Supplemental material

Evaluating integrated care for people with complex needs

***Panagiotis Kasteridis**, Research Fellow, Centre for Health Economics, University of York, UK*

***Anne Mason**, Senior Research Fellow, Centre for Health Economics, University of York, UK*

***Andrew Street**, Professor, Department of Health Policy, London School of Economics and Political Science, UK*

Correspondence to: Panagiotis Kasteridis, Centre for Health Economics, University of York, Heslington, York, YO10 5DD, UK, Phone: (+44) 01904 321429, E-mail: panos.kasteridis@york.ac.uk

Appendix 1: Econometric methods

We employ propensity score matching to match those exposed to the IC interventions to out-of-area controls. Let N1 be the number of individuals in South Somerset receiving the particular IC intervention, N2 the number of individuals in South Somerset not receiving IC, and N3 the number of individuals in the other parts of Somerset. The aim is to match the N1 exposed cases with unexposed controls from the pool of N3 out of area potential controls. The matching process involves three steps.

In step 1, for each of the (N1+N2) individuals in South Somerset the probability of being in the intervention group, given a set of individual characteristics, is predicted using a logistic model.

$$p_i^{SS} = \frac{\exp(X_i\beta)}{1+\exp(X_i\beta)} \quad (1)$$

We use the N1 predicted probabilities to identify unexposed controls in step 3.

In step 2, we use the estimated coefficients from equation (1) to predict propensity scores for out of area individuals in pool N3.

$$p_i^{OOA} = \frac{\exp(Z_i\hat{\beta})}{1+\exp(Z_i\hat{\beta})} \quad (2)$$

We use all N3 predicted probabilities from this step in step 3.

In step 3, a matching algorithm is used to match propensity scores from steps 1 and 2. Specifically we use 'Single Nearest Neighbour without replacement' which selects as a

match an unexposed control whose propensity score is the closest to the exposed case. This unexposed control is then removed from subsequent matches for the cohort. Therefore for each cohort we obtain N1 matched pairs. Note that, as the matching algorithm is run separately for each 6-month cohort, individuals selected as controls for one cohort may also be selected for one or more other cohorts.

We also subjected our analysis to the following propensity score matching algorithms: single nearest neighbour with replacement, nearest neighbours with k neighbours, nearest neighbour with calliper, and coarsened exact matching. However, none of these had material impact on the results.

In 2014/15 and 2015/16, information on long term conditions was reported quarterly so we imputed monthly values from the January, March, June and December values. Some cases had no midpoint data available so values were imputed for these individuals where possible, otherwise the case was dropped.

Difference-in-differences (DiD) regression analysis

Difference-in-differences (DiD) regression analysis is coupled with propensity score matching to compare the mean utilisation before and after the intervention for the matched case and control groups. We implement a simple DID design with a single IC intervention and two time periods - pre and post enrolment:

$$y_{it} = a + \beta POST_t + \gamma(IC_i) + \delta(DiD_{it}) + k(SPQS_i) + u_{it} \quad (3)$$

where $i = 1, \ldots, N$ indicates the individual patients, $t = 1, 2$ indicates the pre-enrolment and post-enrolment periods and y_{it} is a specific service utilisation measure for patient i in period t (i.e. monthly values are aggregated over period t), *POST* is a binary indicator which takes a value of 1 in the post enrolment period, *IC* is a binary indicator for whether an individual belongs to the intervention group, *DiD* is the interaction term which takes a value of 1 for intervention patients in the post enrolment period and 0 otherwise, and $u_i \sim N(0, \sigma_u^2)$. We also account for whether the patient's practice was a participant in the Somerset Quality Practice Scheme (*SPQS*), a local alternative to the national Quality and Outcomes Framework. The *DiD* estimate δ captures the change in outcome between the two periods for the intervention patients minus the change in outcome between the same two time periods for the non-intervention patients, that is:

$$\delta = [E(y_{T=1} \mid X, IC = 1) - E(y_{T=0} \mid X, IC = 1)] - [E(y_{T=1} \mid X, IC = 0) - E(y_{T=0} \mid X, IC = 0)] \quad (4)$$

For the EPC cohorts, the following model was estimated as a sensitivity analysis to account for prior CCT use by EPC patients:

$$y_{it} = a + \beta POST_t + \gamma_1(EPC_i) + \gamma_2(DUAL_i) + \delta_1(DiD^1{}_{it}) + \delta_2(DiD^2{}_{it}) + k(SPQS_i) + u_{it} \quad (5)$$

In the above equation, the binary variable EPC indicates whether the individual was in the EPC group (but not in the CCT group). DUAL is a binary indicator of whether the individual was in both CCT and EPC groups. DiD^1 and DiD^2 are interactions of EPC and DUAL with the POST variable.

Test for the parallel trends assumption

Visual inspection of the pre-intervention trends (left shaded areas in figures 1 and 2) shows that in most cases the parallel trends assumption of the DiD models holds true. We tested the hypothesis that the outcome trends of the cases and controls are parallel prior to intervention by using pre-intervention monthly data to estimate a model that allows for group specific trends. For instance, to test the assumptions for CCT cohort 1 we estimate the following model using the 11 months prior to intervention (April 2014 – February 2015).

$$y_{it} = a + \beta_2 t_2 + \ldots + \beta_{11} t_{11} + \gamma(IC_i) + \delta_2(t_2 * IC_i) + \ldots + \delta_{11}(t_{11} * IC_i) + u_{it} \quad (6)$$

The coefficient γ captures the difference in outcome between cases and controls in the first pre-intervention month (11 months prior to intervention) while the coefficients $\delta_2, \ldots, \delta_{11}$ capture differences in the trends of cases and controls. The null hypothesis that the δ-coefficients are jointly zero is a test of the parallel trends assumption.

In total we estimated 245 δ-coefficients for the 5 CCT cohorts and 5 outcomes. Of those only 17 were significant at 5% significance level. The F-test rejected the null hypothesis of jointly zero coefficients at 5% level in 4 models: outpatient visits for cohorts 2 and 3 and A&E and non-elective admissions for cohort 5. At 1% level the null hypothesis was rejected only for A&E attendances in cohort 5.

We estimated 51 δ-coefficients for the 2 EPC cohorts and 3 outcomes. Of those only 1 was significant at 5% significance level. The null hypothesis of jointly zero coefficients was not rejected in any of the models at 5% level.

Appendix 2: Descriptive statistics and balance graphs

Balance graphs

Descriptive statistics comparing cases and controls on the matching variables at baseline (midpoint) are in Tables A2.1 and A2.2.

Figure A2.1 shows the balance graphs from the matching of CCT cohorts with the corresponding graphs for EPC cohorts in Figure A2.2. These compare matching variables for the 6-month cohorts for cases and controls, with the black dots showing the pre-matching comparability (unmatched controls), and the crosses showing the balance after matching (matched controls). If a characteristic is perfectly matched, its cross will lie on the vertical line. If the cross lies to the right (left) of the vertical line this indicates a positive (negative) bias.

CCT cohort 1 (top left hand corner) compares 86 cases with 86 controls. For each characteristic, the cross is closer to the line of zero bias than its corresponding black dot. This shows that, in general, the matching has worked, implying that the cases and controls are well balanced. However, cases were on average younger (as the cross lies to the left of the line of zero bias) than controls. Over the 12 months before enrolment, cases typically had fewer prescriptions and more GP contacts than controls. Cases in cohort 1 were more likely to live in a care home (the cross is on the right of the vertical line). For the CCT cohorts in general, cases lived in less deprived areas. Overall, the descriptive statistics and balance graphs show that matching has worked well for both care models.

Table A2.1: Matching variables: baseline comparability of CCT groups

CCT1	control (N=86)		case (N=86)	
	mean	sd	mean	sd
Count of prescriptions in last 12m	25.08	14.06	23.64	13.57
Count of GP practice contacts in last 12m	28.83	23.04	29.08	20.51
RISC score	18.23	11.43	18.57	11.67
Age (years)	78.81	10.08	76.03	10.21
Measure of deprivation	16.41	8.70	15.18	8.14
sex = male	0.44	0.50	0.52	0.50
Count of 8 LTCs	2.53	1.40	2.67	1.48
Resident in a care home	0.06	0.24	0.07	0.26

CCT2	control (N=47)		case (N=47)	
	mean	sd	mean	sd
Count of prescriptions in last 12m	23.91	11.10	22.28	10.18
Count of GP practice contacts in last 12m	29.81	19.16	28.87	16.24
RISC score	19.63	10.71	19.97	12.08
Age (years)	74.96	12.11	75.30	13.04
Measure of deprivation	21.14	10.99	16.68	7.83
sex = male	0.68	0.47	0.72	0.45
Count of 8 LTCs	2.60	1.35	2.68	1.58
Resident in a care home	0.04	0.20	0.04	0.20

CCT3	control (N=90)		case (N=90)	
	mean	sd	mean	sd
Count of prescriptions in last 12m	20.27	10.63	21.28	12.85
Count of GP practice contacts in last 12m	30.17	22.43	29.23	20.28
RISC score	18.22	10.86	18.87	11.51
Age (years)	77.34	11.43	73.41	14.52
Measure of deprivation	19.73	10.98	16.55	9.81
sex = male	0.44	0.50	0.49	0.50
Count of 8 LTCs	2.19	1.27	2.34	1.44
Resident in a care home	0.01	0.11	0.03	0.18

CCT4	control (N=209)		case (N=209)	
	mean	sd	mean	sd
Count of prescriptions in last 12m	17.94	10.35	18.47	10.36
Count of GP practice contacts in last 12m	21.17	15.89	22.84	16.45
RISC score	14.19	11.57	15.42	11.90
Age (years)	78.02	11.94	73.33	15.38
Measure of deprivation	19.07	10.23	17.42	9.34
sex = male	0.40	0.49	0.47	0.50
Count of 8 LTCs	1.75	1.37	1.78	1.32
Resident in a care home	0.06	0.24	0.06	0.23

CCT5	control (N=132)		case (N=132)	
	mean	sd	mean	sd
Count of prescriptions in last 12m	19.76	11.54	19.05	9.04
Count of GP practice contacts in last 12m	29.64	20.87	31.46	18.05
RISC score	17.17	12.81	17.52	11.60
Age (years)	77.69	11.16	74.52	16.69
Measure of deprivation	19.65	9.47	16.99	9.49
sex = male	0.43	0.50	0.48	0.50
Count of 8 LTCs	2.28	1.52	2.17	1.43
Resident in a care home	0.06	0.24	0.10	0.30

Table A2.2: Matching variables: baseline comparability of EPC groups

EPC1	control (N=603)		case (N=603)	
	mean	sd	mean	sd
Count of prescriptions in last 12m	13.17	9.89	13.85	9.24
Count of GP practice contacts in last 12m	16.52	14.50	18.50	13.36
RISC score	7.06	8.15	7.94	8.96
Age (years)	70.88	15.19	67.91	17.49
Measure of deprivation	18.70	10.42	16.89	8.68
sex = male	0.32	0.47	0.36	0.48
Count of 8 LTCs	1.37	1.25	1.37	1.28
Resident in a care home	0.01	0.09	0.01	0.11

EPC2	control (N=231)		case (N=231)	
	mean	sd	mean	sd
Count of prescriptions in last 12m	10.22	7.52	10.54	7.23
Count of GP practice contacts in last 12m	13.06	10.76	15.48	9.38
RISC score	4.82	5.97	5.12	6.97
Age (years)	68.25	14.42	64.68	18.64
Measure of deprivation	20.14	10.65	17.44	10.60
sex = male	0.32	0.47	0.32	0.47
Count of 8 LTCs	1.04	1.07	1.06	1.12
Resident in a care home	0.01	0.09	0.00	0.07

Figure A2.1: Balance graphs of covariates used for matching: CCT

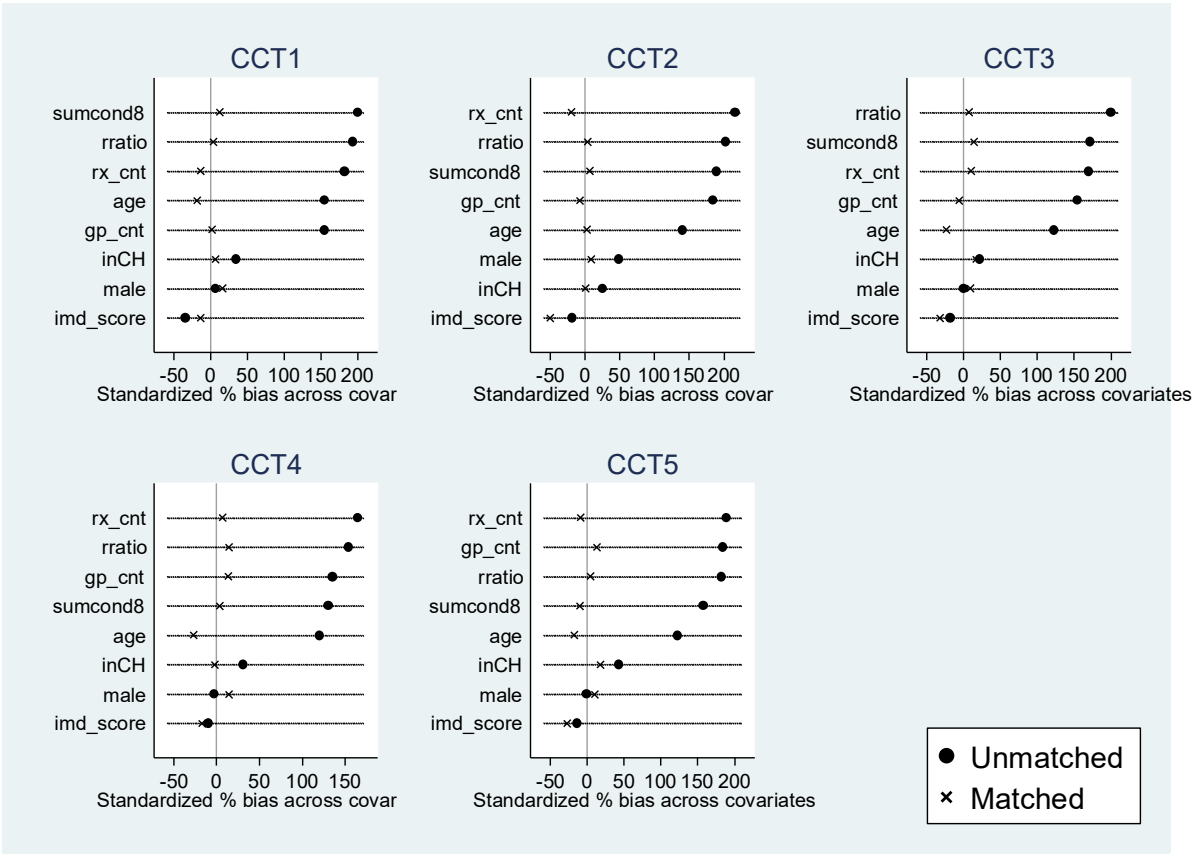
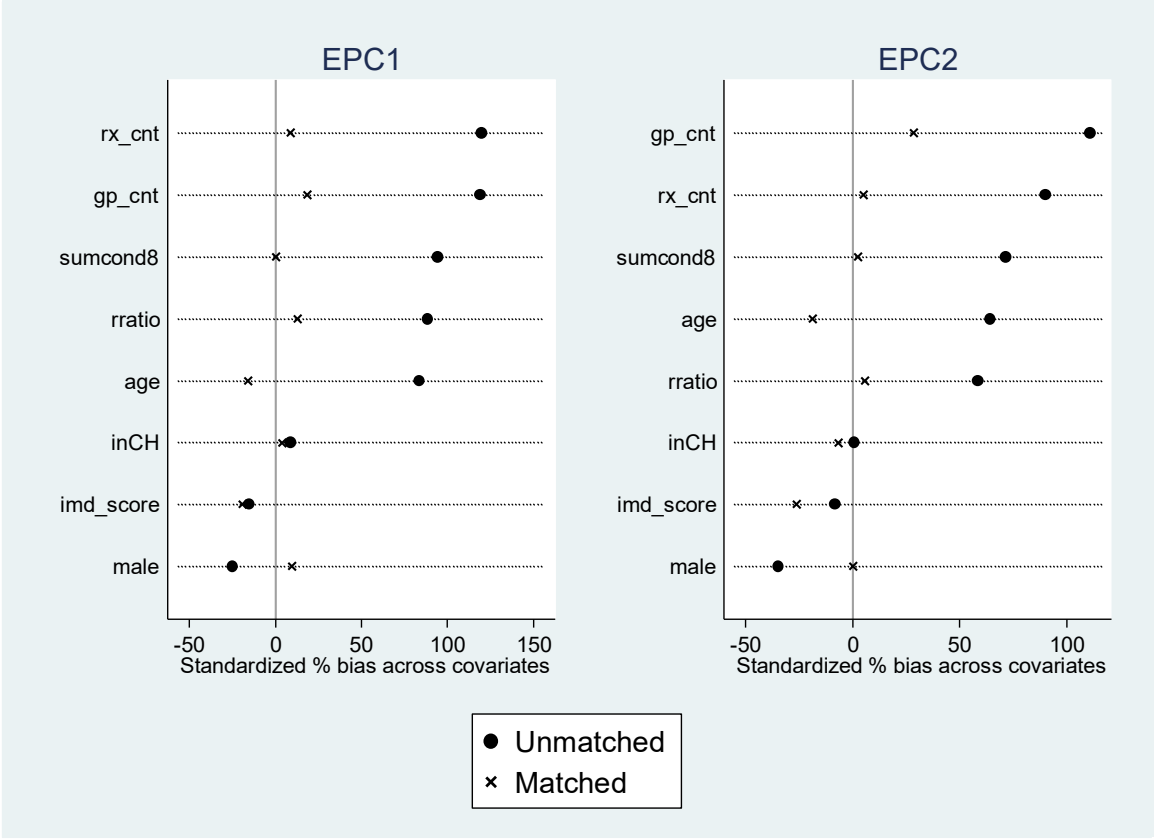


Figure A2.2: Balance graphs of covariates used for matching: EPC



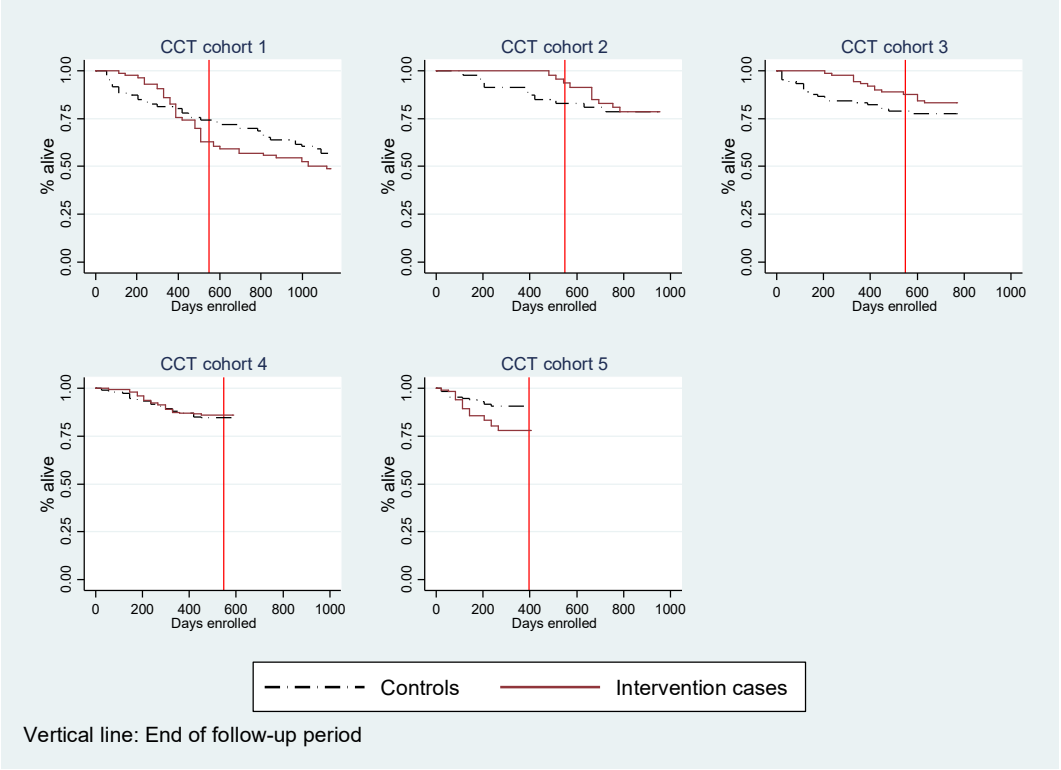
Appendix 3: Sensitivity analysis

Table A3.1: Death rates in CCT and EPC cases and controls

		Death=0	Death=1	% died
CCT1				
	Controls	49	37	43.02
	Intervention cases	42	44	51.16
CCT2				
	Controls	37	10	21.28
	Intervention cases	37	10	21.28
CCT3				
	Controls	70	20	22.22
	Intervention cases	75	15	16.67
CCT4				
	Controls	177	32	15.31
	Intervention cases	180	29	13.88
CCT5				
	Controls	120	12	9.09
	Intervention cases	103	29	21.97
EPC1				
	Controls	545	58	9.62
	Intervention cases	568	35	5.8
EPC2				
	Controls	229	2	0.87
	Intervention cases	224	7	3.03

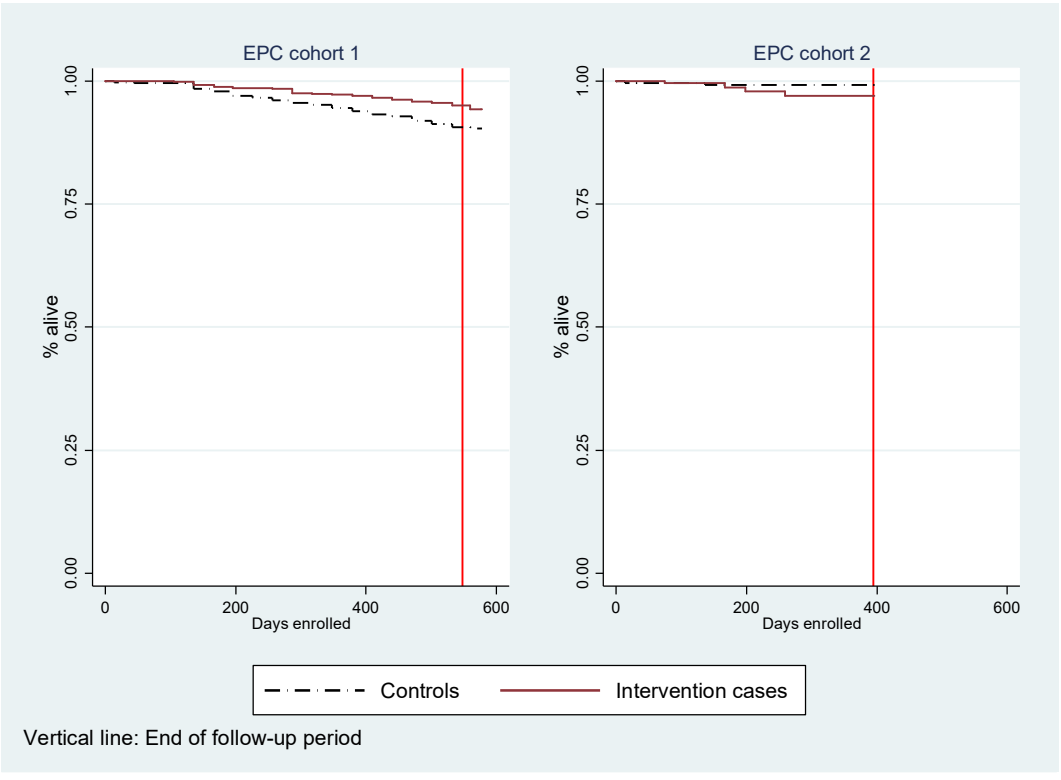
Note: based on data to March 2018

Figure A3.1: Survival curves for CCT cases and controls



Note: the end of follow-up is 12 months except for CCT5 (7 months)

Figure A3.2: Survival curves for EPC cases and controls



Note: the end of follow-up is 12 months (EPC1) or 7 months (EPC2)

Table A3.2: DiD results for CCT cohorts – sensitivity analysis for survivors

6-month cohort	Cases used in the base case analysis	Cases used in the analysis of survivors
CCT1	86	54
CCT2	47	44
CCT3	90	79
CCT4	209	172
CCT5	132	95
EPC1	603	573
EPC2	231	222

Table A3.3: DiD results for CCT cohorts – sensitivity analysis for survivors

Cohort 1 (N=216)	Acute Outpatient visits		A&E attendances		Non-elective admissions		Bed days		Total Monthly Cost	
	Coef	P-val	Coef	P-val	Coef	P-val	Coef	P-val	Coef	P-val
POST	-1.22	0.332	-0.82	0.066	-0.30	0.475	2.07	0.697	-196	0.946
IC	2.35	0.069	-0.21	0.635	-0.20	0.646	-1.69	0.757	2,815	0.341
DiD	0.52	0.771	0.63	0.314	0.19	0.752	0.24	0.975	237	0.954
SPQS	-1.11	0.456	0.06	0.913	0.05	0.913	1.02	0.871	4,263	0.211

Cohort 2 (N=176)	Acute Outpatient visits		A&E attendances		Non-elective admissions		Bed days		Total Monthly Cost	
	Coef	P-val	Coef	P-val	Coef	P-val	Coef	P-val	Coef	P-val
POST	-1.48	0.349	0.32	0.705	-0.52	0.418	-5.16	0.255	244	0.926
IC	**3.68**	0.025	0.58	0.502	-0.001	0.999	1.41	0.764	5,151	0.059
DiD	-0.07	0.976	-0.46	0.702	1.07	0.243	8.68	0.176	850	0.819
SPQS	1.09	0.562	0.53	0.598	0.40	0.601	-3.40	0.529	-91	0.977

Cohort 3 (N=316)	Acute Outpatient visits		A&E attendances		Non-elective admissions		Bed days		Total Monthly Cost	
	Coef	P-val	Coef	P-val	Coef	P-val	Coef	P-val	Coef	P-val
POST	-0.56	0.707	-0.76	0.251	-0.34	0.369	-1.79	0.551	-440	0.841
IC	**3.07**	0.044	-0.15	0.826	**0.87**	0.026	**7.73**	0.012	**4,884**	0.030
DiD	2.19	0.297	0.14	0.882	-0.01	0.981	-2.96	0.484	1,365	0.660
SPQS	-1.06	0.561	0.05	0.946	-0.68	0.148	0.51	0.889	-235	0.930

Cohort 4 (N=688)	Acute Outpatient visits		A&E attendances		Non-elective admissions		Bed days		Total Monthly Cost	
	Coef	P-val	Coef	P-val	Coef	P-val	Coef	P-val	Coef	P-val
POST	-0.24	0.716	-0.38	0.367	-0.36	0.138	-3.06	0.163	-493	0.760
IC	**1.45**	0.040	0.25	0.571	0.32	0.215	1.98	0.389	1,201	0.477

DiD	0.86	0.365	-0.08	0.900	0.06	0.852	2.50	0.420	**5,594**	0.015
SPQS	*-1.89*	0.012	0.14	0.770	-0.09	0.743	-1.49	0.544	-438	0.809

Cohort 5 (N=380)	Acute Outpatient visits		A&E attendances		Non-elective admissions		Bed days		Total Monthly Cost	
	Coef	P-val	Coef	P-val	Coef	P-val	Coef	P-val	Coef	P-val
POST	-1.14	0.258	-0.63	0.115	-0.40	0.129	-3.57	0.120	-2,042	0.156
IC	1.07	0.305	0.57	0.169	0.34	0.211	2.53	0.286	823	0.580
DiD	-0.48	0.733	0.42	0.456	0.31	0.412	1.24	0.702	**4,514**	0.027
SPQS	-0.92	0.446	0.14	0.771	-0.11	0.721	0.70	0.801	212	0.902

Note: significant (p<0.05) results in bold. N is the number of observations (4 observations per patient).

Table A3.4: DiD results for EPC cohorts - sensitivity analysis for survivors

Cohort 1 (N=2,292)	Acute Inpatient		Acute Outpatient		Proportion of out of hospital costs	
	Coef	P-val	Coef	P-val	Coef	P-val
POST	-0.09	0.681	-0.07	0.822	**-0.46**	0.015
IC	0.33	0.165	**1.74**	<0.001	**1.24**	<0.001
DiD	0.39	0.225	0.36	0.435	0.31	0.251
SPQS	-0.08	0.742	-0.52	0.147	0.01	0.956

Cohort 2 (N=888)	Acute Inpatient		Acute Outpatient		Proportion of out of hospital costs	
	Coef	P-val	Coef	P-val	Coef	P-val
POST	0.01	0.881	-0.04	0.905	-0.16	0.291
IC	0.12	0.212	0.47	0.134	0.16	0.330
DiD	-0.01	0.972	-0.02	0.958	0.41	0.060
SPQS	-0.02	0.865	-0.07	0.843	0.06	0.765

Note: significant (p<0.05) results in bold. N is the number of observations (4 observations per patient).

Table A3.5: DiD results for EPC cohort – sensitivity analysis to account for prior CCT use

Cohort 1 (N=2,412)	Acute inpatient		Acute Outpatient		Proportion of out of hospital costs	
	Coef	P-val	Coef	P-val	Coef	P-val
POST	-0.28	0.130	-0.56	0.067	**-1.23**	<0.001
IC	-0.23	0.255	0.22	0.499	0.43	0.021
IC-dual	**3.55**	<0.001	**8.67**	<0.001	**10.87**	<0.001
DiD	0.32	0.246	0.66	0.139	**0.82**	0.001
DiD-dual	**1.69**	0.002	0.30	0.736	**-1.21**	0.017
SPQS	0.06	0.785	0.27	0.439	-0.37	0.060

Cohort 2 (N=924)	Acute inpatient		Acute Outpatient		Proportion of out of hospital costs	
	Coef	P-val	Coef	P-val	Coef	P-val
POST	-0.03	0.764	-0.44	0.127	**-0.26**	0.042
IC	-0.06	0.535	0.19	0.525	-0.01	0.948
IC-dual	**1.34**	<0.001	**4.21**	<0.001	**6.13**	<0.001
DiD	0.04	0.749	0.36	0.383	0.28	0.129
DiD-dual	-0.26	0.474	-0.56	0.638	-0.23	0.678
SPQS	0.14	0.144	-0.17	0.583	-0.07	0.605

Note: significant (p<0.05) results in bold. N is the number of observations (4 observations per patient).