
Article:
Mason, Anne, Liu, Dan, Kasteridis, Panagiotis et al. (2018) Investigating the impact of primary care payments on underdiagnosis in dementia: a difference-in-differences analysis. International journal of geriatric psychiatry. GPS-17-0432.R1. pp. 1090-1097. ISSN: 0885-6230

https://doi.org/10.1002/gps.4897

Title: Investigating the impact of primary care payments on underdiagnosis in dementia: a difference-in-differences analysis

Running head: Payments to tackle underdiagnosis in dementia

Keywords: Dementia, Reimbursement, Incentive, Primary Health Care

Key points:

1. Receiving a timely formal diagnosis of dementia can allow patients and their carers to access appropriate care and support packages, prevent avoidable health crises and plan ahead more effectively.
2. The combined effect of two incentive schemes was to increase GP dementia registers nationally by around 40,000 cases; this figure would have been almost 50,000 if all practices had taken part.
3. The schemes had the intended impact on dementia care, suggesting that financial incentives can enhance performance in primary care, and may be useful for other disease areas where underdiagnosis is problematic.

Author names: Anne Mason, Dan Liu, Panagiotis Kasteridis, Maria Goddard, Rowena Jacobs, Raphael Wittenberg, Gerard McGonigal

Names of Institutions: Centre for Health Economics, University of York; Personal Social Services Research Unit, London School of Economics and Political Science; Department of Medicine for the Elderly, York Teaching Hospital NHS Foundation Trust.

Corresponding author: A Mason: Centre for Health Economics, University of York, York YO10 5DD; E: anne.mason@york.ac.uk; T: 01904 321432

Funder: NIHR Policy Research Programme (Economics of Social and Health Care; PRP Ref: 103 0001).

Word count (excl abstract, tables, captions and refs): 3493

Abstract

Objective: In England, two primary care incentive schemes were introduced to increase dementia diagnosis rates to two-thirds of expected levels. This study assesses the effectiveness of these schemes.

Methods: We used a difference-in-differences framework to analyse the individual and collective impacts of the incentive schemes: (1) Directed Enhanced Service 18 (DES18: facilitating timely diagnosis of and support for dementia); (2) the Dementia Identification Scheme (DIS). The dataset included 7,529 English general practices, of which 7,142 were active throughout the 10-year study period (April 2006 to March 2016). We controlled for a range of factors, including a contemporaneous hospital incentive scheme for dementia. Our dependent variable was the percentage of expected cases that was recorded on practice dementia registers (the 'rate').

Results: From March 2013 to March 2016, the mean rate rose from 51.8% to 68.6%. Both DES18 and DIS had positive and significant effects. In practices participating in the DES18 scheme, the rate increased by 1.44 percentage points more than the rate for non-participants; DIS had a larger effect, with an increase of 3.59 percentage points. These combined effects increased dementia registers nationally by an estimated 40,767 individuals. Had all practices fully participated in both schemes, the corresponding number would have been 48,685.

Conclusion: The primary care incentive schemes appear to have been effective in closing the gap between recorded and expected prevalence of dementia, but the hospital scheme

had no additional discernible effect. This study contributes additional evidence that financial incentives can motivate improved performance in primary care.

Introduction

Dementia is a devastating long-term condition that is projected to place increasing demands on health and care services.[1] In the absence of curative treatments, efforts are focused on reducing risk, timely diagnosis and early intervention.[2] General practitioners (GPs) are uniquely placed to co-ordinate health and social care services for people with dementia and to address the support needs of the family and friends who care for them.

The English Department of Health's Dementia strategy (2009) [3] and the Dementia Challenge (2012)[4] highlighted the problem of 'underdiagnosis': it was estimated that around half of those with dementia did not have a formal diagnosis. Anticipated benefits of a formal diagnosis included improved access to relevant care and support services; empowering patients and their families to plan their lives better; prevention of avoidable health crises and further cognitive decline (when these are due to vascular risk factors);[5] and improvements in the delivery of care and in communication between providers, patients and carers. [6]

NHS England announced a £90m package to improve dementia diagnosis and care.[7] The raft of measures included two financial incentive schemes in primary care and one hospital scheme. The aim of these 'tools and levers' was to increase diagnosis rates to the level of 67% of the expected number of people with the condition by March 2015 (the so-called 'two-thirds ambition').[8] Whilst some interventions were designed to improve dementia care directly, financial incentives have been shown to be powerful levers in effecting behavioural changes in primary and secondary care.[9 10] The aim of this study was to evaluate the impact of these financial incentives on diagnostic rates of dementia in primary care.

Incentive schemes

The two primary care schemes for tackling underdiagnosis were the Directed Enhanced Service (DES18) and the Dementia Identification scheme (DIS). The schemes were facilitated by a separate pay-for-performance scheme, the Quality and Outcomes Framework (QOF). Since 2006, the QOF has incentivised good quality care for people with dementia, primarily via a face-to-face annual review,[11-13] and requires practices to maintain a dementia register. We measured the schemes' effectiveness in tackling underdiagnosis by the gap between the 'reported' (recorded) and 'expected' numbers on practices' QOF dementia registers.[14]

DES18 ran from April 2013 to March 2016.[15] The scheme encouraged a proactive approach to timely assessment of individuals at risk of dementia, followed-up by advanced care planning for newly diagnosed patients and a health check for carers. Participating practices received an upfront payment, and an annual end-of-year payment based on the proportion of national assessments the practice undertook. These payments were funded centrally by annual budgets of £21m for each of the two payments, making a total budget of £126m over the three years DES18 operated.

DIS operated for 6 months from 30 September 2014 to 31 March 2015 and was intended to support and complement DES18.[16] NHS England paid GP practices £55 for each additional patient included on the QOF dementia register, based on the differential between the register at 30 September 2014 and 31 March 2015. Funding available for this scheme totalled £5m.[17]

A third scheme that incentivised hospitals (FAIR) ran in parallel with the primary care schemes, and we controlled for this in our analyses.

Methods

Data

Details of the datasets analysed are in Appendix A, and summary statistics for the outcome and control variables in our model are in Table 1.

TABLE 1 HERE

Study sample

To be included in our study, practices had to have a QOF dementia register so that recorded and expected numbers of dementia patients could be calculated. We compiled a panel of all eligible English practices that were open during the study period 2006-07 to 2015-16.

For our base case analyses, our sample was a balanced panel of 7,142 practices that contributed data in all ten years. We undertook two sensitivity analyses. First, we re-estimated using an unbalanced panel of 7,529 practices totalling 74,241 practice-year observations: this includes practices that closed, opened, split or merged during the study period. Second, we tested the implications of assuming that the effect of DES18 persisted after a practice had exited the scheme.

Dependent variable

For two practices with identical dementia registers but with very different 'expected' registers, the risk of an 'event' (adding a patient to the dementia register) can vary considerably because practices with larger expected registers have greater capacity to improve. We defined our dependent variable as the percentage of expected cases of dementia that was recorded on the dementia register (the 'rate').

The numerator was the number of people recorded on the GP practice's dementia register. The denominator was the expected number of patients aged 65 and over with dementia, which was based on the number, age and sex of a practice's registered patients living in a nursing home; and on the number, age and sex of the *remaining* practice patients. We distinguished nursing home patients from community-dwelling patients because the prevalence of dementia differs between the two groups.[18]

The General and Personal Medical Services dataset publishes annual data on the number, age and sex of a practice's registered patients. NHS Digital publishes annual data on the number of nursing home patients in a practice, but not by age and sex. We therefore estimated the number of nursing home patients in each age / sex band using values for the national care home population taken from the 2011 Census. Appendix A details the data sources used for these calculations.

Defining Participation

Our key explanatory variables were practice participation in the two schemes. We used the following rules to define participation.

Practices were deemed to have participated in DES18 in a particular year in the period 2013-14 to 2015-16 if they reported data on the number of dementia assessments undertaken that year, even if that number was zero. Practices not reporting assessment data were deemed to be non-participants.

Practices participating in DIS were required to report monthly data on recorded dementia diagnoses for September 2014 and for at least one month from October 2014 to March 2015. [16] However, some practices that submitted monthly data did not take part in DIS.

NHS England provided us with a DIS participant list based on information collected by Local Area Teams for payment purposes.

Covariates

One of the Commissioning for Quality and Innovation (CQUIN) national targets,[19] the hospital incentive scheme 'FAIR' was also designed to increase diagnostic rates for dementia.

For all patients aged 75 and over who had an emergency admission involving a hospital stay of at least 72 hours, FAIR rewarded hospitals according to their performance on three indicators (1) **F**ind, (2) **A**ssess & **I**dentify and (3) **R**efer individuals for specialist diagnosis and follow up. Each indicator was scored 0-100%, with payment triggered by achieving at least 90% on all three indicators in any consecutive three months.

To control for the effect of FAIR on QOF dementia registers, we derived a time-varying measure of hospital effort based on the first two FAIR indicators only, because the third indicator ('Refer') was defined differently in the final year and its performance data were not published.

We converted the two hospital trust-level scores to weighted GP practice average values. To match the CQUIN target population, we extracted Hospital Episode Statistics (HES) data on the number of emergency admissions in each GP practice for all people 75 and over with inpatient stays of at least 72 hours. We attributed hospital 'effort' to the practice as the weighted average CQUIN scores, where the weights were the proportion of each practice's emergency admissions (as defined above) to each hospital. The CQUIN scheme operated

from 2012-13 but data were not collected that year. Therefore, this variable was set to zero for all practices for the period before 2013-14.

As dementia registers are affected by factors other than incentive schemes, the analysis also adjusted for the following time-varying practice characteristics: practice list size (i.e. number of registered patients); the proportion of patients aged 65 and over; a measure of overall achievement on the QOF clinical domains;[20] whether the practice had a GMS (General Medical Services) contract; deciles of the practice doctor-patient ratio (full time equivalent (FTE) GPs per 1000 registered patients); practice deprivation (the percentage of practice patients living in the 20% most deprived small areas in England); and a measure of access (the percentage of patients living in urban areas).

To adjust for regional effects, we included variables for each practice's Clinical Commissioning Group (CCG) using NHS England's list of active practices. CCGs for practices that had closed were identified by linking a National Audit Office mapping file to the ONS Postcode Directory.

Statistical modelling

Our unit of analysis was the GP practice. We modelled the two practice schemes, DES18 and DIS, as binary participation indicators and evaluated their impact on the rate as defined above. Our econometric design needed to accommodate multiple incentive schemes as well as the different times the schemes were introduced and taken up.

We identified different types of participants for the 3-year DES18 scheme and for the 6-month DIS scheme, distinguishing practices into categories according to the number and

order of years of participation (Table 2). For example, a practice that only participated in the first two years of DES18 (but not the third year) was categorised as 'Y/Y/N'.

TABLE 2 HERE

Our methodological framework was a 'difference-in-differences' (DID) design.[21] We compared the difference in rates before and after the introduction of the schemes by participation type using linear mixed effects models. These models assume that, in the absence of the intervention, outcome differences between participants and non-participants are constant over time. Therefore, any differences in rates observed in the post-intervention period over and above the time trend can be attributed to the incentive scheme. This effect is measured by the coefficient on the policy variable**.** We applied a DID model with multiple periods[22-24] (technical details are in Appendix B).

The post-estimation 'predict' function was used to derive predicted rates under hypothetical participation scenarios, enabling us to estimate the national impact on dementia registers. Analyses were undertaken in Stata v14.2.

Results

Descriptive analysis

From March 2013 to March 2016, the total number of people listed on GP dementia registers in England increased from 309,461 to 432,727, i.e. a net rise of 123,266 individuals. The number diagnosed will be higher than this figure, because some newly diagnosed patients replaced individuals on the register who died.

Figure 1 shows how the gap between the mean expected and mean recorded dementia registers varied over time. There was an upward trend in recorded dementia disease registers, whereas the rate of increase in expected values was lower. Consequently, the gap between recorded and expected registers has narrowed. The periods when DES18 and DIS were active are shown as shaded areas.

FIGURE 1 HERE

From March 2013 to March 2016, the mean percentage of expected cases that was recorded on GP dementia registers increased from 51.8% to 68.6%. Figure 2 shows how this rate varied by participation in (a) DES18 and (b) DIS. By March 2016, practices participating in DES18 in all three years had a smaller gap between recorded and expected registers (i.e. higher outcome rate) on average than other practices. When comparing participation in DIS, the unadjusted data show a distinct divergence in trends around the time the intervention was introduced.

FIGURE 2 HERE

Regression analysis

While the unadjusted data suggested that practices participating in the schemes closed the gap between their recorded and expected registers at a faster rate than non-participants, the difference-in-differences analysis tested whether the observed differences were explained by confounding factors.

TABLE 3 HERE

Table 3 shows results from the linear random effects regression model applied to the balanced panel. The upward trend in the rates shown in Figure 2 is reflected in the increasing coefficients of the year dummies (beta coefficients, Appendix B). Relative to its value in 2006-07, the rate increased by 0.35 percentage points in 2007-08, by 16.4 percentage points by 2012-13 and by 31.0 percentage points by 2015-16.

The estimates for the DES18 participation groups showed no difference between the rates of practices that never participated in DES18 and the other practice groups in the pre-intervention period, with the exception of practices that participated only in the final year of the scheme (participation variables are the gamma coefficients, Appendix B). Similarly, the rates for DIS participants did not differ significantly from those of non-participants in the pre-intervention period.

The policy variables (delta coefficients, Appendix B) for DES18 were positive and significant. The DES18 scheme increased the rate for the intervention practices by 1.44 percentage points more than the increase in the rate for non-participating practices. DES18 had a significant effect in reducing the gap between recorded and expected registers ($P < 0.001$). The effect of DIS was larger with an estimated 3.59 percentage points increase in the rate ($P < 0.001$).

The effect of the hospital scheme (FAIR) was not statistically significant. Higher overall achievement on the QOF clinical domain presumably reflected better overall practice quality which helped close the gap between the recorded and expected prevalence of dementia. Practices with larger proportions of patients living in urban areas and practices with more disadvantaged patients had smaller gaps between recorded and expected dementia registers (i.e. higher rates). Practices with a higher proportion of individuals aged 65 and

above had significantly lower rates ($P < 0.001$), as did practices with a GMS contract ($P < 0.05$).

To quantify the added value of the schemes, we predicted the rates under hypothetical participation scenarios. Figure 3 shows the effects of the schemes for the 4,594 practices that participated in DES18 in all three years and that also participated in DIS. The black line shows the mean recorded rate. The other four lines depict the predicted rates under four scenarios of practice participation: i) both in DES18 and DIS; ii) only in DIS; iii) only in DES18; iv) neither in DES18 nor in DIS.

The first scenario is the mean predicted rate assuming practices participated fully in both DES18 and DIS (as they did in this subsample). The last three scenarios are hypothetical (predicted) counterfactuals: for instance, the fourth scenario predicts the rates that would have been observed had these practices not participated in either scheme.

Had all practices in the unbalanced panel participated fully in both schemes, these predicted values suggest that national dementia registers would have increased by 48,685. As participation levels were suboptimal, the net effect of the schemes was to increase registers by 40,767 (59% of which was attributable to DES18).

FIGURE 3 HERE

Sensitivity analysis

The results were robust to two sensitivity analyses (results are shown in Appendix C). First, we applied the model to the unbalanced panel of 7,529 practices totalling 74,241 practice-year observations. Both policy variables remained significant with the size of the effects very similar to the estimates from the balanced panel analysis.

The base case analysis assumed that the effects of the schemes did not persist beyond the period of active participation. In the second sensitivity analysis, we estimated a model that assumed the effect of the DES18 persisted after the practice exited the scheme. In this specification, four types of practices were defined by the year in which the practice entered the scheme (if at all). Under this design, the change in rate between 2012-13 and 2015-16 for each of the participating groups relative to the change in rate for the non-participating group did not vary by participation status each year, as in our base model. The DES18 policy effect (1.38) was significant and similar in size to the effect estimated in our base model (1.44).

Discussion

This national study of two primary care financial incentive schemes provides evidence that they helped to tackle the problem of underdiagnosis in dementia. On average, a practice's QOF dementia register rose from 28 individuals (March 2007) to 42 prior to the first scheme's introduction (March 2013), and stood at 59 when the schemes ended (March 2016). Participation in DES18, which incentivised timely assessment and support by general practice, contributed to these numbers by increasing dementia registers amongst participating practices by 1.17 individuals each year on average. Participation in the Dementia Identification Scheme (DIS), which paid practices £55 for each 'net' addition to the dementia register over a 6-month period, had an even larger impact, delivering an average net increase in registers of 2.98.

In common with most evaluations of pay-for-performance schemes, this study faced several methodological challenges[9 10] which we discuss below.

Ideally, participation in the schemes would have been randomly allocated to minimise the risk of known and unknown biases affecting results. However, difference-in-differences (DID) analysis is a good alternative when randomisation is not possible because policies have been rolled out nationally. DID assumes the intervention groups have a common trend with the control group, and the regression analysis (participation coefficients) supports that assumption. We controlled for practice characteristics we believed could affect diagnosis rates, but cannot rule out the possibility that other factors we could not measure, such as the availability of memory clinics, may have influenced results.

A key challenge in this study was defining participation in the schemes. Some practices could be clearly identified as participants or non-participants, but others were 'grey' practices that signed up to the DES18 scheme but then, apparently, did nothing – or so the assessments data suggest. Are these practices 'failed' participants (as we assumed) or non-participants? This matters because our models presuppose a clear distinction between the intervention and control groups. For DIS, NHS England provided a list of participants. The list was based on data provided by their Local Area Teams for payment purposes and was subject to numerous checks.

Our study relied on administrative datasets which are subject to the usual challenges in relation to coding errors and missing data. Data on FAIR were only available for two of the three indicators in 2015/16, so our measure only partially captures hospitals' efforts in diagnosing dementia patients. For approximately 15% of practices that had fewer than six patients in nursing homes, data were suppressed to prevent disclosure. We imputed these missing data with random values between 1 and 5.[1] In addition, the age / sex distribution

[1] Numbers of practices with imputed random value: 2009/10: 1085 (15.2%); 2010/11: 1107 (15.5%); 2011/12: 1102 (15.4%).

of nursing home patients in practices is unknown so we imputed national distributions (Appendix A).

We do not know of any previous studies quantifying the impact of schemes to boost diagnosis rates of dementia. However, the targeting of financial incentives on GPs in order to achieve quality improvements underpins the major policy initiative of the QOF programme. Research on the QOF suggests that overall this policy has been successful in promoting quality improvements – although at relatively modest levels which tend to reduce over time – in the incentivised conditions. [12 13 25 26] In our study, both DES18 and the DIS schemes appeared effective. The impact of DIS is unsurprising given the direct and time-limited nature of the incentive, which was designed to focus attention on the issue of underdiagnosis of dementia. There were calls from doctors for DIS to be withdrawn,[27] criticising it as "cash for diagnosis",[28] and "unethical and dangerous for patients";[29] nonetheless, over three-quarters of practices opted in. We also found evidence suggesting the effects of both schemes persisted after practices had exited the schemes, which supports findings from an evaluation of the withdrawal of QOF indicators. [30]

The hospital CQUIN scheme, 'FAIR', appears not to have had the expected trickledown effect on GP registers. Previous research has found little evidence of any effect of CQUIN schemes aside from those involving hip fracture.[31]

NHS England achieved its two-thirds ambition for dementia in November 2015.[5] During the years when the schemes were active, total numbers on the dementia registers increased by 123,266. However, only one third (40,767) of these additional cases are attributable to the two schemes. The schemes' effect on the number of newly diagnosed individuals will be

higher than this figure, because some additions to the register replace individuals who have died.

Total expenditure on the schemes has not been published, but we estimate the budget to be around £131m, comprising £5m for DIS [17] and £42m available in each of the three years for DES18.[32] Despite the controversy over DIS, our results illustrate that direct, targeted and time-limited financial incentives for GPs work and, as a result, quality of care has likely been enhanced for those individuals whose dementia was identified through the schemes. We also found evidence suggesting that the impact of the schemes persists after they ended, although our evaluation had limited follow-up. Policy makers may consider repeating this approach either for dementia or for other disease areas where early diagnosis is considered beneficial.

Remaining gaps in the evidence base include the wider benefits and unintended consequences of the schemes, and the true cost of delivering the schemes, as opposed to the budgeted expenditure. Although our study demonstrated the schemes were successful in closing the diagnosis gap, a comprehensive assessment of the cost-effectiveness of using financial incentives to improve diagnosis rates would require further research in these two key areas.

References

1. Etkind SN, Bone AE, Gomes B, et al. How many people will need palliative care in 2040? Past trends, future projections and implications for services. *BMC Medicine* 2017;15(1):102. doi: 10.1186/s12916-017-0860-2
2. Robinson L, Tang E, Taylor JP. Dementia: timely diagnosis and early intervention. *Bmj* 2015;350:h3029. [published Online First: 2015/06/17]
3. Department of Health. Living Well With Dementia: a national dementia strategy. London: Department of Health 2009.
4. Department of Health. The Prime Minister's Challenge on Dementia: delivering major improvements in dementia care and research by 2015. London: Department of Health 2012.
5. Burns A, Bagshaw P. A new dementia currency in primary care, 2 March. blog: NHS England, 2016.
6. NHS England. Dementia Identification Scheme: Guidance and Frequently Asked Questions. In: Operations NEC, ed. Gateway reference number: 02504. 12 November 2014 ed, 2014:9.
7. Department of Health. NHS to tackle long waits for dementia assessments. *Press Release* 2014(28 February)
8. NHS England. Everyone counts : planning for patients 2014/15 to 2018/19. Leeds: NHS England, 2013.
9. Flodgren G, Eccles MP, Shepperd S, et al. An overview of reviews evaluating the effectiveness of financial incentives in changing healthcare professional behaviours and patient outcomes. *Cochrane Database Syst Rev* 2011(7):CD009255.
10. Scott A, Sivey P, Ait Ouakrim D, et al. The effect of financial incentives on the quality of health care provided by primary care physicians. *Cochrane Database Syst Rev* 2011(9):CD008451.
11. Goddard M, Kasteridis P, Jacobs R, et al. Bridging the gap: The impact of quality of primary care on duration of hospital stay for people with dementia. *Journal of Integrated Care* 2016;24(1):15-25. doi: doi:10.1108/JICA-11-2015-0045
12. Kasteridis P, Mason A, Goddard M, et al. The influence of primary care quality on hospital admissions for people with dementia in England: a regression analysis. *PloS One* 2015;10(3):e0121506.
13. Kasteridis P, Mason A, Goddard M, et al. Risk of Care Home Placement following Acute Hospital Admission: Effects of a Pay-for-Performance Scheme for Dementia. *PLoS ONE [Electronic Resource]* 2016;11(5):e0155850. doi: http://dx.doi.org/10.1371/journal.pone.0155850
14. NHS England. General Practice Outcome Standards: Methodology for Assessing Variation (Version 2.1). An Introduction to an England Approach to Improve Quality, Access and Patient Experience in General Practice: NHS England, 2016:13.
15. NHS Commissioning Board. Enhanced Service Specification: Facilitating timely diagnosis and support for people with dementia: NHS Commissioning Board, 2013:10.
16. NHS England. Enhanced Service Specification: Dementia Identification Scheme. Leeds: NHS England, 2014:15.
17. Millett D. Practices to earn £55 per extra patient diagnosed with dementia. GP Online, 2014.

18. Matthews FE, Arthur A, Barnes LE, et al. A two-decade comparison of prevalence of dementia in individuals aged 65 years and older from three geographical areas of England: results of the Cognitive Function and Ageing Study I and II. *Lancet* 2013;382(9902):1405-12.
19. NHS Commissioning Board. Commissioning for quality and innovation (CQUIN): 2013/14 guidance, 2012:37.
20. Santos R, Gravelle H, Propper C. Does Quality Affect Patients' Choice of Doctor? Evidence from England. *The Economic Journal* 2017;127(600):445–94 doi: 10.1111/ecoj.12282 [published Online First: 23 February 2016]
21. Ashenfelter O, Card D. Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs. *Review of Economics and Statistics* 1985;67(4):648-60.
22. Bertrand M, Duflo E, Mullainathan S. How Much Should We Trust Differences-in-Differences Estimates? *Quarterly Journal of Economics* 2004;119(1):249-75.
23. Hansen CB. Generalized Least Squares Inference in Panel and Multilevel Models with Serial Correlation and Fixed Effects. *Journal of Econometrics* 2007;140(2):670-94.
24. Wooldridge J. New Developments in Econometrics. Lecture 11: Difference-in-Differences Estimation. *Cemmap Lecture Notes* 2009
25. Campbell SM, Reeves D, Kontopantelis E, et al. Effects of pay for performance on the quality of primary care in England. *New England Journal of Medicine* 2009;361(4):368-78.
26. Doran T, Kontopantelis E, Valderas JM, et al. Effect of financial incentives on incentivised and non-incentivised clinical activities: Longitudinal analysis of data from the UK Quality and Outcomes Framework. *Bmj* 2011;342(7814):d3590. doi: http://dx.doi.org/10.1136/bmj.d3590 [published Online First: 28 June]
27. Brunet M. An open letter to Simon Stevens, NHS chief executive, and Alistair Burns, national clinical lead for dementia. *BMJ : British Medical Journal* 2014;349 doi: https://doi.org/10.1136/bmj.g6666
28. Kmietowicz Z. Doctors condemn "unethical" 55 payment for every new dementia diagnosis. *Bmj* 2014;349:g6424.
29. Kmietowicz Z. Axe the £55 payment for dementia diagnosis, say doctors. *Bmj* 2014;349:g6614.
30. Kontopantelis E, Springate D, Reeves D, et al. Withdrawing performance indicators: retrospective analysis of general practice performance under UK Quality and Outcomes Framework. *Bmj* 2014;348:g330. [published Online First: 2014/01/29]
31. McDonald R, Zaidi S, Todd S, et al. Evaluation of the Commissioning for Quality and Innovation Framework: Final Report: University of Nottingham, University of Manchester, 2013.
32. NHS England. NHS England, Government and BMA agree new GP contract for 2016/17. *NHS England News* 2016;19 February

Acknowledgements

We thank participants at the Health Economics Study Group (HESG) meeting in Birmingham (January 2017) for helpful comments on an earlier draft of this paper; our advisory groups for valuable insights and advice; and NHS England for providing the list of DIS participants. We are also grateful to two reviewers for their insightful comments.

The paper is based on independent research commissioned and funded by the NIHR Policy Research Programme (Economics of Social and Health Care; PRP Ref: 103 0001). The views expressed in the publication are those of the authors and not necessarily those of the NHS, the NIHR, the Department of Health, arm's length bodies or other government departments.

Conflict of interest

The authors have no financial or personal conflicts to declare.

Ethical approval

Not required

Table 1: Descriptive statistics for the outcome and explanatory variables: Balanced panel, 2006-07 to 2015-16

Variable	Mean	Std. Dev.	Min	Max	N
Recorded dementia register	39.75	36.46	0	631	71,420
Expected dementia register	80.91	64.34	0.02	1135.91	71,420
Mean 'rate' (100*recorded / expected)	49.07	21.28	0	100	71,420
DES18 participation (%): 3 years	79.11				71,420
DES18 participation (%): 2 years	15.93				71,420
DES18 participation (%): 1 year	3.43				71,420
DIS participation (%)	75.93				71,420
Hospital effort (2013-14 to 2015-16 only) *	86.06	17.14	0	100	21,426
Practice list size (1,000)	7.28	4.23	0.01	60.38	71,420
% practice patients 65 or older	16.05	5.74	0.00	47.99	71,420
Weighted achievement on the QOF clinical domain	80.73	4.63	0.05	99.79	71,420
GMS contract	0.59	0.49	0	1	71,420
Full-time equivalent GPs† per 1000 patients	0.57	1.01	0.01	266.67	71,420
% patients living in 20% most deprived areas	23.12	26.20	0.00	99.65	71,420
% patients living in urban areas	82.71	32.45	0.00	100	71,420

*Hospital effort assumed to be zero from 2006-07 to 2012-13; N = practice-years

†excluding retainers/registrars

Table 2: Participation in DES18 or DIS: balanced panel, 2006-07 to 2015-16

		Practice-years	Percent	Mean dementia register
DES18 participation	Years of participation: 3	56,500	79.11	42.67
	Y/Y/Y	56,500	79.11	42.67
	Years of participation: 2	11,380	15.93	29.29
	Y/Y/N	1,280	1.79	33.08
	Y/N/Y	1,420	1.99	28.31
	N/Y/Y	8,680	12.15	28.89
	Years of participation: 1	2,450	3.43	25.54
	Y/N/N	440	0.62	31.82
	N/Y/N	700	0.98	22.85
	N/N/Y	1,310	1.83	23.63
	No participation	1,090	1.53	31.09
	N/N/N	1,090	1.53	31.09
	Total	71,420	100	39.75
DIS participation	No	17,190	24.07	34.45
	Yes	54,230	75.93	41.43
	Total	71,420	100	39.75

Note: as this is a balanced panel, the number of practices contributing data can be inferred by dividing practice-years by 10.

Table 3: Linear random effects results: Balanced panel, 2006-07 to 2015-16

Variable	Coefficient	95% CI
FY is 2006-07 (ref.)		
FY is 2007-08	0.345**	[0.096, 0.593]
FY is 2008-09	2.397***	[2.073, 2.721]
FY is 2009-10	5.795***	[5.427, 6.162]
FY is 2010-11	7.908***	[7.508, 8.307]
FY is 2011-12	12.556***	[12.121, 12.992]
FY is 2012-13	16.419***	[15.934, 16.903]
FY is 2013-14	19.022***	[17.563, 20.482]
FY is 2014-15	26.562***	[24.814, 28.311]
FY is 2015-16	30.977***	[29.329, 32.624]
Practice participation in DES18 in 2013-14/2014-15/ 2015-16 (participation is indicated by Yes (Y), non-participation by No (N)		
N/N/N (ref.)		
Y/Y/Y	2.010	[-0.638, 4.658]
Y/Y/N	1.275	[-2.411, 4.960]
Y/N/Y	-0.207	[-3.562, 3.148]
Y/N/N	-0.909	[-5.114, 3.295]
N/Y/Y	-0.720	[-3.523, 2.082]
N/Y/N	-1.843	[-6.382, 2.695]
N/N/Y	-3.438*	[-6.843, -0.033]
Participation in DIS	0.770	[-0.030, 1.570]
Policy variable (DES18)	1.439***	[0.669, 2.210]
Policy variable (DIS)	3.594***	[2.785, 4.403]
Hospital effort (FAIR)	0.008	[-0.007, 0.024]
Practice list size (in 1,000)	0.255***	[0.172, 0.338]
% of practice patients 65 or older	-0.559***	[-0.651, -0.467]
QOF achievement in the clinical domain	0.301***	[0.253, 0.349]
GMS contract	-0.650*	[-1.187, -0.112]
Deciles of FTE GPs per 1,000 patients		
Decile 1 (ref.)		
Decile 2	0.096	[-0.590, 0.781]
Decile 3	-0.013	[-0.702, 0.675]
Decile 4	0.077	[-0.609, 0.764]
Decile 5	-0.066	[-0.756, 0.624]
Decile 6	0.182	[-0.515, 0.879]
Decile 7	0.294	[-0.397, 0.985]
Decile 8	0.168	[-0.534, 0.871]
Decile 9	0.385	[-0.348, 1.118]
Decile 10	0.518	[-0.287, 1.322]
% of practice patients living in 20% most deprived areas	0.033**	[0.012, 0.054]
% of practice patients living in urban areas	0.019**	[0.007, 0.031]
Within R-squared	0.489	
Between R-squared	0.196	
Overall R-squared	0.360	
Standard deviation of practice random effect (sigma_u)	12.204	
Intraclass correlation (rho)	0.508	

95% confidence intervals in brackets; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; Models also adjust for CCG (results not shown); R2=36%

Figure 1: Gap between mean recorded dementia register and mean expected dementia register

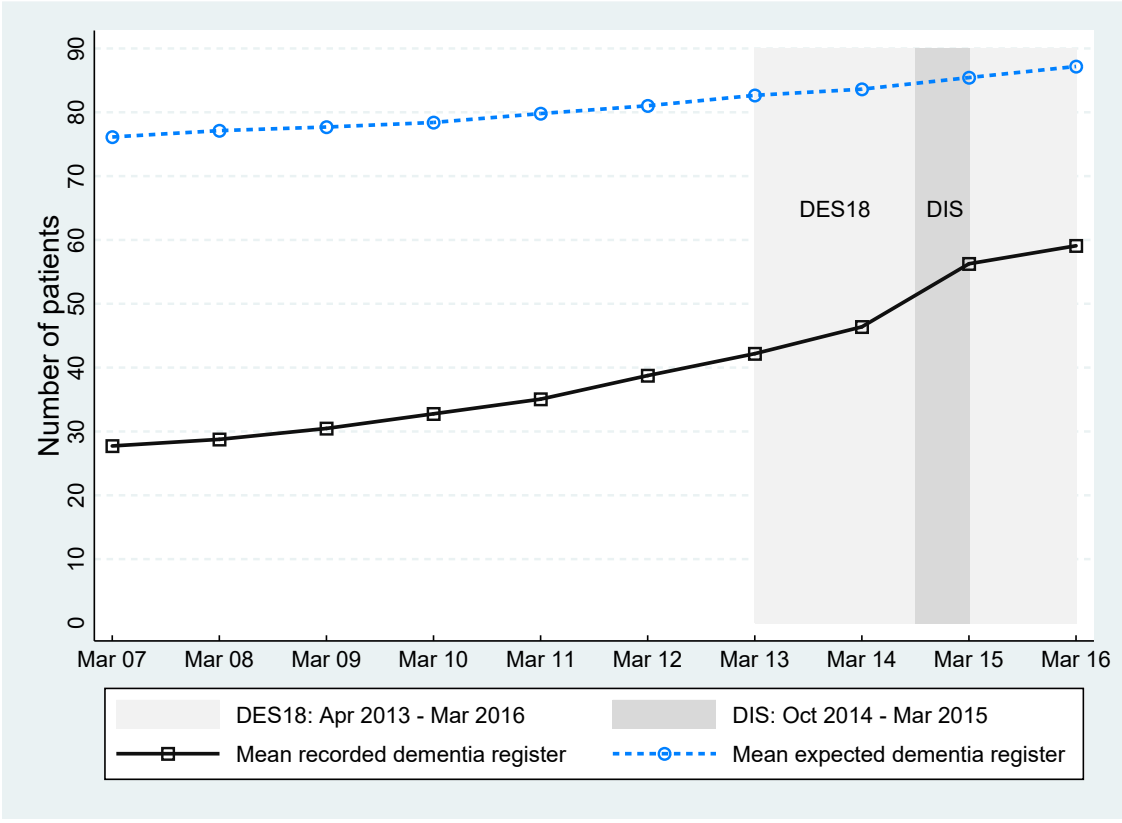


Figure 2: Trends in mean practice outcome rates by years of participation in (a) DES18 (b) DIS

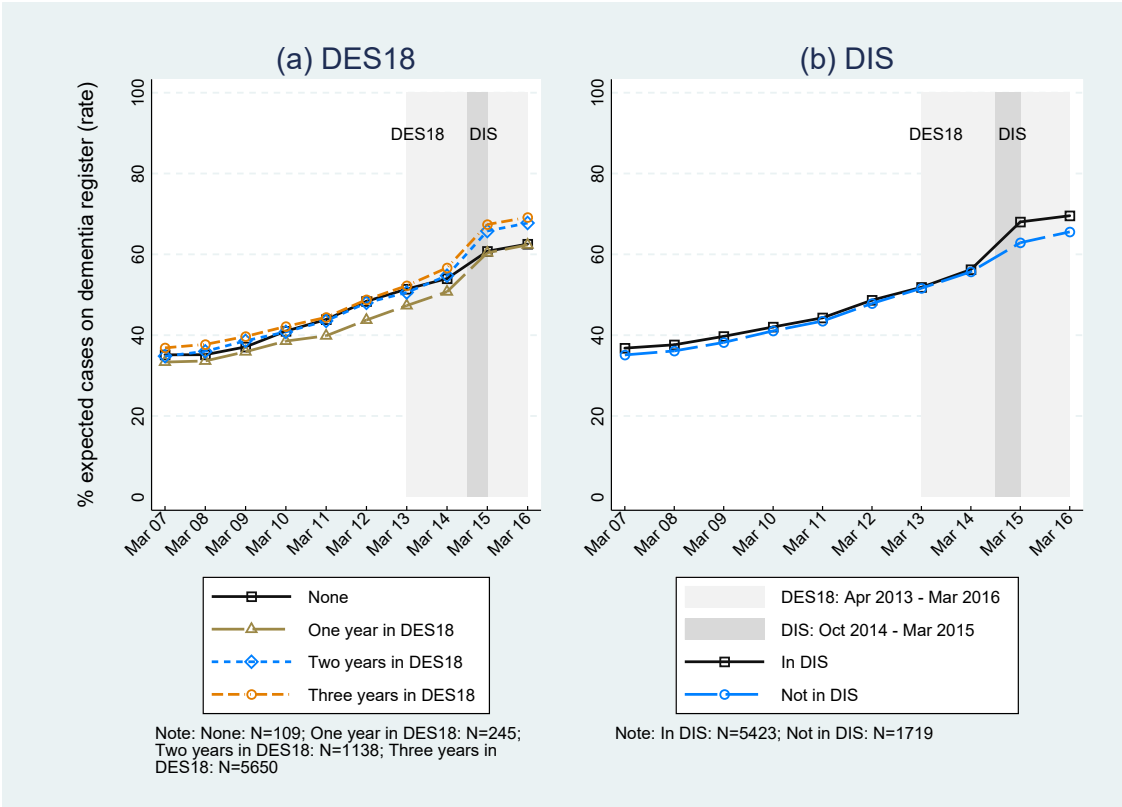
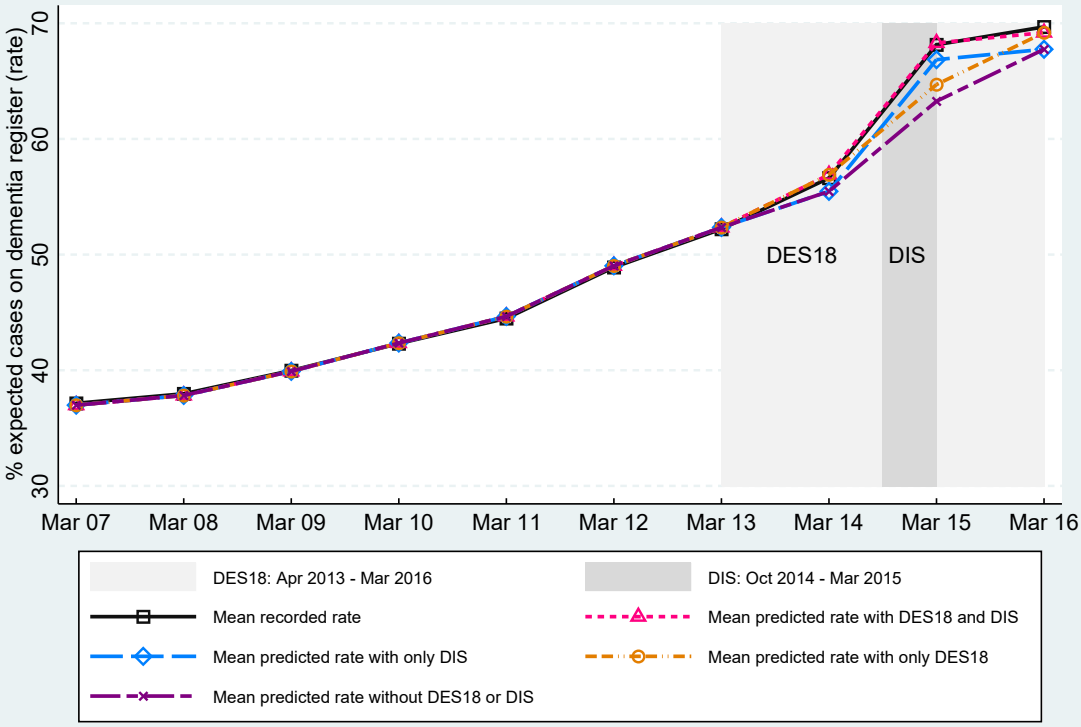


Figure 3: Trends in mean of the recorded and predicted practice outcome rates: DES18 and DIS



Note: Graph shows practices participating in DES18 for three years and also in DIS (N=4594)

Appendix A: Datasets used for the analysis

Dataset	Reporting level	Years	Type of variable(s) derived	Details of variables
Quality and Outcomes Framework (QOF)	GP practice	2006-07 – 2015-16	Dependent variable (numerator)	Source of dementia register counts
			Control variables	Overall QOF achievement on clinical domain Practice list size % practice patients 65+
Cognitive Function and Ageing Study II (CFAS II) [1]	Residential setting	2008 –2011 [a]	Dependent variable (denominator)	Consensus estimates of dementia prevalence by age, gender and setting (nursing home / community). Used to derive expected dementia registers.
Numbers of Nursing Home GP Patients by practice in England [b]	GP practice	2006 – 2015	Dependent variable (denominator)	Used to derive expected dementia registers.
ONS – census data	Residential setting	2011	Dependent variable	Proportion of patients in different age and gender bands (65+) in care homes. Used to derive expected dementia registers.
General and Personal Medical Services dataset (GMS) [c]	GP practice	2011/12 – 2015-16	Dependent variable (denominator).	Proportion of practice patients in different age and gender bands (65+) used to derive expected dementia registers. GMS contract status.
Dementia (facilitating timely diagnosis and support for people with dementia) – assessments data	GP practice	2013-14 – 2015-16	Policy variable	Used to identify participation in Directed Enhanced Services (DES18): Facilitating Timely Diagnosis and Support for People with Dementia
Dementia Assessment and Referral Data Collection	NHS Trust	2013-14 – 2015-16	Policy variable	Published monthly and quarterly. Used to identify NHS Trusts participating in the FAIR (Find; Assess & Investigate; Refer) CQUIN scheme
Attribution Dataset (ADS)	GP practice	2006-07 – 2015-16	Control variables	Numbers of practice patients in each LSOA. Used to generate practice-level weighted averages of rurality and deprivation
Hospital Episode Statistics (HES)	Patients	2013-14 – 2015-16	Control variable	Used to calculate weights for deriving practice-level FAIR variable.
ONS – urban	LSOA	2004, 2011	Control variable	Source of urban classifications. Combined with ADS to derive practice rurality measure. 2004 data used for missing values in unbalanced panel.
ONS –deprivation	LSOA	2010, 2015	Control variable	Source of IMD classifications. Combined with ADS to derive practice deprivation measure. 2010 data used for missing

Dataset	Reporting level	Years	Type of variable(s) derived	Details of variables
CCG code	GP practice	2006-07 – 2015-16	Control variable	values in unbalanced panel. Practice CCG code

[a] These are the survey years; a unique value was derived for each age/gender/setting band. CFAS II prevalence rates are used in NHS England's performance framework for CCGs to identify areas where dementia diagnosis rates are below expected levels.[2]

[b] Positive values <6 are suppressed, so were imputed using a random value between 1 and 5

[c] The method of GMS data collection changed in 2015-16 and data are missing for around 15% of practices.

Acronyms: ADS: attribution dataset; CCG: clinical commissioning group; CQUIN: Commissioning for Quality and Innovation; DSA: Data Sharing Agreement; GMS: General and Personal Medical Services dataset; IMD: index of multiple deprivation; LSOA: lower-layer super output area; ONS: Office for National Statistics; QOF: quality and outcomes framework

Data sharing

All the data used in the analyses can be freely downloaded from the web, with the following exceptions:

1. General and Personal Medical Services dataset (GMS) (2006/7 to 2012/13) were accessed via a Data Sharing Agreement between the Centre for Health Economics, University of York, and NHS Digital;
2. Numbers of practice patients in nursing homes, 2006/07 to 2012/13: supplied by NHS England;
3. Hospital Episode Statistics (HES), 2006/07 to 2015/16: provided by NHS Digital under a Data Sharing Agreement;
4. Attribution Dataset (ADS), 2006/07 to 2013/14: provided under a Data Sharing Agreement with NHS England;
5. List of DIS participants (supplied by NHS England).

References:

1. Matthews FE, Arthur A, Barnes LE, et al. A two-decade comparison of prevalence of dementia in individuals aged 65 years and older from three geographical areas of England: results of the Cognitive Function and Ageing Study I and II. *Lancet* 2013;382(9902):1405-12.
2. NHS England. CCG improvement and assessment framework 2016/17: Technical Annex, 2016.

Appendix B: Difference-in-differences (DID) model

The simplest DID design is a linear mixed effects model with two time periods – pre and post intervention years – and a single intervention. In our setting, practices indexed by i, can participate in DES18 schemes at any or all of three time periods (2013-14, 2014-15, 2015-16) and can concurrently participate in DES18 and DIS in 2014-15. To account for this, we apply a DID model with multiple periods. [1-3] The specification includes a full set of time period dummies, seven binary indicators indicating different types of DES18 participants for the 3-year DES18 scheme (as defined in Supplementary Table 2), a binary indicator of DIS participation, and two binary policy variables: P_t^1 for DES18 and P_t^2 for DIS (in analogy to the interaction term in the simple case) defined as unity for practices and time periods subject to the interventions.

The random error u_i is introduced to capture practice specific time invariant effects and follows a normal distribution $N(0,\sigma_u^2)$

$$\begin{aligned} y_{it} &= \alpha + \beta_1 t_{2006} + \ldots + \beta_{10} t_{2015} \\ &+\gamma_1 DES_{YYY} + \gamma_2 DES_{YYN} + \gamma_3 DES_{YNY} + \gamma_4 DES_{YNN} + \gamma_5 DES_{NYY} + \gamma_6 DES_{NYN} + \gamma_7 DES_{NNY} + \gamma_8 DIS \\ &+\delta_1 P_t^1 + \delta_2 P_t^2 + bX_{it} + u_i + \varepsilon_{it} \end{aligned} \qquad (1)$$

where $u_i \sim N(0,\sigma_u^2)$, $\varepsilon_{it} \sim N(0,\sigma_\varepsilon^2)$ and ε_{it} is independent of u_i.

The compound error term $e_{it} = u_i + \varepsilon_{it}$ is independent across practices but not within a practice. The within practice correlation is the sum of intraclass correlation

$\rho = \dfrac{\sigma_u^2}{\sigma_u^2 + \sigma_\varepsilon^2}$ plus serial correlation $E(\varepsilon_{it}\varepsilon_{is})$

To account for within practices serial correlation we cluster standard errors at the same level as the random effect (i.e. at practice level).

The beta coefficients of the time period dummy variables $\beta_1,\ldots,\beta_{10}$ explain the effect of aggregate temporal factors that cause changes in the outcome (i.e. the rate), even in the absence of interventions. The gamma coefficients $\gamma_1,\ldots,\gamma_7$ capture differences in rates with respect to the DES18 reference group (never participating in DES18, NNN) prior to the intervention.

For instance, the difference in average outcomes between practices in the group YYY and practices in the group NNN in year 2007 after adjusting for practice characteristics is:

$$\bar{Y}_{t=2007}^{YYY} - \bar{Y}_{t=2007}^{NNN} = (\alpha + \beta_2 + \gamma_1) - (\alpha + \beta_2) = \gamma_1$$

Similarly, γ_8 captures differences in rates between participants and non-participants in DIS prior to the intervention. The policy effects of the incentive schemes are captured by the delta coefficients δ_1 and δ_2.

For instance, the difference in outcomes for the group YYY between 2012-13 and 2013-14 is:

$\bar{Y}_{t=2013}^{YYY} - \bar{Y}_{t=2012}^{YYY} = (\alpha + \beta_8 + \gamma_1 + \delta_1) - (\alpha + \beta_7 + \gamma_1) = (\beta_8 - \beta_7) + \delta_1$ and the difference in outcomes for the group NNN between 2012-13 and 2013-14 is:

$$\bar{Y}_{t=2013}^{NNN} - \bar{Y}_{t=2012}^{NNN} = (\alpha + \beta_8) - (\alpha + \beta_7) = (\beta_8 - \beta_7)$$

Therefore, the difference-in differences estimate (which we call DES18 policy variable) is:

$$(\bar{Y}_{t=2013}^{YYY} - \bar{Y}_{t=2012}^{YYY}) - (\bar{Y}_{t=2013}^{NNN} - \bar{Y}_{t=2012}^{NNN}) = \delta_1$$

Note, that γ_1 appears in the calculation of the average outcome for the participating group in the post-intervention year 2013 ($\bar{Y}_{t=2013}^{YYY}$) but cancels out in the DiD estimate.

The base model allows for practices entering the DES18 scheme at different time periods and the effects of DES18 are assumed not to persist once the practice exits the scheme. As an example, Table 1 presents the year-to-year effects for two types of practices: those that enter the scheme in 2013/14, continue in 2014/15, and exit the scheme in 2015-16 ($DES18_{YYN}$); and those that take up DES18 in 2013/14, interrupt their participation in 2014/15, and re-join the scheme in 2015-16 ($DES18_{YNY}$). For the $DES18_{YYN}$ group, the scheme is in effect in 2014-15 but not in 2015-16 and therefore the change in rate from 2014/15 to 2015-16 compared to the non-participating control group is $-\delta_1$ (third row, third column). For the $DES18_{YNY}$ group the change is δ_1 (third row, sixth column) implying a 2 δ_1 percentage points higher rate for the $DES18_{YNY}$ group compared to the $DES18_{YYN}$ group.

Appendix Table 1: Year to year effects of the DES18 scheme for the $DES18_{YYN}$ and $DES18_{YNY}$ groups

	$DES18_{YYN}$			$DES18_{YNY}$		
	2013-14	2014-15	2015-16	2013-14	2014-15	2015-16
2011-13 (Pre-DES18)	δ_1	δ_1	0	δ_1	0	δ_1
2013-14		0	$-\delta_1$		$-\delta_1$	0
2014-15			$-\delta_1$			δ_1

References

1. Bertrand M, Duflo E, Mullainathan S. How Much Should We Trust Differences-in-Differences Estimates? *Quarterly Journal of Economics* 2004;119(1):249-75.
2. Hansen CB. Generalized Least Squares Inference in Panel and Multilevel Models with Serial Correlation and Fixed Effects. *Journal of Econometrics* 2007;140(2):670-94.
3. Wooldridge J. New Developments in Econometrics. Lecture 11: Difference-in-Differences Estimation. *Cemmap Lecture Notes* 2009

Appendix C

Sensitivity analysis 1: results using an unbalanced panel, 2006-07 to 2015-16

Variable	Coefficient	95% CI
FY is 2006-07 (ref.)		
FY is 2007-08	0.363**	[0.110, 0.615]
FY is 2008-09	2.409***	[2.081, 2.736]
FY is 2009-10	5.913***	[5.544, 6.282]
FY is 2010-11	7.997***	[7.588, 8.406]
FY is 2011-12	12.576***	[12.134, 13.019]
FY is 2012-13	16.349***	[15.856, 16.843]
FY is 2013-14	18.669***	[17.156, 20.183]
FY is 2014-15	25.997***	[24.179, 27.815]
FY is 2015-16	30.393***	[28.681, 32.105]
Practice participation in DES18 in 2013-14/2014-15/ 2015-16 (participation is indicated by Yes (Y), non-participation by No (N)		
N/N/N (ref.)		
Y/Y/Y	2.314	[-0.574, 5.203]
Y/Y/N	2.017	[-1.942, 5.976]
Y/N/Y	1.378	[-2.191, 4.946]
Y/N/N	-0.148	[-4.698, 4.403]
N/Y/Y	-0.239	[-3.284, 2.806]
N/Y/N	-1.319	[-5.762, 3.124]
N/N/Y	-3.249	[-6.834, 0.335]
Participation in DIS	0.802	[-0.017, 1.622]
Policy variable (DES18)	1.405***	[0.619, 2.191]
Policy variable (DIS)	3.583***	[2.768, 4.399]
Hospital effort (FAIR)	0.012	[-0.004, 0.028]
Practice list size (in 1,000)	0.217***	[0.132, 0.302]
% of practice patients 65 or older	-0.449***	[-0.565, -0.333]
QOF achievement in the clinical domain	0.299***	[0.250, 0.348]
GMS contract	-0.950***	[-1.497, -0.403]
Deciles of FTE GPs per 1,000 patients		
Decile 1 (ref.)		
Decile 2	0.119	[-0.558, 0.797]
Decile 3	-0.165	[-0.845, 0.516]
Decile 4	-0.007	[-0.688, 0.675]
Decile 5	-0.183	[-0.867, 0.501]
Decile 6	0.136	[-0.556, 0.828]
Decile 7	0.200	[-0.488, 0.889]
Decile 8	0.057	[-0.642, 0.757]
Decile 9	0.301	[-0.430, 1.033]
Decile 10	0.755	[-0.057, 1.567]

Variable	Coefficient	95% CI
% of practice patients living in 20% most deprived areas	0.044***	[0.021, 0.067]
% of practice patients living in urban areas	0.022***	[0.009, 0.034]
Within R-squared	0.468	
Between R-squared	0.202	
Overall R-squared	0.347	
Standard deviation of practice random effect (sigma_u)	12.646	
Intraclass correlation (rho)	0.508	

95% confidence intervals in brackets; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; Models also adjust for CCG (results not shown)

Sensitivity analysis 2: results assuming persistent effects. Balanced panel over 2006-07 to 2015-16

Variable	Coefficient	95% CI
FY is 2006-07 (ref.)		
FY is 2007-08	0.343**	[0.095, 0.592]
FY is 2008-09	2.395***	[2.071, 2.719]
FY is 2009-10	5.794***	[5.427, 6.162]
FY is 2010-11	7.905***	[7.506, 8.305]
FY is 2011-12	12.553***	[12.117, 12.989]
FY is 2012-13	16.414***	[15.930, 16.899]
FY is 2013-14	19.058***	[17.544, 20.572]
FY is 2014-15	26.516***	[24.692, 28.339]
FY is 2015-16	30.968***	[29.246, 32.690]
Practice participation in DES18 in2013-14/2014-15/ 2015-16 (participation is indicated by Yes (Y), non-participation by No (N)		
N/N/N (ref.)		
Y/-/- (i.e. Y/N/N or Y/N/Y or Y/Y/N or Y/Y/Y)	1.756	[-0.885, 4.397]
N/Y/- (i.e. N/Y/N or N/Y/Y)	-0.915	[-3.706, 1.875]
N/N/Y	-3.480*	[-6.885, -0.075]
Participation in DIS	0.966*	[0.174, 1.757]
Policy variable (DES18)	1.379**	[0.473, 2.285]
Policy variable (DIS)	3.659***	[2.848, 4.470]
Hospital effort (FAIR)	0.009	[-0.007, 0.024]
Practice list size (in 1,000)	0.258***	[0.176, 0.341]
% of practice patients 65 or older	-0.556***	[-0.648, -0.464]
QOF achievement in the clinical domain	0.302***	[0.254, 0.349]
GMS contract	-0.656*	[-1.194, -0.117]
Decile 1 (ref.)		
Decile 2	0.093	[-0.593, 0.779]
Decile 3	-0.014	[-0.703, 0.674]
Decile 4	0.073	[-0.614, 0.760]
Decile 5	-0.068	[-0.758, 0.623]

Variable	Coefficient	95% CI
Decile 6	0.180	[-0.517, 0.878]
Decile 7	0.293	[-0.398, 0.985]
Decile 8	0.168	[-0.535, 0.871]
Decile 9	0.385	[-0.347, 1.118]
Decile 10	0.517	[-0.287, 1.322]
% of practice patients living in 20% most deprived areas	0.033**	[0.012, 0.055]
% of practice patients living in urban areas	0.019**	[0.007, 0.031]
Within R-squared	0.489	
Between R-squared	0.195	
Overall R-squared	0.360	
Standard deviation of practice random effect (sigma_u)	12.206	
Intraclass correlation (rho)	0.508	

95% confidence intervals in brackets; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; Models also adjust for CCG (results not shown)