



OPEN ACCESS

Evaluation of the NHS England evidence-based interventions programme: a difference-in-difference analysis

Michael Anderson ,¹ Aoife Molloy,² Laia Maynou,^{1,3,4} Ilias Kyriopoulos,¹ Alistair McGuire,¹ Elias Mossialos ^{1,5}

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/bmjqs-2021-014478>).

For numbered affiliations see end of article.

Correspondence to

Dr Michael Anderson, Department of Health Policy, The London School of Economics and Political Science, London WC2A 2AE, UK; M.Anderson5@lse.ac.uk

Received 14 November 2021
Accepted 21 March 2022
Published Online First
7 April 2022



► <http://dx.doi.org/10.1136/bmjqs-2022-014977>



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Anderson M, Molloy A, Maynou L, et al. *BMJ Qual Saf* 2023;**32**:90–99.

ABSTRACT

Background The NHS England evidence-based interventions programme (EBI), launched in April 2019, is a novel nationally led initiative to encourage disinvestment in low value care.

Method We sought to evaluate the effectiveness of this policy by using a difference-in-difference approach to compare changes in volume between January 2016 and February 2020 in a treatment group of low value procedures against a control group unaffected by the EBI programme during our period of analysis but subsequently identified as candidates for disinvestment.

Results We found only small differences between the treatment and control group after implementation, with reductions in volumes in the treatment group 0.10% (95% CI 0.09% to 0.11%) smaller than in the control group (equivalent to 16 low value procedures per month). During the month of implementation, reductions in volumes in the treatment group were 0.05% (95% CI 0.03% to 0.06%) smaller than in the control group (equivalent to 7 low value procedures). Using triple difference estimators, we found that reductions in volumes were 0.35% (95% CI 0.26% to 0.44%) larger in NHS hospitals than independent sector providers (equivalent to 47 low value procedures per month). We found no significant differences between clinical commissioning groups that did or did not volunteer to be part of a demonstrator community to trial EBI guidance, but found reductions in volume were 0.06% (95% CI 0.04% to 0.08%) larger in clinical commissioning groups that posted a deficit in the financial year 2018/19 before implementation (equivalent to 4 low value procedures per month).

Conclusions Our analysis shows that the EBI programme did not accelerate disinvestment for procedures under its remit during our period of analysis. However, we find that financial and organisational factors may have had some influence on the degree of responsiveness to the EBI programme.

INTRODUCTION

Low value care can be defined as the ‘use of an intervention where evidence suggests it confers no or very little benefit on patients, or risk of harm exceeds likely benefit, or, more broadly,

Key messages

What is already known?

⇒ In April 2019, NHS England launched the evidence-based interventions programme, a novel approach to disinvest in low value care which aims to reduce avoidable harm to patients, maximise value and avoid waste by reducing unnecessary interventions.

What does this study add?

⇒ By comparing trends against a control group, we provide early evidence that the evidence-based interventions programme has not accelerated disinvestment for low value procedures in the first 11 months after implementation.

How might this study affect research, practice or policy?

⇒ Despite the lack of effectiveness to date, the evidence-based interventions programme has developed a structured and transparent approach to identify candidates for disinvestment.
⇒ Moving forward, the evidence-based interventions programme needs to consider how to balance both bottom-up and top-down implementation to achieve sustainable and consistent reductions in low value care.

the added costs of the intervention do not provide proportional added benefits”¹ Minimising low value care, and tackling unwarranted clinical variation is a major issue for all healthcare systems. It has been estimated that 25% of healthcare expenditure in the USA is spent on low value procedures.²

Equivalent figures for the UK do not exist, but disinvestment in low value care has been high on the policy agenda over the last two decades.^{3–8} Despite several initiatives to disinvest in low value care in the UK and internationally, analyses to date have indicated that many initiatives fail to achieve their objectives.⁹ This is typically attributed to factors such as poor engagement from the public and clinicians, and lack of data collection and monitoring arrangements.^{10 11} Over the last few years, NHS England has sought to develop a more coordinated and structured approach to disinvest in low value care through the evidence-based interventions (EBI) programme. The stated aims of the EBI programme is to reduce avoidable harm to patients, maximise value and avoid waste by reducing unnecessary interventions.¹² After undertaking a review of evidence and consulting with the public, commissioners and healthcare professionals, the first phase of the programme identified 17 low value procedures within two categories: category 1 interventions which have been shown to be ineffective and should no longer be offered to patients, and category 2 interventions which are only appropriate in certain circumstances.

The EBI programme is distinctly different to previous initiatives to disinvest in low value care in several respects. First, targets have been set for each local commissioning organisation, known as clinical commissioning groups (CCGs), to reduce the number of category 1 interventions to ‘near zero’ and category 2 interventions to 25% of baseline levels nationally. Second, the EBI programme introduced a zero tariff for category 1 interventions, and recommended all CCGs implement a prior approval process for category 2 interventions. Finally, progress in meeting agreed targets is monitored and fed back to hospitals and CCGs using a publicly available dashboard to allow benchmarking.¹³ Statutory guidance for the first phase of the programme was published in November 2018,¹² and the programme was officially launched in April 2019. A second phase of the EBI programme has since been developed involving 31 low value procedures, with statutory guidance published in November 2020.¹⁴ However, it has not been feasible to evaluate the effectiveness of the second phase of the EBI programme in this paper due to the significant impact of the COVID-19 pandemic on elective care provision in England.

To our knowledge, this paper is the first quantitative analysis of the EBI programme. The primary aim of this paper is to evaluate the impact of the first phase of the EBI programme in reducing the provision of low value procedures compared with a control group of other low value procedures unaffected by the EBI programme during our period of analysis. The secondary aims of this paper are to

establish if the effect of the EBI programme differs according to different organisational or financial factors.

METHOD

Study design

To test our primary aim, we used a difference-in-difference (DiD) approach, which is a quasi-experimental method commonly used to estimate the causal effect of a policy intervention against a comparator group that can adjust for both observable and non-observable confounding factors.¹⁵ While we cannot evaluate the effectiveness of the second phase of the EBI programme in this paper (the period of analysis in our paper ends in February 2020 which is several months before the second phase of EBI begins), the second phase does create an opportunity to construct a control group of low value procedures which were unaffected by the EBI programme during our period of analysis. To ascertain if there is evidence of adaptation after implementation of the EBI programme, we adapted a methodology used by Cooper *et al* to apply spline regression to analyse differences in trends between the treatment and control group.¹⁶

The two key assumptions which underlie the use of DiD analysis are that the treatment and control group have parallel trends before implementation of the policy intervention, and that the control group remains unaffected by the treatment after implementation.¹⁵ We relied on visual inspection of trends to test the parallel trend assumption. While data-driven approaches to test for parallel trends do exist, there is consensus among the literature that there is no perfect approach to test the parallel trend assumption and pre-trend testing is not a substitute for logical reasoning why parallel trends should hold for treatment and control group.^{17–19} In our case, we argue that the parallel trend assumption should hold as both the treatment and control groups are considered low value procedures. Moreover, one review of DiD studies found that pre-trend testing is rarely used in studies with large numbers of observations as even small differences in pre-implementation trends are likely to be significant.¹⁸ The authors of this review highlight how pre-trend testing is more often used in studies with small numbers of observations when the test could fail to reject parallel trends because it is underpowered rather than due to parallel trends.¹⁸ For the second assumption, we constructed a control group of low value procedures that should not have been impacted by disinvestment in the low value procedures targeted by the EBI programme during our period of analysis.

To test our secondary aims, we used a method developed to ascertain the difference between two DiD estimators known as triple difference estimation.²⁰ This

was used to expand on our DiD model to ascertain whether the implementation of the EBI programme was associated with a larger decrease in the provision in low value procedures in certain groups including between CCGs with different levels of financial performance, CCGs that were or were not part of a demonstrator community that volunteered to trial EBI guidance before implementation and NHS hospitals compared with independent sector hospitals.

Study cohort and data sources

We analysed pseudoanonymised individual patient level data between 1 January 2016 and 28 February 2020 retrieved from the NHS Digital Secondary Uses Services (SUS) database. We did not analyse data beyond this point as elective care activity was significantly impacted by the COVID-19 pandemic. Data access was provided by the NHS England EBI programme team to researchers from the London School of Economics and Political Science as part of a service development evaluation exercise. The identification of low value procedures was undertaken using combinations of primary and secondary procedure and diagnosis codes for each low value procedure, developed by the EBI programme, based on feedback from stakeholders including CCGs, hospitals and specialty organisations. These codes are publicly available and mentioned in EBI guidance.^{12 14} A summary of these codes is also mentioned in online supplemental tables 1 and 2. The NHS SUS database is classified according to finished consultant episodes, which relates to the clinician responsible for the respective aspect of care, and hospital spells, which encompass all activity from hospital admission to discharge, including ward transfers of patients. To avoid multiple counting of low value procedures, procedures were identified by applying these criteria to each hospital spell rather than finished consultant episode. For each low value procedure, information on individual patient characteristics was extracted including age, gender, deprivation and comorbidities. Comorbidities were classified according to the Charlson Comorbidity Index (CCI), based on code written by Quan *et al* that uses International Classification of Diseases codes.²¹ Deprivation was classified according to the English Index of Multiple Deprivation (IMD) 2019.²²

The primary treatment group for our analysis was activity data for category 2 procedures under the remit of phase I of the EBI programme (online supplemental table 3), with the exception of removal of benign skin lesions. Removal of benign skin lesions was removed from the primary analysis as it was seen as not comparable to other procedures under the remit of the EBI programme. It is a relatively minor procedure that often takes place in outpatient clinics rather than in surgical theatres, and as a relatively high-volume procedure its inclusion could bias results. The decision was made to not include category 1 procedures

in the primary analysis as category 1 procedures are not recommended to be conducted in any circumstances, and therefore are likely to have experienced a significantly larger decline than category 2 procedures before implementation of the EBI programme. To ascertain if these omissions significantly alter our results, we include both removal of benign skin lesions and category 1 procedures in the treatment group in separate robustness analyses.

We also constructed separate treatment groups for procedures grouped according to whether they were classified as high-cost or low-cost procedures. This was determined by whether estimated potential annual savings for each procedure exceed £10 000 000 per annum. High-cost procedures in the primary treatment group included hysterectomy for heavy menstrual bleeding, shoulder decompression, carpal tunnel syndrome release and Dupuytren's contracture release (online supplemental table 4). Estimated savings were calculated by the EBI team using baseline expenditure for each procedure (total expenditure for each procedure in 2017/18 divided by the number of associated hospital spells), and the assumption that all CCGs meet their targets to reduce provision. These estimated savings are mentioned in online supplemental table 4.

The control group for our analysis consists of four category A procedures: repair of minimally symptomatic inguinal hernia, surgical intervention for chronic rhinosinusitis, kidney stone surgery and surgical intervention for benign prostatic hypertrophy. These procedures were selected from category A low value interventions that are subject to phase II of the EBI programme (online supplemental table 3). Other category A procedures were not included in the control group as they are either diagnostics rather than surgical procedures, or considered to be vulnerable to potential spill-over effect of the first phase of EBI programme. Category B or C interventions were excluded as the EBI programme team have not yet developed reliable definitions to publish activity levels.¹⁴

Data on CCG financial performance was retrieved from NHS England.²³ CCGs were coded as being in deficit if their expenditure exceed their allocation in the financial year before the implementation of the EBI programme. During our period of analysis, many CCGs underwent mergers and the number of CCGs reduced from 191 in 2018/19 to 135 in 2019/20.²⁴ To overcome this, we consistently coded CCGs according to their CCG status in 2019/20 and the financial performance of CCGs that underwent a merger in 2019/20 was estimated using the total surplus or deficit for merged CCGs. The EBI team provided information on whether a CCG was coded as part of the demonstrator community that volunteered to trial EBI guidance before implementation. Hospital providers were coded as either NHS or independent sector hospitals according to their

organisation code classified by the NHS Digital Organisation Data Service.²⁵

Statistical analysis

Equation 1 shows the ordinary least squares (OLS) regression model for our DiD analysis that tests our primary aim:

$$\begin{aligned} \log Y_{ccgt} = & \beta_0 + \beta_1 t + \beta_2 \{t \geq EBI\} + \beta_3 (Phase1 \times t) + \\ & \beta_4 (Phase1 \times \{t \geq EBI\}) + \beta_5 Phase1 + \beta_6 EBI + \beta_7 (EBI \times Phase1) \\ & + \beta_8 yeardummy_t + \beta_9 monthdummy_t + \beta_{10} Z_{ccgt} + \varepsilon_{ccgt} \end{aligned} \quad (1)$$

$\log Y_{ccgt}$ is the dependent variable, the log of the number of total low value procedures undertaken at *ccg*, during month *t*. *Phase I* is the number of low value procedures in the treatment group; *t* indicates a running counter of months from January 2016 to February 2020. *EBI* is the break point in the spline corresponding to the policy start point, which is 1 from April 2019, and 0 before (the month which phase I of the EBI programme was launched). $\{t \geq EBI\}$ indicates a running counter of months from April 2019 to February 2020. Z_{ccgt} is a vector of CCG controls including aggregate patient characteristics (age, gender, CCI and IMD). The model was run using fixed effects which differenced out all time-invariant CCG characteristics out of the equation. Year and month dummies were also added to account for year and seasonal variation.

Setting $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ gives rises to our preferred standard DiD specification when the coefficient β_7 captures the treatment effect of the EBI programme, specifically the difference in the average change in volumes of low value procedures in the treatment group before and after the implementation of the EBI programme minus the difference in the average change in volumes of low value procedures in the control group. We ran this model with all phase I category 2 low value procedures with the exception of removal of benign skin lesions, and then separately according to whether procedures were classified as high or low cost. Setting $\beta_6 = \beta_7 = 0$ gives rises to our spline-based DiD regression specification when the coefficient β_4 captures the difference in the average monthly rate of change in volumes of low value procedures in the treatment group before and after the implementation of the EBI programme minus the difference in the average monthly rate of change in volumes of low value procedures for the control group. Relaxing these assumptions allows a combination of these estimators. This specification allows a step change in policy, and a change in trends. As these additional specifications did not show significantly different results, we did not repeat these specifications with procedures classified according to whether they are high or low cost.

A number of robustness checks were performed to test the reliability of our results. First, with the treatment group including volumes for both category 1 and 2 procedures. Second, with the treatment group including volumes for removal of benign skin lesions. Third, to account for potential anticipatory behaviour change before implementation of the EBI programme, the implementation period was changed to November 2018 (the month which the statutory guidance for the EBI programme was published).

Equation 2 shows the OLS equation which uses a triple difference estimator to test our secondary aims:

$$\begin{aligned} \log Y_{it} = & \beta_0 + \beta_1 Phase1 + \beta_2 EBI + \beta_3 X + \beta_4 (EBI \times Phase1) + \\ & \beta_5 (EBI \times X) + \beta_6 (Phase1 \times X) + \beta_7 (EBI \times Phase1 \times X) + \beta_8 \\ & yeardummy_t + \beta_9 monthdummy_t + \beta_{10} Z_{it} + \varepsilon_{it} \end{aligned} \quad (2)$$

$\log Y_{ccgt}$ is the dependent variable, the log of the number of total low value procedures undertaken at CCG or hospital, *i*, during month *t*. *X* is a binary variable which reflects different CCG or hospital characteristics. The equation was run three separate times. First, with *X* being 1 for CCGs which posted a deficit in financial year 2018/19 and 0 for those which did not. Second, with *X* being 1 for CCGs which were part of the demonstrator community that volunteered to trial EBI guidance before implementation and 0 for those which were not. Third, with *X* being 1 for NHS hospitals and 0 for independent sector providers. The coefficient β_6 captures the difference in the average change in low value procedures in the treatment group between these CCG and hospital characteristics. The other components of equation 2 are the same as equation 1. We also performed the same robustness checks for equation 1 and equation 2. As classifying procedures in the treatment group according to their cost did not produce significantly different results in our standard DiD specification, we did not repeat this with the triple difference estimator.

Patient and public involvement

No patients were involved in the development of the research question, design or implementation of this study. The authors plan to disseminate results to patients and policymakers through outreach activities and platforms provided by NHS England, and the Royal College of Surgeons in England.

RESULTS

Descriptive statistics

When visually inspecting pre-implementation trends between the treatment and control group (figure 1), the assumption of parallel trends appears to hold in the pre-treatment period. Similarly, the pre-implementation trends between treatment and control group do not appear to significantly differ for alternative compositions of the treatment group discussed

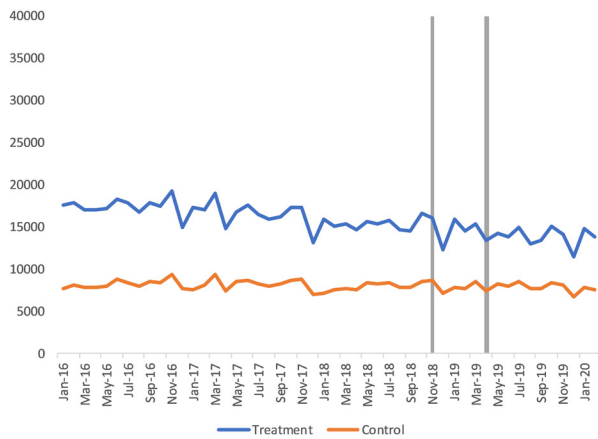


Figure 1 Trends in evidence-based interventions programme (EBI) treatment and control group procedures. *Treatment group includes all EBI phase I category 2 procedures with the exception of removal of benign skin lesions. Control group includes EBI phase II procedures including repair of minimally symptomatic inguinal hernia, surgical intervention for chronic rhinosinusitis, kidney stone surgery and surgical intervention for benign prostatic hypertrophy. The grey lines reflect the launch of the EBI programme in April 2019, and the publication of statutory guidance in November 2018. This figure was created by the coauthors of this manuscript.

in the ‘Methods’ section (online supplemental figures 1–4).

When focusing on the 11 months before and after the implementation of the EBI programme (table 1), the proportion of procedures undertaken after implementation was similar in the treatment group (48.4%), and control group (49.2%). There was more variation in the proportion of individual procedures undertaken after implementation in the treatment group (41.3%–53.4%), compared with the control group (48.5%–50.9%).

There were no significant changes in the proportion of patients, which were male or female before and after the implementation of the EBI programme for any of groups of procedures (online supplemental table 5). Most patients in the control group were male, which likely reflects how prostate surgery is performed exclusively for male patients and hernia repair is performed more frequently for male patients. There were significant increases in the average age of all groups of procedures, with largest increases in category 1 procedures at 1.42 years (95% CI 1.07 to 1.77), and the smallest increase in the control group at 0.79 years (95% CI 0.64 to 0.94). The average IMD score increased for all groups of procedures, although this increase was not significant in any groups of procedures. There were significant increases in average CCI for all groups of procedures. However, these increases were very small and the largest increase was only 0.09 points (95% CI 0.07 to 0.10) in category 1 procedures.

While 38 out of 135 CCGs (28.1%) posted a deficit in the financial year 2018/19 before implementation, a higher proportion of low value procedures were

undertaken in these CCGs for both the treatment (37.0%) and control groups (39.0%) (online supplemental table 6). Forty-eight out of 135 CCGs (35.6%) were coded as having volunteered to be part of the demonstrator community to trial EBI guidance before implementation, and a similar proportion of low value procedures were undertaken in these CCGs for both the treatment group (36.2%) and control group (35.6%). While 226 out of 388 hospitals (58.2%) were independent sector hospitals, only 18.2% of procedures in the treatment group, and 20.5% of procedures in the control group were conducted in these hospitals. Although it should be noted that independent sector hospitals are typically smaller than NHS hospitals and have much less capacity.

Difference-in-difference analysis

In our primary DiD model, the coefficient that reflects the treatment effect of the EBI programme was 0.10 (95% CI 0.09 to 0.11), and significant at the $p < 0.001$ level (table 2). This indicates that reductions in the provision of low value procedures in the treatment group were 0.10% smaller than reductions in the control group, which is equivalent to 16 low value procedures per month. This coefficient remained positive for the high-cost and low-cost treatment groups, indicating that the effectiveness of the EBI programme was not influenced by estimated potential annual savings for individual procedures. In our spline-based and combination-based DiD, the coefficients that reflect differences in monthly changes in volume were not significant. The co-efficient that reflects differences in the step-change in policy was 0.05 (95% CI 0.03 to 0.06) and significant at the $p < 0.001$ level. This indicates the step-change reduction in the provision of low value procedures in the treatment group was 0.05% smaller than the control group, which is equivalent to 7 low value procedures. These findings are consistent across all robustness analyses outlined in our ‘Methods’ section (online supplemental tables 7–9).

Triple difference estimation

When focusing on the results of the triple difference estimation, we found that the coefficient which reflected differences in reductions in the treatment group after implementation of the EBI programme for CCGs that posted a deficit in the financial year 2018/19, was -0.06 (95% CI -0.08 to -0.04), and significant at the $p < 0.001$ level (table 3). This indicates that reductions in low value procedures after the implementation of the EBI programme were 0.06% larger in CCGs posting a deficit in the baseline year (2018/19) than those which did not, which is equivalent to 4 low value procedures per month. We found no significant differences after implementation of the EBI programme for CCGs that were part of the demonstrator community compared with those that

Table 1 Number of procedures for phase I and phase II of EBI programme*

	Before EBI	After EBI	Total
Phase I procedures (treatment)			
Category 1			
Intervention for snoring (not OSA)	667 (56.1%)	521 (43.9%)	1188
Dilatation and curettage for heavy menstrual bleeding	217 (46.6%)	249 (53.4%)	466
Knee arthroscopy with osteoarthritis	3966 (55.8%)	3147 (44.2%)	7113
Injection for non-specific low back pain without sciatica	13 022 (58.7%)	9164 (41.3%)	22 186
Total	17 872 (57.7%)	13 081 (42.3%)	30 953
Category 2			
Breast reduction	938 (53.8%)	806 (46.2%)	1744
Removal of benign skin lesions	94 427 (50.3%)	93 347 (49.7%)	187 774
Grommets	7007 (53.1%)	6182 (46.9%)	13 189
Tonsillectomy	28 382 (53.9%)	24 260 (46.1%)	52 642
Haemorrhoid surgery	7929 (51.4%)	7495 (48.6%)	15 424
Hysterectomy for heavy bleeding	17 790 (51.4%)	16 822 (48.6%)	34 612
Chalazia removal	4761 (52.9%)	4239 (47.1%)	9000
Shoulder decompression	8947 (56.6%)	6860 (43.4%)	15 807
Carpal tunnel syndrome release	39 162 (51.0%)	37 677 (49.0%)	76 839
Dupuytren's contracture release	13 965 (51.7%)	13 050 (48.3%)	27 015
Ganglion excision	5035 (51.0%)	4844 (49.0%)	9879
Trigger finger release	7264 (51.4%)	6865 (48.6%)	14 129
Varicose vein surgery	25 693 (52.7%)	23 044 (47.3%)	48 737
Total	261 300 (51.6%)	245 491 (48.4%)	506 791
Total (category 1 and 2)	279 172 (51.9%)	258 572 (48.1%)	537 744
Phase II procedures (control)			
Hernia repair	50 748 (51.5%)	47 771 (48.5%)	98 519
Sinus surgery	11 712 (49.7%)	11 853 (50.3%)	23 565
Kidney stone surgery	13 483 (51.1%)	12 896 (48.9%)	26 379
Benign prostatic hypertrophy surgery	13 502 (49.1%)	13 988 (50.9%)	27 490
Total	89 445 (50.8%)	86 508 (49.2%)	175 953

*Percentage in parentheses indicates the proportion of procedures undertaken in the 11 months before and after the implementation of EBI in April 2019 for illustrative purposes. However, it should be noted that our difference-in-difference analysis considers longer pre-implementation trend from 1 January 2016 to 31 March 2019. This table was created by the coauthors of this manuscript.
OSA, Obstructive sleep apnoea.

were not. The co-efficient which reflected differences in reductions in the treatment group for NHS hospitals after implementation of the EBI programme was -0.35 (95% CI -0.45 to -0.26), and significant at $p < 0.001$ level. This indicates that reductions in low value procedures after the implementation of the EBI programme were 0.35% larger in NHS hospitals compared with independent sector hospitals, which is equivalent to 47 low value procedures per month. These findings are consistent according to all robustness analyses outlined in our 'Methods' section (online supplemental tables 10–12).

DISCUSSION

Summary of findings

This analysis demonstrates that for the first 11 months after implementation, the EBI programme did not successfully achieve its aim of accelerating disinvestment for the low value procedures under its remit. Conversely, on the understanding that the control group provides a counterfactual scenario whereby

the EBI programme did not exist, we found statistically significant evidence that the implementation EBI programme was associated with a small increase in the volumes of low value procedures under its consideration. This finding is consistent irrespective of whether we change the composition of the treatment group according to procedures with estimated potential annual savings of above or below £10 000 000. When analysing organisational and financial factors which may have influenced implementation of the EBI programme, we found that CCGs which posted a deficit in the financial year before implementation had larger reductions in low value procedures than CCGs which did not. This may be because CCGs which posted a deficit in the year before implementation felt the need to more pro-actively engage with the EBI programme as one mechanism to save costs and reduce their deficit in the subsequent year. Despite approximately a third of CCGs volunteering to be part of a demonstrator community which trialled EBI recommendations before implementation, there were

Table 2 Results for DiD analysis (%)

	Model 1 (standard DiD)	Model 2 (standard DiD)	Model 3 (high cost procedures)	Model 4 (low cost procedures)	Model 5 (time trend analysis)	Model 6 (combination)
<i>Phase I</i>	0.54*** (0.53 to 0.56)	0.55*** (0.54 to 0.57)	0.92*** (0.89 to 0.94)	0.92*** (0.89 to 0.94)	0.53*** (0.52 to 0.54)	0.54*** (0.52 to 0.55)
<i>EBI</i>	-11.82*** (-12.89 to -10.75)	-12.50*** (-14.01 to -10.99)	-8.65*** (-10.40 to -6.90)	-11.21*** (-13.00 to -9.42)	-	-4.98** (-7.20 to -2.75)
<i>EBI</i> × <i>phase I</i>	0.10*** (0.09 to 0.11)	0.10*** (0.09 to 0.11)	0.15*** (0.13 to 0.16)	0.17*** (0.16 to 0.19)	-	0.05*** (0.03 to 0.06)
<i>t</i>	-	-	-	-	-0.27*** (-0.37 to -0.16)	-0.22*** (-0.33 to -0.12)
<i>t</i> ≥ <i>EBI</i>	-	-	-	-	-0.87 (-1.15 to -0.59)	-0.36* (-0.71 to -0.02)
<i>t</i> × <i>phase I</i>	-	-	-	-	0.00*** (0.00 to 0.00)	0.00** (0.00 to 0.00)
<i>t</i> ≥ <i>EBI</i> × <i>phase I</i>	-	-	-	-	0.01 (0.00 to 0.01)	0.00 (-0.00 to 0.00)
Year dummies	No	Yes	Yes	Yes	Yes	Yes
Month dummies	No	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes	Yes	Yes
Constant	4.31*** (4.29 to 4.32)	4.18*** (4.08 to 4.29)	3.89*** (3.78 to 4.01)	3.85*** (3.75 to 3.96)	4.21*** (4.10 to 4.32)	4.20*** (4.09 to 4.32)
Observations	6750	6750	6750	6750	6750	6750
Units of observation	135 CCGs	135 CCGs	135 CCGs	135 CCGs	135 CCGs	135 CCGs

Co-efficients are reported as percentages, and can be understood as follows: *phase I* is the percentage difference in change in volumes between the treatment and control group. *EBI* is the percentage difference in volumes for all procedures before and after the implementation of the EBI programme. In models 1–4, *EBI*×*phase I* represents the treatment effect of the EBI programme and is the percentage DiD of volumes before and after implementation between the treatment and control groups. In model 6, *EBI*×*phase I* is the percentage difference in changes in volumes between the treatment and control groups during implementation of the EBI programme; *t* reflects monthly percentage change in volumes for all procedures; *t*≥*EBI* reflects the monthly percentage change in volumes for all procedures after the implementation of the EBI programme; *t*≥*EBI*×*phase I* reflects the DiD in the monthly percentage change in volumes between the treatment and control groups after the implementation of the EBI programme. 95% CIs are mentioned in parentheses. All models used fixed effects, therefore errors are clustered at the CCG level. This table was created by the coauthors of this manuscript.
p*<0.05; *p*<0.01; ****p*<0.001.
–, not applicable to this model; CCG, clinical commissioning group; DiD, difference-in-difference; EBI, evidence-based interventions programme.

no significant differences between changes in volumes of low value procedures between demonstrator and non-demonstrator CCGs. Finally, we found that NHS hospitals had significantly larger reductions in low value procedures than independent sector hospitals. This may be because NHS hospitals have an institutional culture which is more amenable to NHS England-led national quality improvement initiatives, whereas independent sector hospital hospitals may be more motivated by their respective corporate-level objectives and strategies.

Strengths and limitations

There are several strengths to this analysis. First, to our knowledge this is the only quantitative evaluation of the impact of EBI programme. This provides important information on the early impact of the EBI programme so far which can be used to inform its planned expansion. Second, this evaluation uses a DiD analysis which is a robust method for casual inference. This is a valuable addition to the literature on empirical evaluations of national disinvestment initiatives which typically use weaker methods such as interrupted-time series, or before and after analyses.⁹ Third, we extend our analysis to focus on a range of organisational and

financial factors which may have influenced implementation and use a series of robustness tests to assess the reliability of our findings.

Despite these strengths, the robustness of this analysis is heavily dependent on the ability of our selected control group to meet the two key assumptions to undertake a DiD.¹⁵ First, that there are parallel trends pre-implementation. Second, that the control group remains unaffected by the treatment after intervention. For the first assumption, the degree to which parallel trends exist is frequently debated in DiD analyses.¹⁸ However, we can be reassured that on visual inspection of pre-implementation trends the assumption of parallel trends appears to hold. Moreover, the coefficient which reflects differences in trends between the treatment and control groups is very small (table 3). For the second assumption, we constructed the control group to minimise the potential for any spill-over effect of the EBI programme. However, we cannot completely exclude any possibility of a spill-over effect as the EBI programme may have generally encouraged a culture of disinvestment in hospitals and CCGs.

There are other minor limitations of our analysis to consider when interpreting the findings of this study. First, we used a coding algorithm for the CCI

Table 3 Results for triple difference estimation (%)

logTotal	Model 1 (CCG deficit)	Model 2 (CCG demonstrator)	Model 3 (NHS hospitals)
<i>Phase I</i>	0.69*** (0.67 to 0.71)	0.57*** (0.56 to 0.59)	3.90*** (3.79 to 3.99)
<i>EBI</i>	-13.07*** (-14.74 to 11.40)	-12.03*** (-13.75 to -10.31)	-9.69*** (-12.48 to -6.90)
<i>EBI</i> × <i>phase I</i>	0.13*** (0.11 to 0.14)	0.10*** (0.09 to 0.11)	0.48*** (0.39 to 0.56)
<i>EBI</i> × <i>X</i>	2.26 (-0.02 to 4.53)	-1.24 (-3.43 to 0.94)	-2.79 (-6.85 to 1.27)
<i>Phase I</i> × <i>X</i>	-0.30*** (-0.32 to 0.28)	-0.05*** (-0.07 to -0.02)	-3.08*** (-3.19 to -2.98)
<i>EBI</i> × <i>phase I</i> × <i>X</i>	-0.06*** (-0.08 to 0.04)	0.00 (-0.02 to 0.02)	-0.35*** (-0.45 to -0.26)
Constant	4.15*** (4.05 to 4.25)	4.18*** (4.08 to 4.29)	2.84*** (2.76 to 2.93)
Observations	6750	6750	16 559
Units of observation	135 CCGs	135 CCGs	382 hospitals

Co-efficients are reported as percentages, and can be understood as follows: *phase I* is the percentage difference in change in volumes between the treatment and control groups. *EBI* is the percentage difference in volumes for all procedures before and after the implementation of the EBI programme. *EBI*×*phase I* is the percentage DiD of volumes before and after the implementation of the EBI programme between the treatment and control groups. *EBI*×*X* is the average percentage DiD in volumes for all procedures before and after the implementation of the EBI programme between different organisational characteristics defined by *X*. *Phase I*×*X* is average percentage difference in changes in volumes for the treatment group for different organisational characteristics defined by *X*. *EBI*×*phase I*×*X* is the average percentage DiD in volumes before and after the implementation of the EBI programme for the treatment group between different organisational characteristics defined by *X*. In model 1, *X* is 1 for CCGs which posted a deficit in financial year 2018/19, and 0 for those which did not. In model 2, *X* is 1 for CCGs which were part of the demonstrator community, and 0 for those which were not. In model 3, *X* is 1 for NHS hospitals and 0 for independent sector providers. 95% CIs in parentheses. All models used fixed effects, therefore errors are clustered at the CCG or hospital level. This table was created by the coauthors of this manuscript.
*P<0.05; **p<0.01; ***p<0.001.
CCG, clinical commissioning group; EBI, evidence-based interventions programme; NHS, National Health Service.

developed by Quan *et al.*²¹ There is a modified version developed for use with hospital administrative data collected in England which is arguably more applicable to our dataset but this was not available on STATA.²⁶ Second, we classified spells according to hospital spells which do not take account of transfers between hospitals. Therefore, it is possible we have counted some procedures twice if an interhospital transfer has occurred. Although, this is unlikely to have substantially impacted our results as we expect very few interhospital transfers took place for procedures in our analysis as they are generally of low complexity.

Policy implications and conclusions

The findings of this analysis are consistent with many other empirical analyses of national disinvestment initiatives that often conclude they have not successfully achieved their aims.⁹ Despite broad consensus from the healthcare community in the UK on the importance of disinvestment in low value care,⁸ other evaluations on the effectiveness of interventions such as the National Institute for Health and Care Excellence guidance and 'do not do' recommendations have also found limited impact.^{4 27 28} While it is frustrating that early evidence from the EBI programme has indicated limited effectiveness, the EBI programme still represents a major step forward as it has developed a

structured and transparent approach to identify candidates for disinvestment that involved broad consultation with specialty organisations, patient groups and CCGs. It is possible that the EBI programme has taken an approach to implementation that is too top-down. The barriers to disinvestment in low value care are complex and involve a range of patient, clinician, organisational and health system factors.²⁹ Moreover, the success of the EBI programme is heavily reliant on successful collaboration between CCGs, hospitals and primary care in developing and effectively implementing prior approval processes. Moving forward, the EBI programme will need to focus on developing strategies to facilitate and monitor these collaborations at the local level to balance both bottom-up and top-down implementation in a manner that could foster more sustainable and consistent reductions in low value care.

Author affiliations

¹Department of Health Policy, The London School of Economics and Political Science, London, UK

²Health Inequalities Improvement Programme, NHS England, London, UK

³Department of Econometrics, Statistics and Applied Economics, Universitat de Barcelona, Barcelona, Spain

⁴Center for Research in Health and Economics, Universitat Pompeu Fabra, Barcelona, Spain

⁵Institute of Global Health Innovation, Imperial College London, London, UK

Correction notice This article has been corrected since it was first published online. In the abstract, 0.35% (95% CI 0.26% to 0.445%) was amended to 0.35% (95% CI 0.26% to 0.44%) and 0.06% (95% CI 0.045% to 0.08%) was amended to 0.06% (95% CI 0.04% to 0.08%).

Acknowledgements We would like to express our gratitude to members of the NHS England evidence-based interventions programme team who provided access to NHS England Secondary Uses Service data for the purposes of this analysis and provided insights regarding their experiences interacting with hospitals and commissioning bodies during implementation of the EBI programme.

Contributors MA undertook the analysis and drafted the manuscript. All other authors were collaboratively involved in the design of the analysis and provided comments and edits to iterative versions of the manuscript. MA is the guarantor for this manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests AM, was until August 2021, Clinical Lead of the EBI Programme. MA was also granted the honorary position of clinical fellow at NHS England between January 2021 and December 2021. The other authors have no relevant conflicts of interest to declare related to this manuscript.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from NHS Digital and are not publicly available.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Michael Anderson <http://orcid.org/0000-0002-8454-4640>

Elias Mossialos <http://orcid.org/0000-0001-8664-9297>

REFERENCES

- Scott IA, Duckett SJ. In search of professional consensus in defining and reducing low-value care. *Med J Aust* 2015;203:179–81.
- Shrank WH, Rogstad TL, Parekh N. Waste in the US health care system: estimated costs and potential for savings. *JAMA* 2019;322:1501–9.
- Donaldson L. *On the state of the public health: annual report of the chief medical officer 2005*, 2005.
- Chamberlain CA, Martin RM, Busby J, *et al.* Trends in procedures for infertility and caesarean sections: was NICE disinvestment guidance implemented? NICE recommendation reminders. *BMC Public Health* 2013;13:112.
- National Audit Office. *Progress in making NHS efficiency savings: department of health*. The Stationery Office, 2012.
- The Royal College of Surgeons of England. Is access to surgery a Postcode lottery, 2014. Available: <https://www.rcseng.ac.uk/-/media/files/rcs/news-and-events/is-access-to-surgery-a-postcode-lottery.pdf> [Accessed 27 Feb 2022].
- Coronini-Cronberg S, Bixby H, Laverty AA, *et al.* English National Health Service's savings plan may have helped reduce the use of three 'low-value' procedures. *Health Aff* 2015;34:381–9.
- Malhotra A, Maughan D, Ansell J, *et al.* Choosing wisely in the UK: the Academy of medical Royal colleges' initiative to reduce the harms of too much medicine. *BMJ* 2015;350:h2308.
- Chambers JD, Salem MN, D'Cruz BN, D'Cruz BN, *et al.* A review of empirical analyses of disinvestment initiatives. *Value Health* 2017;20:909–18.
- van Dulmen SA, Naaktgeboren CA, Heus P, *et al.* Barriers and facilitators to reduce low-value care: a qualitative evidence synthesis. *BMJ Open* 2020;10:e040025.
- Colla CH, Mainor AJ, Hargreaves C, *et al.* Interventions aimed at reducing use of low-value health services: a systematic review. *Med Care Res Rev* 2017;74:507–50.
- NHS England. Evidence-Based interventions: guidance for CCGs, 2018. Available: <https://www.evidence.nhs.uk/document?id=2128996&returnUrl=search%3Fq%3Dpolicy%2Bon%2Bchronic%2Bshoulder%2Bpain> [Accessed 27 Feb 2022].
- NHS Business Services Authority. Catalyst - public insight portal. Available: <https://www.nhsbsa.nhs.uk/new-catalyst-public-insight-portal-released-public> [Accessed 27 Feb 2022].
- Academy of Medical Royal Colleges. Evidence-Based interventions list 2 guidance, 2020. Available: https://www.aomrc.org.uk/wp-content/uploads/2020/12/EBI_list_2_guidance_150321.pdf [Accessed 27 Feb 2022].
- Blundell R, Costa Dias M, Dias MC. Evaluation methods for Non-Experimental data. *Fisc Stud* 2000;21:427–68.
- Cooper Z, Gibbons S, Jones S, *et al.* Does Hospital competition save lives? Evidence from the English NHS patient choice reforms. *Econ J* 2011;121:F228–60.
- Kahn-Lang A, Lang K. The promise and pitfalls of Differences-in-Differences: reflections on 16 and pregnant and other applications. *J Bus Econ Stat* 2019:1–14.
- Roth J. Pre-Test with caution: Event-study estimates after testing for parallel trends. American economic association, 2019. Available: <https://www.aeaweb.org/articles?id=10.1257/aeri.20210236> [Accessed 27 Feb 2022].
- Jaeger DA, Joyce TJ, Kaestner R. A cautionary tale of evaluating identifying assumptions: did reality TV really cause a decline in teenage childbearing? *Journal of Business & Economic Statistics* 2020;38:317–26.
- Olden A, Møen J. The triple difference estimator. social science research network, 2020. Available: <https://papers.ssrn.com/abstract=3582447> [Accessed 27 Feb 2022].
- Quan H, Sundararajan V, Halfon P, *et al.* Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care* 2005;43:1130–9.
- UK Government. English indices of deprivation, 2020. Available: <https://www.gov.uk/government/collections/english-indices-of-deprivation> [Accessed 17 Feb 2022].
- NHS England. Financial performance report 2018/19: quarter 4 (January to March). Available: <https://www.england.nhs>

- uk/wp-content/uploads/2019/10/financial-information-4th-quarter-2018-19.pdf [Accessed 27 Feb 2022].
- 24 NHS England. Clinical commissioning group details, 2021. Available: <https://www.england.nhs.uk/ccg-details/> [Accessed 27 Feb 2022].
 - 25 NHS Digital. Organisation data service. Available: <https://digital.nhs.uk/services/organisation-data-service> [Accessed 27 Feb 2022].
 - 26 Bottle A, Aylin P. Application of AHRQ patient safety indicators to English hospital data. *Qual Saf Health Care* 2009;18:303–8.
 - 27 Ryan J, Piercy J, James P. Assessment of NICE guidance on two surgical procedures. *Lancet* 2004;363:1525–6.
 - 28 Dietrich ES. Effects of the National Institute for health and clinical excellence's technology appraisals on prescribing and net ingredient costs of drugs in the National health service in England. *Int J Technol Assess Health Care* 2009;25:262–71.
 - 29 Norton WE, Chambers DA. Unpacking the complexities of de-implementing inappropriate health interventions. *Implement Sci* 2020;15:2.