

Predicting hospital costs for patients receiving renal replacement therapy to inform an economic evaluation

Short title: Predicting hospital costs for patients receiving renal replacement therapy

Bernadette Li, John Cairns, James Fotheringham, Rommel Ramanan on behalf of the ATTOM Study Group

B. Li (corresponding author)

Department of Health Services Research and Policy

London School of Hygiene and Tropical Medicine

15-17 Tavistock Place, London WC1H 9SH, UK

E-mail: bernadette.li@lshtm.ac.uk

Tel: +44 (0)20 7958 8292

J. Cairns

Department of Health Services Research and Policy, London School of Hygiene and Tropical Medicine, London, UK

J. Fotheringham

Sheffield Kidney Institute, Sheffield, UK

R. Ramanan

Richard Bright Renal Unit, Southmead Hospital, Bristol, UK

A full list of members of the Access to Transplantation and Transplant Outcome Measures (ATTOM) Study Group is provided in the Acknowledgments.

Note: this is the author's accepted version of this article created by Bernadette Li. It reflects all changes made during the peer review process, but does not incorporate any modifications made at the proof stage. The final publication is available at Springer via <http://dx.doi.org/10.1007/s10198-015-0705-x>

ABSTRACT

Objective: To develop a model to predict annual hospital costs for patients with established renal failure, taking into account the effect of patient and treatment characteristics of potential relevance for conducting an economic evaluation, such as age, comorbidities and time on treatment. The analysis focuses on factors leading to variations in inpatient and outpatient costs and excludes fixed costs associated with dialysis, transplant surgery and high cost drugs.

Methods: Annual costs of inpatient and outpatient hospital episodes for patients starting renal replacement therapy in England were obtained from a large retrospective dataset. Multiple imputation was performed to estimate missing costs due to administrative censoring. Two-part models were developed using logistic regression to first predict the probability of incurring any hospital costs before fitting generalised linear models to estimate the level of cost in patients with positive costs. Separate models were developed to predict inpatient and outpatient costs for each treatment modality.

Results: Data on hospital costs were available for 15,869 incident dialysis patients and 4,511 incident transplant patients. The two-part models showed a decreasing trend in costs with increasing number of years on treatment, with the exception of dialysis outpatient costs. Age did not have a consistent effect on hospital costs, however, comorbidities such as diabetes and peripheral vascular disease were strong predictors of higher hospital costs in all four models.

Conclusion: Analysis of patient-level data can result in a deeper understanding of factors associated with variations in hospital costs and can improve the accuracy with which costs are estimated in the context of economic evaluations.

Key words: hospital costs, established renal failure, regression, patient-level data, two-part model

INTRODUCTION

Analysts involved in carrying out economic evaluations in healthcare are accustomed to expending considerable effort and resources to identify, collect, extrapolate and synthesise data to fully quantify the health consequences associated with different treatment approaches. However, when it comes to estimating costs, it is not uncommon to rely on readily available average unit costs that are assumed to apply uniformly to all patients or remain constant over time. If appropriate data sources can be identified, it would be beneficial to develop more precise ways to estimate the costs of managing patients with specific diseases and to explore in greater detail whether costs vary with patient and treatment characteristics of interest.

Treatment options for patients with established renal failure (ERF) include dialysis and transplantation. For many patients, transplantation can result in increased life expectancy and better quality of life compared to chronic dialysis [1]. Treatment of ERF is resource intensive for the health service. While costs of dialysis and transplantation may be comparable in the first year of treatment, costs for transplant recipients following surgery drop considerably in subsequent years, while the cost of maintenance dialysis sessions remains constant [2]. In England, payment to providers for dialysis is covered under a fixed national tariff as part of the Payment by Results (PbR) system. A similar approach is underway to introduce a fixed tariff for kidney transplant surgery. However beyond the provision of dialysis and transplant surgery, patients with ERF may incur additional hospital costs for monitoring of their condition, management of comorbidities or infections, maintenance of vascular access or post-operative follow-up. More than half of patients starting renal replacement therapy (RRT) have one or more comorbidities [3] which, alongside other factors such as age, may lead to variations in hospitalisation rates [4].

As part of the Access to Transplantation and Transplant Outcome Measures (ATTOM) study, an economic evaluation is being developed to compare alternative schemes for allocating kidneys to patients with ERF who are awaiting transplantation in the United Kingdom. Different approaches to kidney allocation can impact the length of time that patients with different characteristics will spend

on dialysis. This in turn can have an impact on the level of costs incurred. The objective of the current analysis is to develop a model to predict hospital inpatient and outpatient costs for patients with ERF, taking into account relevant patient and treatment characteristics such as age, comorbidities and time on treatment. The analysis will focus on characterising variations in hospital costs and therefore exclude fixed costs associated with routine dialysis, transplant surgery and high cost drugs. The approach to this analysis is guided by the intended use of the results as inputs for an economic evaluation that will compare the costs and consequences of alternative kidney allocation schemes.

METHODS

Data source

In England, all admissions to NHS hospitals are captured in the Hospital Episodes Statistics (HES) dataset. Patient demographics and information about type and length of stay are collected during a patient's time at hospital and are submitted to allow hospitals to be paid for the care they deliver [5]. Data on inpatient admissions have been routinely captured in HES since 1998 and outpatient attendances since 2003 [6]. The UK Renal Registry (UKRR) collects data provided by renal centres on all incident renal replacement therapy (RRT) patients, including demographics, comorbidity and treatment information [7]. In 2011, a pilot study was carried out in which UKRR data was linked to HES data for incident patients ≥ 18 years of age who started dialysis or received a kidney transplant between 2002 and 2006. HES only began collection of outpatient attendances in April 2003, therefore the sample for this analysis was restricted to those patients who started RRT between April 2003 and December 2006. The linked dataset captured hospital episodes until the end of December 2009 [6].

Linkage of HES data to UKRR data enhances the variables available in the separate datasets and facilitates analysis of hospital episodes by RRT modality. Taking the start of dialysis or date of transplant surgery as the index date, annual costs for each patient were generated by applying the appropriate 2011-12 PbR tariff to each inpatient admission (based on healthcare resource group) or outpatient appointment (based on treatment function code) [8]. Four separate datasets were created to capture dialysis inpatient, dialysis outpatient, transplant inpatient, and transplant outpatient costs in order to allow for the effect of explanatory variables to differ between regression models depending on treatment modality or type of hospital activity. The datasets included all inpatient admissions and outpatient appointments for any reason except routine dialysis or kidney transplant surgery. High cost drugs such as immunosuppressants following transplant surgery or drugs to treat renal anaemia were not captured in the dataset. The analysis therefore includes hospital costs that may not be specifically related to the management of patients' ERF. It was not considered feasible to distinguish between hospital episodes that were related versus those that were unrelated to the management of ERF in the

current analysis. However in economic evaluations, the focus is on the difference in costs between alternative strategies, therefore the inclusion of both related and potentially unrelated costs is appropriate, provided the same approach is taken for both the dialysis and transplant datasets.

Administrative censoring

Linkage of the HES and UKRR datasets came to an end in December 2009 and therefore no further data on hospital episodes were available beyond this date. This means that in any given year, some patients may only have observed costs for a portion of the year due to administrative censoring.

Rather than exclude these patients from the analysis, multiple imputation was performed to predict costs in the year that administrative censoring occurred under an assumption that data were missing at random (MAR). In the first instance, costs were imputed for the full year in which administrative censoring took place. However, since observed costs were available in these patients for part of the year up until the day of censoring, an additional step was taken to generate a hybrid imputed cost in order to make use of as much observed data as possible. Hybrid imputed costs were generated by using the imputed cost for the full year to calculate an imputed cost per day and multiplying this by the number of unobserved days for that year, to which the observed costs up until the day of censoring were then added.

Model development

Hospital costs in all four datasets were positively skewed with a varying proportion of zero-cost patients who had no inpatient admissions or outpatient visits in a given year. In order to accommodate these characteristics of the data, a two-part model approach for the regression analyses was taken [9-11]. Part one involved using logistic regression analysis to predict whether or not patients would incur any hospital costs. Part two involved fitting a generalised linear model (GLM) for those patients with positive costs [12,13]. The cluster option was used to take into account the dependence between multiple observations (years of cost data) for the same patient.

Initially age, sex, treatment modality, year since starting RRT, and co-morbidities were all entered into the regression models. Dummy variables were also entered for events including renal recovery, transplant and death in the dialysis models and for graft failure and death in the transplant models. Since costs are expected to be elevated for several months prior to death, inclusion of a dummy variable only in the year of death would not capture the full impact of this event on costs in patients who die at the beginning of the year. Therefore, an additional variable was created to indicate if death occurred in the first half of the following year. Backwards elimination was used to inform variable selection using a P-value threshold of 0.2 [14].

Model performance was assessed by comparing predicted and observed mean costs and calculating the root-mean-square error (RMSE) [10]. In addition, models that were developed based on multiply imputed values were compared to the results of complete-case analyses to provide reassurance of the validity of the MAR assumption.

All analyses were conducted in Stata (Version 13, Stata Corp, College Station, Texas, USA).

RESULTS

Data on inpatient admissions and outpatient appointments during the first year after initiation of RRT were available for 15,869 dialysis patients and 4,511 transplant patients. Administrative censoring occurred in approximately 11% of transplant patients in the first year after surgery and increased to more than 50% by year six. In contrast, no administrative censoring was present in the first three years of the dialysis patient sample, but ranged between 20% and 50% in years three through six. Tables 1a and 1b summarise the number of patients included in the dataset by number of years following initiation of RRT.

Table 1a Dialysis dataset: number of patients by years on dialysis

Years on dialysis	With complete year costs	Dialysis patients				TOTAL PATIENTS
		Death	Transplant	Recovered	Admin censoring	
1	11,894 (75%)	2,798 (17%)	750 (5%)	427 (3%)	0 (0%)	15,869
2	9,472 (80%)	1,488 (12%)	803 (7%)	123 (1%)	0 (0%)	11,886
3	7,501 (79%)	1,246 (13%)	634 (7%)	84 (1%)	0 (0%)	9,465
4	4,205 (56%)	1,063 (14%)	476 (6%)	40 (1%)	1,713 (23%)	7,497
5	1,932 (48%)	659 (16%)	248 (6%)	31 (1%)	1,188 (29%)	4,058
6	596 (33%)	274 (15%)	101 (6%)	5 (0%)	823 (46%)	1,799

Table 1b Transplant dataset: number of patients by years following transplant

Years following transplant	With complete year costs	Transplant patients			TOTAL PATIENTS
		Graft failure	Death	Admin censoring	
1	3,625 (80%)	266 (6%)	122 (3%)	498 (11%)	4,511
2	2,881 (80%)	116 (3%)	33 (1%)	585 (16%)	3,615
3	2,150 (75%)	48 (2%)	35 (1%)	644 (22%)	2,877
4	1,355 (63%)	38 (2%)	22 (1%)	735 (34%)	2,150
5	717 (53%)	16 (2%)	17 (1%)	605 (44%)	1,355
6	239 (33%)	9 (1%)	21 (3%)	448 (63%)	717

Part one: logistic regression analyses

Excluding patients with only part-year cost data, the proportion of patients with zero costs in the first year of RRT was lower in the outpatient setting (2% for dialysis patients and 1% for transplant

patients) than in the inpatient setting (24% for dialysis patients and 27% for transplant patients).

Logistic regression analyses showed that, compared to the first year of RRT, the odds of incurring any hospital costs in subsequent years was lower, with the exception of outpatient appointments for transplant patients (Tables 2a and 2b).

The presence of comorbidities was associated with higher odds of incurring inpatient costs in both dialysis and transplant patients, but the association was less consistent in the outpatient setting.

Table 2a Logistic regression analysis to predict whether or not dialysis patients incur any hospital costs

	n (%) patient-years	Dialysis inpatient		Dialysis outpatient	
		Odds ratio	95% CI	Odds ratio	95% CI
Constant		2.34*	(2.18, 2.51)	18.09*	(15.62, 20.95)
Age group					
< 50 years	10,608 (21%)	Reference		Reference	
50-64 years	13,330 (26%)	0.98	(0.91, 1.05)	1.26*	(1.08, 1.46)
65-75 years	15,393 (30%)	0.91*	(0.85, 0.97)	1.01	(0.88, 1.17)
> 75 years	11,243 (22%)	0.87*	(0.81, 0.94)	0.82*	(0.70, 0.96)
Sex					
Male	31,450 (62%)	Reference		-	-
Female	19,124 (38%)	1.10*	(1.05, 1.16)	-	-
Years on dialysis					
1	15,869 (31%)	Reference		Reference	
2	11,886 (23%)	0.59*	(0.56, 0.62)	0.80*	(0.73, 0.88)
3	9,465 (19%)	0.50*	(0.47, 0.52)	0.69*	(0.62, 0.76)
4	7,497 (15%)	0.58*	(0.54, 0.62)	0.76*	(0.67, 0.85)
5	4,058 (8%)	0.61*	(0.56, 0.67)	0.71*	(0.62, 0.82)
6	1,799 (4%)	0.65*	(0.57, 0.74)	0.72*	(0.59, 0.89)
Dialysis modality					
Haemodialysis	39,730 (79%)	Reference		Reference	
Peritoneal dialysis	10,844 (21%)	0.83*	(0.79, 0.88)	2.36*	(2.07, 2.69)
Comorbidities					
Myocardial infarction	8,347 (17%)	1.22*	(1.14, 1.31)	-	-
Congestive heart failure	8,801 (17%)	1.11*	(1.04, 1.19)	0.88*	(0.79, 0.98)
Peripheral vascular disease	8,204 (16%)	1.33*	(1.24, 1.42)	1.25*	(1.12, 1.41)
Cerebrovascular disease	5,459 (11%)	1.15*	(1.07, 1.24)	0.86*	(0.76, 0.97)
Pulmonary	7,351 (15%)	1.26*	(1.17, 1.35)	1.13*	(1.01, 1.27)
Liver	393 (1%)	-	-	-	-
Diabetes	19,167 (34%)	1.27*	(1.21, 1.34)	1.64*	(1.48, 1.81)
Cancer	4,092 (8%)	1.22*	(1.11, 1.33)	1.40*	(1.20, 1.63)
Hypertension	31,245 (62%)	1.09*	(1.04, 1.14)	1.36*	(1.23, 1.49)

Transplant	3,012 (6%)	1.11*	(1.02, 1.21)	0.25*	(0.21, 0.29)
Recovered renal function	710 (1%)	0.82*	(0.69, 0.96)	0.12*	(0.10, 0.15)
Death	7,528 (15%)	1.94*	(1.81, 2.07)	0.16*	(0.15, 0.18)
Death first half following year	2,521 (5%)	2.61*	(2.34, 2.92)	1.16	(0.93, 1.44)

*p<0.05

Table 2b Logistic regression analysis to predict whether or not transplant patients incur any hospital costs

	n (%) patient-years	Transplant inpatient		Transplant outpatient	
		Odds ratio	95% CI	Odds ratio	95% CI
Constant		1.89*	(1.65, 2.16)	104.02*	(72.08, 150.12)
Age group					
< 35 years	3,352 (22%)	Reference		-	-
36-45 years	3,950 (26%)	0.81*	(0.72, 0.92)	-	-
46-55 years	3,886 (25%)	0.73*	(0.64, 0.82)	-	-
> 55 years	4,037 (27%)	0.76*	(0.67, 0.87)	-	-
Sex					
Male	9,575 (63%)	Reference		Reference	
Female	5,650 (37%)	1.35*	(1.22, 1.49)	1.53*	(1.09, 2.16)
Years following transplant					
1	4,511 (29%)	Reference		Reference	
2	3,615 (24%)	0.21*	(0.19, 0.23)	1.17	(0.85, 1.62)
3	2,877 (19%)	0.18*	(0.16, 0.20)	1.60*	(1.06, 2.43)
4	2,150 (14%)	0.19*	(0.17, 0.22)	1.79*	(1.06, 3.04)
5	1,355 (9%)	0.19*	(0.16, 0.23)	1.08	(0.64, 1.84)
6	717 (5%)	0.18*	(0.14, 0.22)	1.06	(0.45, 2.51)
Transplant type					
Deceased donor	9,874 (65%)	Reference		Reference	
Living donor	5,351 (35%)	0.82*	(0.75, 0.90)	0.71	(0.49, 1.03)
Comorbidities					
Myocardial infarction	1,238 (8%)	1.47*	(1.24, 1.73)	-	-
Congestive heart failure	932 (6%)	1.48*	(1.22, 1.79)	-	-
Peripheral vascular disease	1,676 (11%)	1.87*	(1.62, 2.16)	1.56*	(1.00, 2.42)
Cerebrovascular disease	975 (6%)	1.38*	(1.16, 1.65)	-	-
Pulmonary	2,050 (13%)	1.24*	(1.09, 1.40)	-	-
Liver	119 (1%)	2.18*	(1.37, 3.47)	-	-
Diabetes	4,000 (26%)	1.62*	(1.46, 1.80)	1.80*	(1.21, 2.66)
Cancer	614 (4%)	1.62*	(1.31, 2.01)	3.45*	(1.32, 8.99)
Hypertension	11,251 (74%)	1.33*	(1.21, 1.46)	-	-
Graft failure	493 (3%)	-	-	0.02*	(0.02, 0.03)
Death	250 (2%)	1.62*	(1.14, 2.31)	0.02*	(0.01, 0.03)
Death first half following year	79 (0.5%)	4.55*	(2.47, 8.39)	-	-

*p<0.05

Part two: generalised linear models

Generalised linear models with an identity link function and gamma distribution were fitted to the subset of patients with non-zero costs. The model results shown in Tables 3a and 3b include imputed values that were generated using the hybrid approach to predict missing costs due to administrative censoring.

Mean inpatient costs were higher for dialysis patients compared to transplant patients with a trend towards decreasing costs in both patient groups over time. In contrast, mean outpatient costs were initially higher in the first year for transplant patients compared to dialysis patients, but decreased at a faster rate in subsequent years with dialysis outpatient costs overtaking those of transplant patients by the third year.

For dialysis patients, cost differed by treatment modality; haemodialysis was associated with higher costs in the inpatient setting, whereas peritoneal dialysis was associated with higher costs in the outpatient setting. Similarly in the transplant datasets, living donor transplants were associated with lower costs in the inpatient setting (although not statistically significant) and higher costs in the outpatient setting.

The increase in mean annual costs associated with various comorbidities ranged between £321 - £1,682 in the dialysis inpatient setting and between £264 and £2,093 in the transplant inpatient setting. Of the comorbidities included in the final models, peripheral vascular disease and diabetes were the only two that were consistently associated with significantly higher costs in both dialysis and transplant patients as well as in both inpatient and outpatient settings. The proportion of patients in both the dialysis and transplant datasets who had peripheral vascular disease was approximately 12% and the proportion who had diabetes was approximately 30%.

Table 3a Mean annual costs (£) for dialysis patients (generalised linear model)

	Dialysis inpatient		Dialysis outpatient	
	Coeff	95% CI	Coeff	95% CI
Constant	7782*	(7423, 8140)	1379*	(1331, 1428)
Age group				
< 50 years	Reference		Reference	
50-64 years	-170	(-489, 149)	-25	(-79, 29)
65-75 years	-181	(-513, 151)	-167*	(-219, -115)
> 75 years	-444*	(-806, -83)	-320*	(-376, -264)
Sex				
Male	Reference		-	-
Female	208*	(-23, 439)	-	-
Years on dialysis				
1	Reference		Reference	
2	-1189*	(-1487, -891)	-159*	(-186, -131)
3	-1434*	(-1729, -1140)	-112*	(-145, -80)
4	-1848*	(-2166, -1530)	-438*	(-85, -1)
5	-1709*	(-2099, -1319)	-13	(-66, 40)
6	-2270*	(-2774, -1767)	134*	(36, 232)
Dialysis modality				
Haemodialysis	Reference		Reference	
Peritoneal dialysis	-612*	(-838, -385)	334*	(296, 373)
Comorbidities				
Myocardial infarction	390*	(96, 683)	-	-
Congestive heart failure	321*	(58, 584)	-40	(-81, 0)
Peripheral vascular disease	721*	(423, 1019)	117*	(66, 168)
Cerebrovascular disease	506*	(174, 838)	-	-
Pulmonary	412*	(128, 696)	46	(0, 93)
Liver	1682*	(-161, 3524)	-	-
Diabetes	1191*	(929, 1453)	248*	(211, 284)
Cancer	-	-	139*	(72, 206)
Hypertension	-	-	-	-
Transplant	-1863*	(-2140, -1585)	-552*	(-602, -501)
Recovered renal function	1293*	(513, 2073)	-348*	(-454, -243)
Death	2403*	(2152, 2654)	-377*	(-414, -341)
Death first half following year	4415*	(3926, 4904)	200*	(138, 262)

*p<0.05

Table 3b Mean annual costs (£) for transplant patients (generalised linear model)

	Transplant inpatient		Transplant outpatient	
	Coeff	95% CI	Coeff	95% CI
Constant	4735*	(4331, 5138)	4053*	(3961, 4145)
Age group				
< 35 years	Reference		Reference	
36-45 years	-318	(-664, 29)	-123*	(-193, -53)
46-55 years	-310	(-676, 56)	-151*	(-224, -78)
> 55 years	-91	(-487, 306)	-126*	(-195, -57)
Sex				
Male	Reference		Reference	
Female	190	(-76, 455)	126*	(76, 175)
Years following transplant				
1	Reference		Reference	
2	-1576*	(-1881, -1271)	-2671*	(-2731, -2610)
3	-1919*	(-2228, -1611)	-2935*	(-3000, -2869)
4	-2138*	(-2485, -1790)	-3018*	(-3088, -2948)
5	-2061*	(-2502, -1620)	-3089*	(-3166, -3011)
6	-2654*	(-3212, -2096)	-3105*	(-3204, -3006)
Transplant type				
Deceased donor	Reference		Reference	
Living donor	-223	(-486, 39)	130*	(78, 182)
Comorbidities				
Myocardial infarction	641*	(145, 1138)	130*	(17, 242)
Congestive heart failure	1248*	(646, 1851)	159*	(35, 284)
Peripheral vascular disease	1222*	(729, 1715)	256*	(157, 354)
Cerebrovascular disease	898*	(271, 1524)	88	(-21, 197)
Pulmonary	264	(-87, 616)	179*	(99, 258)
Liver	2093*	(30, 4155)	524*	(200, 849)
Diabetes	1046*	(734, 1359)	593*	(515, 671)
Cancer	485*	(2, 969)	273*	(134, 411)
Hypertension	324*	(56, 592)	144*	(91, 197)
Graft failure	2438*	(1723, 3152)	-309*	(-451, -167)
Death	4924*	(3726, 6123)	-216*	(-426, -5)
Death first half following year	5725*	(3350, 8100)	629*	(321, 936)

*p<0.05

Model performance and predicted costs

Table 4 summarises observed and predicted mean annual cost estimates for each of the final two-part models. The results were compared with models that were developed based on complete-case

analyses, in which patients who were administratively censored were removed from the dataset. In each case, RMSE was found to be similar between the model based on complete-case analysis and the model that was developed using multiply imputed values.

Table 4 Observed and predicted mean annual costs

	Number of observations (patient-years)	Mean costs (std dev)	RMSE
Dialysis inpatient			
Observed		£ 5581 (9440)	
Two-part model complete-case analysis	46850	£ 5576 (2120)	9202.92
Two-part model hybrid imputed costs		£ 5578 (2136)	9204.73
Dialysis outpatient			
Observed		£ 1202 (1348)	
Two-part model complete-case analysis	46850	£ 1196 (343)	1291.19
Two-part model hybrid imputed costs		£ 1203 (345)	1291.22
Transplant inpatient			
Observed		£ 2398 (4675)	
Two-part model complete-case analysis	11710	£ 2390 (1931)	4278.45
Two-part model hybrid imputed costs		£ 2468 (1958)	4279.87
Transplant outpatient			
Observed		£ 2388 (2007)	
Two-part model complete-case analysis	11710	£ 2383 (1332)	1459.10
Two-part model hybrid imputed costs		£ 2447 (1386)	1458.98

As the motivation for the analysis was to predict annual hospital costs that can be used inputs in an economic evaluation, the appendix provides a worked example of how the regression results presented above can be used for this purpose.

DISCUSSION

Previous examples of economic evaluations that have compared dialysis and transplantation as treatment alternatives for patients with ERF have taken a variety of approaches to estimating costs. For transplant costs, it is common practice to estimate a cost for the first year of treatment that reflects the cost of surgery, and then assume a constant annual cost to capture resource use such as immunosuppressive therapy or outpatient visits in subsequent years [15-20]. For dialysis costs, some studies restrict the analysis to the cost of routine dialysis only, while others include the cost of hospitalisations, management of complications or drugs. Other than taking into account the cost of vascular access at the start of dialysis, annual costs for dialysis patients are often assumed to be constant [18,16]. However, there are examples of economic evaluations that have introduced an element of variation in costs among dialysis patients by considering factors such as age or time on treatment: de Wit et al [19] presented separate estimates of hospital costs by age group based on data collected at 13 Dutch dialysis centres; Haller et al [15] analysed patient-level cost data from a hospital in Austria and presented separate cost estimates for dialysis patients in the first year, second year and subsequent years of treatment. None of these previous studies have simultaneously considered the impact of treatment modality, length of time on treatment, age and comorbidities on costs.

Collection of patient-level cost data is a resource intensive exercise. The linkage of HES data to UKRR data provides a rare opportunity to analyse a large existing dataset to explore variations in hospital costs specifically among patients receiving RRT in England. HES is, to our knowledge, the most complete source of routinely collected information on admissions and attendances at NHS hospitals in England and linkage to UKRR data facilitates simultaneous exploration of multiple patient and treatment-related factors that may affect costs. The approach to analysing the linked dataset was guided by both the features of the data and the intended use of the results. In this case, the primary objective of the analysis was to predict annual costs for patients with different characteristics for use as inputs in an economic evaluation. Additionally, there were three main features of the cost data that needed to be addressed: 1) missing data due to administrative censoring, 2) the proportion of observations with zero costs and 3) positively skewed distributions.

Multiple imputation was carried out to address the issue of administrative censoring. Multiple imputation has the advantage of making use of all available observed data, while allowing for uncertainty about the missing values [21,22]. In this analysis, the models based on complete cases and the models that included imputed values were very similar, providing confidence that missing data due to administrative censoring did not bias the estimates of cost. This suggests that a complete-case analysis would have been sufficient, but it is unclear if this conclusion can be generalised beyond our dataset. There is a growing body of literature describing other methods to address the common issue of censoring of cost data [23-27]. However, given that the primary objective of the current analysis was to estimate annual (as opposed to lifetime) costs and that cost histories detailing the timing of individual hospital events were not available in the current extract of the dataset, approaches based on survival analysis techniques were not pursued.

The issues of zero costs and positively skewed distributions were addressed by adopting a two-part approach in which a logistic regression was fitted to predict the probability of incurring any hospital costs, followed by fitting a GLM to estimate the level of cost for patients with at least one admission or visit. A potential advantage of the two-part approach is that covariates that are determined to be significant in part one of the model do not have to be the same as those that determine the level of cost in part two. In the present analyses, there was general consistency in terms of the covariates that were included in part one and part two of the final inpatient cost models, but less agreement in the outpatient setting where the percentage of zero costs was lower.

The results of this analysis highlight a number of findings that are relevant when considering variations in hospital costs for patients on RRT in the context of economic evaluation. Firstly, while the cost of transplant surgery can be viewed as a one-time event and the cost of maintenance dialysis sessions generally remains constant from week to week and year to year, hospital costs for patients on RRT showed a decreasing trend over time that extended beyond the first two years on RRT. A plausible explanation for this trend is that patients who survive longer on therapy are on average fitter and healthier and required fewer hospital visits. Secondly, age did not have a consistent effect on costs across all treatment modalities and hospital settings and, in contrast to the approach taken in the

economic evaluation by de Wit et al [19], the current analysis suggests that, controlling for other factors, increasing age alone may be associated with lower rather than higher costs. Thirdly, many of the comorbidities included in the analysis were found to be significant predictors of hospital costs and had a bigger impact than age in the estimation of costs for patients on RRT.

In the absence of evidence to the contrary, it is perhaps most natural to adopt an assumption that costs remain constant either over time or between subgroups of patients with different characteristics.

However, if appropriate patient-level data sources can be identified, a more detailed understanding of patient characteristics and treatment factors that influence costs can help improve the accuracy with which costs are estimated in the context of economic evaluations.

APPENDIX: A WORKED EXAMPLE TO PREDICT HOSPITAL COSTS BASED ON THE FINAL TWO-PART MODEL

To estimate annual inpatient costs for a 55-year-old male patient with diabetes who has been on haemodialysis for three years:

Part 1: probability of incurring any inpatient cost > £0

Taking the natural log of the odds ratios in Table 2a, calculate log odds of incurring any inpatient cost

$$\begin{aligned} & \text{CONSTANT} + (\beta_1 \times \text{AGEGROUP50} - 64) + (\beta_2 \times \text{YEAR3}) + (\beta_3 \times \text{DIABETES}) \\ & = 0.850 + (-0.022 \times 1) + (-0.702 \times 1) + (0.242 \times 1) = 0.368 \end{aligned}$$

Calculate probability from log odds

$$e^{x\beta} / (1 + e^{x\beta}) = e^{0.368} / (1 + e^{0.368}) = 0.591$$

Part 2: estimate level of inpatient cost based on coefficients in Table 3a

$$\begin{aligned} & \text{CONSTANT} + (\beta_1 \times \text{AGEGROUP50} - 64) + (\beta_2 \times \text{YEAR3}) + (\beta_3 \times \text{DIABETES}) \\ & = 7782 + (-170 \times 1) + (-1434 \times 1) + (1191 \times 1) = 7368 \end{aligned}$$

Combine parts 1 and 2: multiply estimated level of inpatient cost by probability of incurring any cost

$$7368 \times 0.591 = \text{£}4,354$$

ACKNOWLEDGMENTS

This article presents independent research commissioned by the National Institute for Health Research (NIHR) under the Programme Grant for Applied Research (RP-PG-0109-10116) entitled Access to Transplantation and Transplant Outcome Measures (ATTOM). The views expressed in this publication are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. JF was funded by a Kidney Research UK Clinical Training Fellowship. The ATTOM Study Group comprises: J. Andrew Bradley, Clare Bradley, John Cairns, Heather Draper, Chris Dudley, John L. Forsythe, Damian G. Fogarty, Rachel J. Johnson, Geraldine Leydon, Wendy Metcalfe, Gabriel C. Oniscu, Rommel Ramanan, Paul Roderick, Charles R. Tomson and Christopher Watson. The authors are grateful to the UK Renal Registry and Hospital Episode Statistics for the linked dataset. Hospital Episode Statistics: Copyright 2014, re-used with the permission of the Health and Social Care Information Centre. All rights reserved.

REFERENCES

1. Tonelli, M., Wiebe, N., Knoll, G., Bello, A., Browne, S., Jadhav, D., Klarenbach, S., Gill, J.: Systematic review: kidney transplantation compared with dialysis in clinically relevant outcomes. *Am. J. Transplant.* **11**(10), 2093-2109 (2011).
2. Department of Health. The National Service Framework for renal services. Part one: dialysis and transplantation.
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/199001/National_Service_Framework_for_Renal_Services_Part_One_-_Dialysis_and_Transplantation.pdf (2004). Accessed 16 July 2014.
3. Rao, A., Steenkamp, R., Caskey, F.: UK Renal Registry 16th annual report: Chapter 5 comorbidities and current smoking status amongst patients starting renal replacement therapy in England, Wales and Northern Ireland from 2011 to 2012. *Nephron. Clin. Pract.* **125**, 99-110 (2013).
4. Metcalfe, W., Khan, I.H., Prescott, G.J., Simpson, K., Macleod, A.M.: Hospitalization in the first year of renal replacement therapy for end-stage renal disease. *QJM* **96**(12), 899-909 (2003).
5. Hospital Episode Statistics, <http://www.hscic.gov.uk/hes>. Accessed 2 April 2015.
6. Fotheringham, J., Fogarty, D., Jacques, R., El Nahas, M., Campbell, M.: Chapter 13 The linkage of incident renal replacement therapy patients in England (2002-2006) to hospital episodes and national mortality data: improved demography and hospitalisation data in patients undergoing renal replacement therapy. *Nephron. Clin. Pract.* **120 Suppl 1**, c247-260 (2012).
7. UK Renal Registry, <https://www.renalreg.org/datasets/the-uk-renal-registry-dataset/>. Accessed 2 April 2015.
8. Department of Health. Confirmation of payment by results (PbR) arrangements for 2011-12, 2011-12 tariff information spreadsheet.
<http://webarchive.nationalarchives.gov.uk/20130507170152/https://www.gov.uk/government/publications/confirmation-of-payment-by-results-pbr-arrangements-for-2011-12>. Accessed 21 January 2014.

9. Mullahy, J.: Much ado about two: reconsidering retransformation and the two-part model in health econometrics. *J. Health. Econ.* **17**(3), 247-281 (1998).
10. Dunn, G., Mirandola, M., Amaddeo, F., Tansella, M.: Describing, explaining or predicting mental health care costs: a guide to regression models. Methodological review. *Br. J. Psychiatry* **183**, 398-404 (2003).
11. Buntin, M.B., Zaslavsky, A.M.: Too much ado about two-part models and transformation? Comparing methods of modeling Medicare expenditures. *J. Health. Econ.* **23**(3), 525-542 (2004).
12. Barber, J., Thompson, S.: Multiple regression of cost data: use of generalised linear models. *J. Health Serv. Res. Policy* **9**(4), 197-204 (2004).
13. Hardin, J.W.H., J.M.: *Generalized linear models and extensions*, 2nd ed. Stata Press, Lakeway Drive, Texas (2007).
14. Kirkwood, B.R., Sterne, J.A.C.: *Essential Medical Statistics*, 2nd ed. Blackwell Publishing Oxford (2003).
15. Haller, M., Gutjahr, G., Kramar, R., Harnoncourt, F., Oberbauer, R.: Cost-effectiveness analysis of renal replacement therapy in Austria. *Nephrol. Dial. Transplant.* **26**(9), 2988-2995 (2011).
16. Howard, K., Salkeld, G., White, S., McDonald, S., Chadban, S., Craig, J.C., Cass, A.: The cost-effectiveness of increasing kidney transplantation and home-based dialysis. *Nephrology* **14**(1), 123-132 (2009).
17. Jassal, S.V., Krahn, M.D., Naglie, G., Zaltzman, J.S., Roscoe, J.M., Cole, E.H., Redelmeier, D.A.: Kidney transplantation in the elderly: a decision analysis. *J. Am. Soc. Nephrol.* **14**(1), 187-196 (2003).
18. Kontodimopoulos, N., Niakas, D.: An estimate of lifelong costs and QALYs in renal replacement therapy based on patients' life expectancy. *Health Policy* **86**, 85-96 (2008).
19. de Wit, G.A., Ramsteijn, P.G., de Charro, F.T.: Economic evaluation of end stage renal disease treatment. *Health Policy* **44**, 215-232 (1998).
20. Wong, G., Howard, K., Chapman, J.R., Chadban, S., Cross, N., Tong, A., Webster, A.C., Craig, J.C.: Comparative survival and economic benefits of deceased donor kidney transplantation and dialysis in people with varying ages and co-morbidities. *PLoS ONE* **7**(1) (2012).

21. Briggs, A., Clark, T., Wolstenholme, J., Clarke, P.: Missing... presumed at random: cost-analysis of incomplete data. *Health Econ.* **12**(5), 377-392 (2003).
22. Sterne, J.A., White, I.R., Carlin, J.B., Spratt, M., Royston, P., Kenward, M.G., Wood, A.M., Carpenter, J.R.: Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* **338**, b2393 (2009).
23. Bang, H.T., Tsiatis, A.A.: Estimating medical costs with censored data. *Biometrika* **87**, 329-343 (2000).
24. Carides, G.W., Heyse, J.F., Iglewicz, B.: A regression-based method for estimating mean treatment cost in the presence of right-censoring. *Biostatistics* **1**(3), 299-313 (2000).
25. Lin, D.Y., Feuer, E.J., Etzioni, R., Wax, Y.: Estimating medical costs from incomplete follow-up data. *Biometrics* **53**(2), 419-434 (1997).
26. O'Hagan, A., Stevens, J.W.: On estimators of medical costs with censored data. *J. Health Econ.* **23**(3), 615-625 (2004).
27. Young, T.A.: Estimating mean total costs in the presence of censoring: a comparative assessment of methods. *Pharmacoeconomics* **23**(12), 1229-1242 (2005).