
N patients in 1000s	129.98	77.07	214.28	131.99	77.07	214.28
Specialisation Index	0.23	0.10	0.40	0.21	0.10	0.40
Survival rate in top quartile	0.33	0	1	0.36	0	1
Survival rate in lower quartile	0.17	0	1	0.12	0	1
Offer decline rate lower quartile	0.04	0	1	0.09	0	1

Table 6 shows the regression results for length of stay by type of admission.

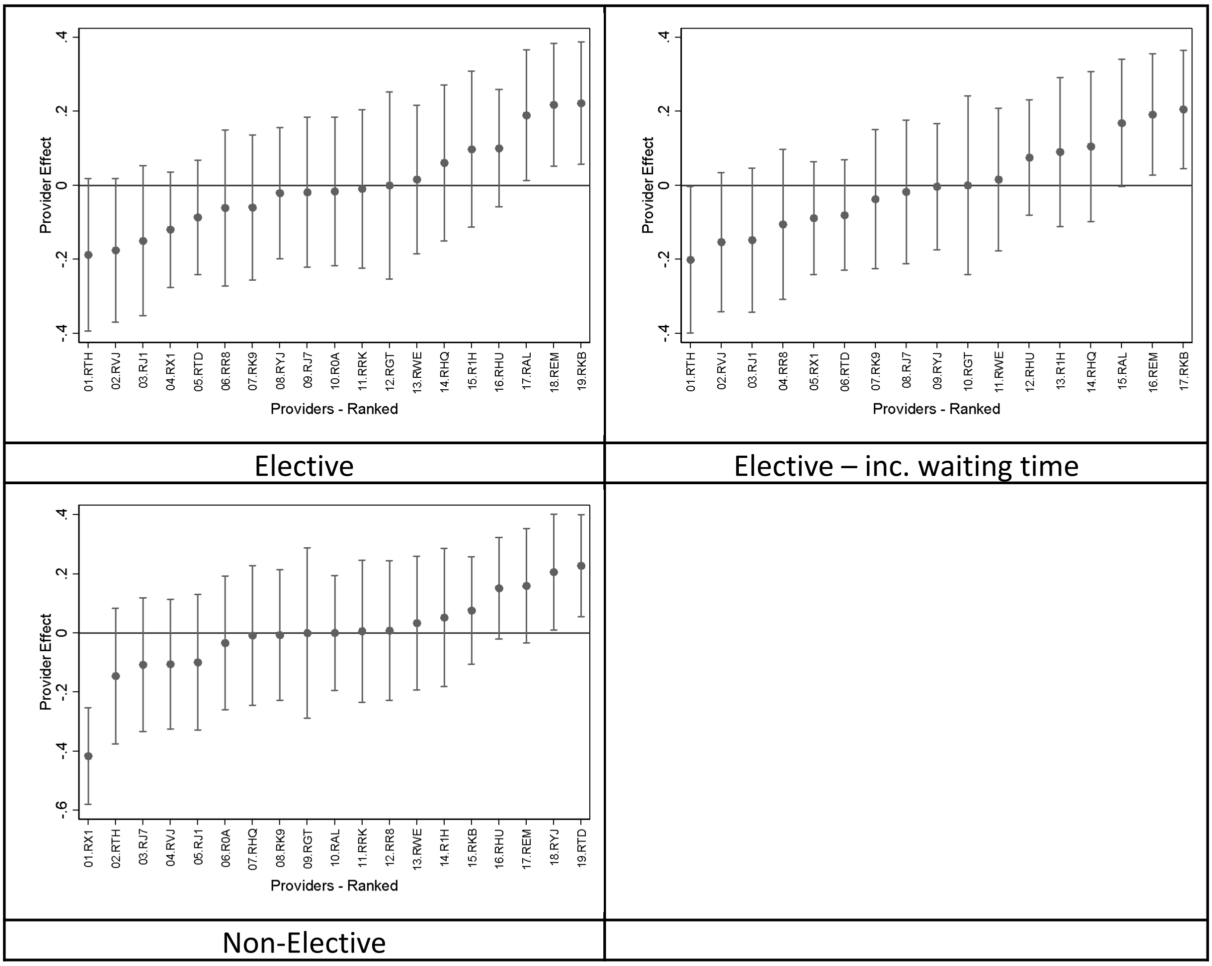
Table 6. Regression Results – by type of admission

	Elective N = 1,605		**Elective with Waiting Time N = 1,418**		**Non-Elective N = 1,839**	
Variable	**Mean**	**P-value**	**Mean**	**P-value**	**Mean**	**P-value**
Demographic						
1-year mortality	-0.058	0.629	-0.148	0.373	0.041	0.471
Age in years (at admission)	0.000	0.713	-0.001	0.219	0.002	0.001
Male	0.005	0.811	0.006	0.765	-0.015	0.375
Index of multiple deprivation						
quintile 1 – Most Deprived (ref)						
quintile 2	-0.004	0.895	0.016	0.641	-0.025	0.315
quintile 3	-0.018	0.575	0.010	0.780	-0.023	0.359
quintile 4	-0.020	0.527	-0.007	0.843	-0.011	0.688
quintile 5	0.035	0.287	0.038	0.281	-0.040	0.186
Clinical						
Waiting Time	N/A		0.000	0.311	N/A	
Number of procedures	0.074	<0.001	0.073	<0.001	0.052	<0.001
Number of diagnoses	0.067	<0.001	0.067	<0.001	0.031	<0.001
Adverse event	-0.121	0.102	-0.124	0.110	0.093	0.062
Post-operation complications	0.087	0.486	0.074	0.565	0.146	0.035
Died in Hospital	-1.209	<0.001		(omitted)	-0.204	0.084
Hospital level variables						
CQC score is good	0.025	0.827	0.048	0.664	0.101	0.423
Av num. of beds (in 100s)	4.808	0.099	3.942	0.161	3.702	0.252
Teaching Trust	-0.063	0.613	-0.132	0.273	-0.050	0.711
Foundation Trust	0.061	0.466	0.051	0.527	0.035	0.708

	Elective N = 1,605		Elective with Waiting Time N = 1,418		Non-Elective N = 1,839	
Variable	**Mean**	**P-value**	**Mean**	**P-value**	**Mean**	**P-value**
Number patients in 1000s	-0.005	0.202	-0.004	0.311	-0.005	0.245
Specialisation Index	-0.071	0.872	-0.313	0.465	0.149	0.762
Survival rate in top quartile	-0.074	0.480	-0.068	0.503	-0.014	0.907
Survival rate in lower quartile	0.170	0.153	0.167	0.143	0.245	0.063
Offer decline rate lower quartile	-0.102	0.573	-0.079	0.653	-0.169	0.391
Between providers variation	0.017	0.007	0.015	0.009	0.022	0.004
Between patients variation	0.165	<0.001	0.159	<0.001	0.130	<0.001
Intraclass correlation	0.093		0.086		0.145	

Figure 4 shows the provider effects for the three regressions reported in Table 9. We observe at most three outliers in each type of admission.

Figure 4. Relative efficiency across providers – by type of admission



Part of

UNIVERSITY of York

CHE
Centre For Health Economics