

Evaluating the Effectiveness of a Local Primary Care Incentive Scheme: A Difference-in-Differences Study

Medical Care Research and Review
2022, Vol. 79(3) 394–403
© The Author(s) 2021



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/10775587211035280
journals.sagepub.com/home/mcr



Esmael Khedmati Morasae^{1*}, Tanith C. Rose^{2*} , Mark Gabbay²,
Laura Buckels³, Colette Morris³, Sharon Poll³, Mark Goodall²,
Rob Barnett⁴, and Ben Barr²

Abstract

National financial incentive schemes for improving the quality of primary care have come under criticism in the United Kingdom, leading to calls for localized alternatives. This study investigated whether a local general practice incentive-based quality improvement scheme launched in 2011 in a city in the North West of England was associated with a reduction in all-cause emergency hospital admissions. Difference-in-differences analysis was used to compare the change in emergency admission rates in the intervention city, to the change in a matched comparison population. Emergency admissions rates fell by 19 per 1,000 people in the years following the intervention (95% confidence interval [17, 21]) in the intervention city, relative to the comparison population. This effect was greater among more disadvantaged populations, narrowing socioeconomic inequalities in emergency admissions. The findings suggest that similar approaches could be an effective component of strategies to reduce unplanned hospital admissions elsewhere.

Keywords

emergency admission, quality improvement scheme, incentive, general practice

Introduction

The United Kingdom has been at the forefront of developing incentive schemes to improve the quality and efficiency of general practice. The Quality and Outcomes Framework (QOF) was introduced nationally in the United Kingdom in 2004, allocating 25% of general practitioners (GP) income based on the achievement of quality targets for the management of chronic and severe conditions, as well as indicators related to practice organization and patient experience (Guthrie et al., 2006; Roland & Guthrie, 2016; see <https://qof.digital.nhs.uk/> for a description of the national indicators).

The QOF has, however, received increasing criticism with several studies suggesting it has not led to improvements in the quality of primary care and population health metrics, and that it is poor value for money (Ashworth & Gulliford, 2017; Campbell et al., 2009; Doran et al., 2006; Doran et al., 2011; Doran et al., 2014; Doran et al., 2017; Fleetcroft et al., 2010; Forbes et al., 2017; Gillam et al., 2012; Grigoroglou et al., 2018; Hamel et al., 2014; Kontopantelis et al., 2015; Lester et al., 2010; Mandavia et al., 2017; Marshall & Roland, 2017; Raleigh, 2018;

Roland, 2016; Roland & Guthrie, 2016; Roland & Olesen, 2016; Ryan et al., 2016; Steel & Shekelle, 2016; Thorne, 2016). There have been a considerable number of studies and systematic reviews investigating the effects of incentive schemes on quality, cost, efficiency, and equity of primary care provision (Alshamsan et al., 2010; Alshamsan et al., 2012; Conrad & Perry, 2009; Damberg et al., 2014; Doran et al., 2006; Eijkenaar et al., 2013; Emmert et al., 2012; Flodgren et al., 2011; Girault et al., 2017; Glidewell et al., 2015; Grimaldi, 2018; Herbst et al., 2018; Houle et al., 2012;

This article, submitted to *Medical Care Research and Review* on May 31, 2020, was revised and accepted for publication on July 2, 2021.

¹University of Exeter Business School, Exeter, UK

²University of Liverpool, Liverpool, UK

³Liverpool Clinical Commissioning Group, Liverpool, UK

⁴Liverpool Local Medical Committee, Liverpool, UK

*Joint first author

Corresponding Author:

Tanith C. Rose, Research Associate in Health Data Analysis, Department of Public Health, Policy, and Systems, Institute of Population Health Sciences, University of Liverpool, Liverpool L69 3BX, UK.
Email: tanith.rose@liverpool.ac.uk

Kondo et al., 2016; Lee et al., 2011; Markovitz & Ryan, 2017; Mendelson et al., 2017; Pandya et al., 2018; Petersen et al., 2006; Petersen et al., 2013; Roland & Dudley, 2015; Scott et al., 2011; Soucat et al., 2017; Van Herck et al., 2010; Walker et al., 2010; Witter et al., 2012; Yuan et al., 2017), often with inconsistent results (Bardach et al., 2013; Emmert et al., 2012; Houle et al., 2012; Roland & Dudley, 2015). Whilst there is some evidence that they can improve process measures, there is a lack of evidence for improvements in patient health indicators (Flodgren et al., 2011). Some experimental economic studies have, however, suggested that such schemes may lead to more optimal use of resources (Green 2014; Hennig-Schmidt et al., 2011; Brosig-Koch et al., 2016; Xi et al., 2019). Studies investigating the health inequalities impact of such schemes have been also mixed (Alshamsan et al., 2010), with some studies indicating that they increase inequalities (Alshamsan et al., 2010; Roland & Dudley, 2015), whilst Doran et al. (2008) showed that QOF was successful in reducing the inequalities indicators for the first 3 years of deployment, though it was not persistent in its positive effects (Dixon & Khachatryan, 2010). Due to this mounting criticism, there have been suggestions to amend the QOF (Doran et al., 2017; Minchin et al., 2018; Pandya et al., 2018), including handing control over to local primary care systems (Hackett et al., 2014; Glidewell et al., 2015). There is, however, a lack of evidence available to inform local incentive and quality improvement schemes (Hackett et al., 2014).

We therefore aimed to investigate the effect on all-cause emergency admissions of a local primary care quality improvement scheme implemented in 92 GP practices in a northern U.K. city (see The Intervention section for further detail and a conceptual framework). Using data on emergency admission rates for small neighborhoods, we investigated the change in rates before and after the intervention in the study site and in a set of matched comparator areas.

New Contributions

Incentive schemes have been viewed as valued ways of funding primary care services amidst overall reductions in primary care financing. Some evidence suggests that national incentive schemes in the United Kingdom have had limited success in improving patient outcomes, for example, reductions in unplanned hospitalizations. This has prompted calls for localized alternatives to capture the specificities of local contexts. Evaluation evidence of localized incentive schemes is lacking, especially in relation to effects on health inequalities. We therefore evaluated a local incentive scheme developed and scaled up in a city in the North West of England. This local scheme was implemented in addition to the nationwide incentive scheme and involved four key components: (1) an increase in the level of funding for general practices, (2) distribution of that funding proportional to the level of need in the populations served by each practice, (3) a set of

locally tailored performance targets, and (4) a governance framework overseen by the GPs themselves. The findings suggest that these components were effective in reducing the level of and inequalities in all-cause emergency hospitalizations. Quality improvement incentive schemes, including these components may be effective in other settings.

Method

Setting

The intervention was implemented across a city in North West of England, with a population of 550,000. It is the fourth most deprived local authority in England based on the Indices of Multiple Deprivation (Department for Communities and Local Government 2015). Healthy life expectancy is only 58 years—significantly lower than the 63 years nationally. The intervention included all 92 GP practices serving this population.

Study Design

This study was a longitudinal matched controlled study using difference-in-differences analysis. This difference-in-differences method controls for all time-invariant differences between the intervention and control populations (supplementary file, Appendix 1, available online). The key assumption of difference-in-differences analysis is the parallel trends assumption. If the trend in the outcome in the intervention and control populations would have been parallel in the absence of the intervention, then the difference between the change in the outcomes and between the two groups provides an unbiased estimate of the interventions effect (Dimick & Ryan, 2014).

Lower Layer Super Output Area (LSOA) years were the units of analysis. LSOAs are small geographical areas used by the UK's Office for National Statistics (ONS), each typically containing a population of about 1,500 people. England is divided into just over 30,000 LSOAs. Two hundred and ninety eight LSOAs cover the entire population of the intervention city. Each of these intervention LSOAs was matched with five control LSOAs located within other districts in the North West region of England, providing 1,490 matched control LSOAs—that is, 1,788 LSOAs in total. In order to identify a matched control group that was likely to satisfy the parallel trends assumption, we used propensity score matching (Rosenbaum & Rubin, 1983) to identify control areas that had experienced similar trends in emergency admissions and predictors of this outcome (age, unemployment, proportion of the population that were female) in the time period before the introduction of the intervention (2005-2010; see online supplementary file, Appendix 2, for full details of the matching variables). These three variables were included in the matching and as time varying controls as they include the main factors that predict health care utilization (i.e., age, sex,

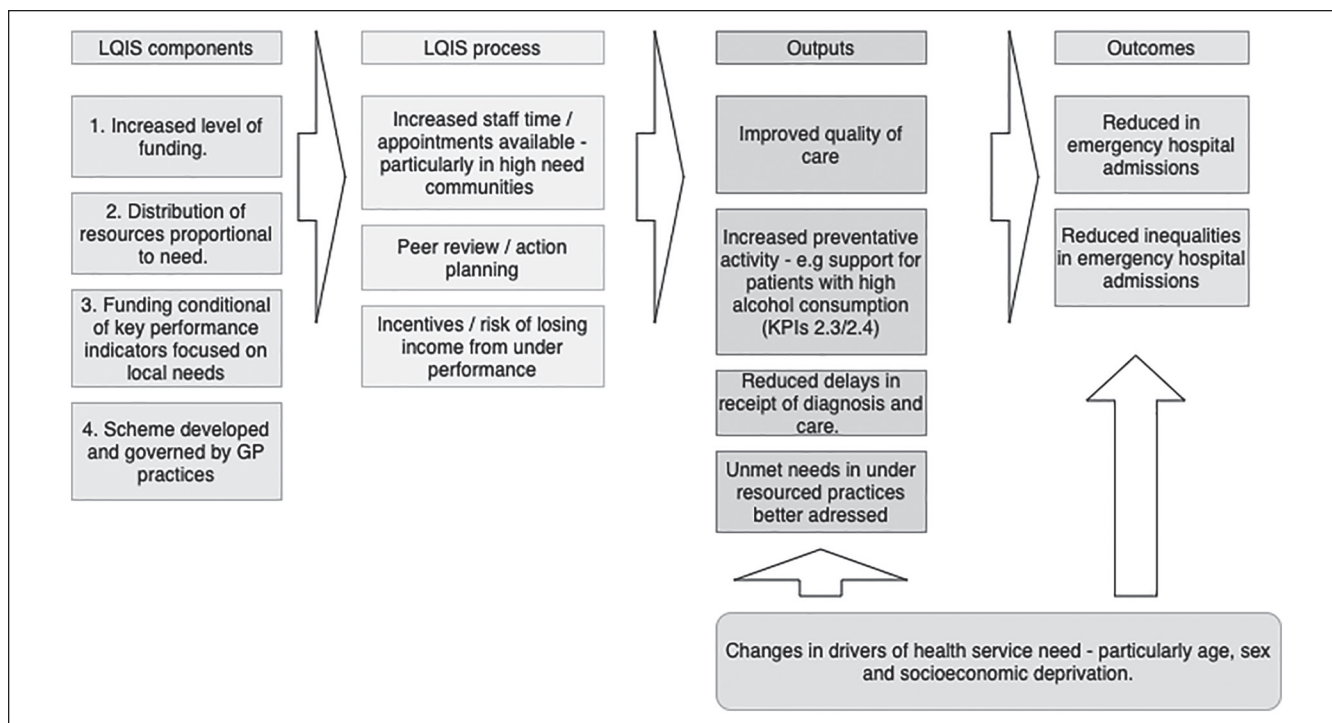


Figure 1. Conceptual framework showing input components and hypothesized process and outputs leading to the outcomes.
 Note. LQIS = Local Quality Improvement Scheme; GP = general practitioner.

and socioeconomic deprivation), and they were available annually for LSOAs (see The Intervention section). Other morbidity measures were considered (e.g., prevalence of chronic conditions), however, these would have introduced bias as the diagnosed prevalence of these conditions could have been affected by the intervention itself through increased access to diagnosis.

We then used a linear regression model to compare the change (difference) in the emergency admission rate in the intervention population with the change (difference) in the emergency admission rate in the matched comparison population, 6 years before (2005-2010) and 6 years after (2011-2016) implementation. We calculated emergency hospital admissions per 1,000 population for each of the 1,788 LSOAs between 2005 and 2016, giving a total sample size of 21,456 LSOA-years. To adjust for time varying factors that could be associated with trends in emergency admission rates, we controlled for trends in the annual average age of the population, the percent female, and the unemployment rate measured as the proportion of the working age population (aged 16-64 years) claiming unemployment benefits. We investigated the parallel trends assumption using graphical methods and regression models to compare trends in emergency admission rates between the intervention and control populations in the preintervention period (2005-2010). To investigate whether there was a difference in effect in more compared with less deprived areas, we conducted subgroup analysis across groups defined by their

level of income deprivation. Parallel trends were investigated for each subgroup.

To test the sensitivity of the analysis to the control group chosen, we repeated the analysis using GP practice—years as the unit of analysis and using controls selected from outside the North West region of England. Additionally as the process of matching can introduce bias related to regression to the mean (Daw & Hatfield, 2018), we conducted supplementary analysis using the synthetic control method for microdata (Robbins & Davenport, 2021). This uses a weighted combination of all of the available “untreated” LSOAs, instead of just a matched sample (see online supplementary file, Appendix 6).

All intervention practices were included in this analysis even if their funding was withdrawn, due to under performance (see below). This intention to treat analysis will provide a more conservative estimate of the intervention effect and is less prone to bias.

The Intervention

The Local Quality Improvement Scheme (LQIS) has four main components as outlined in the conceptual framework in Figure 1.

First, there was an increase in the level of funding. An additional £5 million a year, £30 million in total was invested in the GP practices between 2011 and 2016 through this scheme. Second, these funds were distributed proportional to

the level of need. Prior to the intervention, idiosyncrasies in national contracting arrangements and differences in the uptake of optional enhanced services had led to considerable variation in the funding each practice received ranging from £51 to £207 per population, weighted for need using the Carr-Hill Formula. The Carr-Hill formula is used in the United Kingdom to adjust populations for their relative need for primary care—taking into account the age and sex of the population and the standardized prevalence of limited long-standing illness and the standardized mortality ratio for patients younger than 65 years (British Medical Association, 2020). The LQIS equalized these payments ensuring that each practice received a minimum of £90 per needs weighted population. Third, receipt of this additional funding was conditional on achieving a set of 13 key performance indicators (KPIs) that were monitored on an annual basis (see online supplementary file, Appendix 3, for a list of KPIs). These indicators were developed in collaboration with local GPs and public health specialists. They were designed to address specific health needs relevant to the local population. In particular, this involved increasing primary care access and early identification of chronic conditions, support for people with high alcohol intake, improved medicines management, and increased uptake of childhood vaccinations. These are all areas that are not well covered through the national incentive scheme (QOF). In addition, there was a specific target to reduce unplanned emergency admissions, which was also not included in the national scheme. Finally, a key component was the involvement of the GPs in developing the scheme, ensuring that they felt the measures were relevant and were collectively committed to the scheme's success.

As highlighted in Figure 1, these components increased the level of access to GP practices from an average of 50 appointments to 70 appointments per 1,000 needs-weighted population per day. Where GP practices were not achieving the performance indicator targets, they were required to provide evidence outlining actions they were taking to achieve the target. This evidence was reviewed by a committee that included GPs, lay members of the public, contract managers, and representatives from National Health Service (NHS) England. Incentives were in place to achieve the targets, or risk losing income. To date funds have been withdrawn from 37 practices for nonachievement of KPIs over several validation processes. For example, 56 practices had initial submissions that had not met the KPI standards in 2016. Further investigation showed that 35 practices were not able to provide evidence that they met all the standards and this resulted in a £350,000 fund withdrawal from these practices (from nearly £9m invested that year altogether in the scheme). It is therefore hypothesized that LQIS could have had an impact on the level and inequalities of emergency admissions through improvements in the quality of care, increased preventative action (e.g., brief interventions for people with high alcohol intake), early identification of problems, and through addressing unmet needs in previously under resourced practices.

Data Sources

We used Hospital Episode Statistics provided by NHS Digital along with ONS population estimates to derive our primary outcome: all-cause, all-age emergency hospital admissions per 1,000 population, for all LSOAs from 2005 to 2016. Emergency admissions are defined by NHS digital as those admitted at short notice due to clinical need, generally through an accident and emergency department or through direct request from a GP. Although the KPI used in LQIS was emergency admissions from Ambulatory Care Sensitive Conditions (ACSCs) that are usually managed in primary care, we used the all-cause emergency admission rate as our outcome because previous research suggests that incentive schemes may have an adverse effect on conditions not covered by performance indicators (Damberg et al., 2014; Doran et al. 2008). All-cause emergency admissions would therefore better capture the overall program effect. In sensitivity analysis, we additionally analyzed the impact on emergency admissions for ACSCs. Data on unemployment and demographics was obtained from the ONS. There were no missing data in our study.

Patient and Public Involvement

This study was part of Partners Priority Program in the National Institute for Health Research Collaboration for Leadership in Applied Health Research and Care North West Coast. The Partners Priority Program aimed to create a strong link between academia and local health care organizations and programs so that a research priority was defined by partners, according to their needs. A coalition of academic researchers, partner organization staff, and public advisors undertook the research. Public advisors were engaged in a qualitative investigation of LQIS that informed the design of this study, in the planning of the work and in the dissemination of the findings.

Results

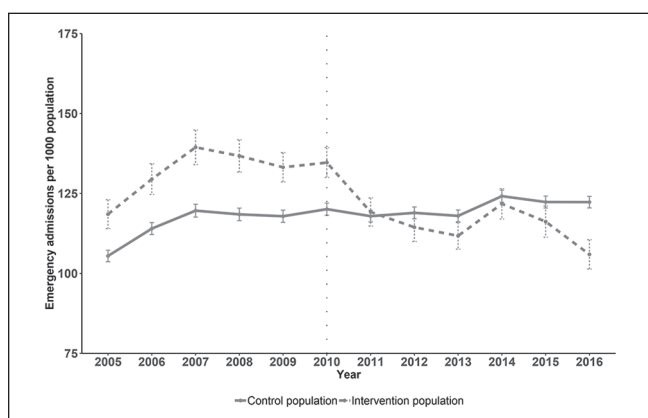
The sociodemographic characteristics and emergency admission rates for the intervention populations and the matched control populations before the intervention are summarized in Table 1. Whilst there was some remaining imbalance between the intervention and control populations on the matching variables, these variables were included as covariates in the difference-in-differences model, and any fixed differences between the populations are accounted for through the difference-in-differences analysis.

Figure 2 shows the trend in emergency admission rates within the intervention and control populations before and after intervention. Whilst the admission rate was higher in the intervention compared with the control populations, prior to the intervention the trends appeared parallel (as suggested by regression analysis in the online supplementary file, Appendix 4). After 2010, when the LQIS was introduced, the

Table 1. Sociodemographic Features of Intervention Population Compared With Unmatched and Matched Populations, in the Time Period Before the Introduction of the Intervention (2005-2010).

Variable	Unmatched sample (2005-2010)			Matched sample (2005-2010)		
	Intervention population, M (SD)	Unmatched population, M (SD)	Standardized mean difference	Intervention population M (SD)	Matched control population, M (SD)	Standardized mean difference
Average age in years	37.53 (4.39)	39.52 (4.75)	0.435	37.53 (4.39)	38.26 (4.94)	0.155
Working age population unemployed (%)	5.63 (2.89)	2.92 (2.34)	1.032	5.63 (2.89)	3.94 (2.79)	0.595
Population female (%)	50.87 (3.55)	51.06 (2.29)	0.065	50.87 (3.55)	51.18 (2.44)	0.104
All-cause emergency hospital admission rate (per 1,000)	131.52 (43.03)	106.59 (35.37)	0.633	131.52 (43.03)	116.37 (38.36)	0.372
Number of LSOAs	298	4,199	—	298	1,490	—

Note. LSOA = Lower Layer Super Output Area. M = mean. SD = standard deviation.

**Figure 2.** Trends in all-cause emergency hospital admission rates per 1,000 in intervention and control populations, before and after the introduction of the intervention.

emergency admission rate fell in the intervention populations to a similar level as the control populations.

The difference-in-difference analysis (Table 2) shows that there was a statistically significant reduction in admissions in the intervention populations compared with control populations following the introduction of the intervention when controlling for other potential confounders. The analysis indicates that the LQIS was associated with a reduction of 19 emergency admissions per 1,000 people (95% confidence interval [CI: 17, 21] per year, compared with control populations. This is a 14% decline from the baseline level (see Table 1), and approximately equivalent to 10,670 fewer emergency admissions across the overall 550,000 intervention population annually after the intervention, and 64,020 over the 6 years of the intervention period (2011-2016) included in this study. Based on the reference costs for short stay emergency admissions (Curtis & Burns, 2016) used by the NHS, this reduction in admissions is the equivalent to a £39 million cost saving to the NHS over these 6 years.

For each deprivation subgroup, regression analysis suggested parallel trends in the outcome between intervention

and control populations before the introduction of the intervention (see online supplementary file, Appendix 5, for details). We found that the effect of the intervention was greater in more deprived areas, with the intervention associated with a decline of 22 emergency admissions per 1,000 population (95% CI [18, 25]) per year in the most deprived third of LSOAs in the study city, compared with a decline of 17 emergency admissions per 1,000 population (95% CI [15, 20]) per year in the least deprived third of LSOAs.

We found similar results when repeating the analysis using GP practices as the units of analysis (decline of 18 emergency admissions per 1,000 population, 95% CI [14, 22] per year), using controls selected from outside the North West region of England (decline of 15 emergency admissions per 1,000 population, 95% CI [13, 17] per year) and using chronic ACSCs as the outcome of interest (decline of 0.8 emergency admissions per 1,000 population, 95% CI [0.7, 0.99] per year; see online supplementary file, Appendix 6).

Discussion

Summary

We found that a local General Practice quality improvement scheme that improved the level and equitable distribution of investment alongside a performance incentive scheme was associated with a reduction in emergency hospital admissions. The scheme had a greater impact in more deprived areas compared with more affluent areas, reducing inequalities. This reduction of admissions represented an approximate £39 million cost saving to the NHS. Given that the intervention was estimated to cost £30 million, this suggests that the intervention was cost saving to the NHS.

Strengths and Limitations

Our study's strengths include evaluating the LQIS in its real-life implementation setting, which makes our findings potentially more externally valid than those implemented in trials.

Table 2. Results of DiD Models Showing Changes in All-Cause Emergency Hospital Admissions per 1,000 People in Intervention Populations Following the Introduction of the Intervention, Compared With the Control Populations, 2005-2016.

Variable	Model 1	Model 2	Model 3	Model 4
	Full model, Coefficient [95% CI]	Subgroup: Low deprivation, Coefficient [95% CI]	Subgroup: Middle deprivation, Coefficient [95% CI]	Subgroup: High deprivation, Coefficient [95% CI]
Annual time trend term	1.88 [1.73, 2.04]	1.65 [1.45, 1.85]	1.66 [1.41, 1.90]	1.63 [1.32, 1.95]
Average age in years	1.86 [1.55, 2.17]	2.57 [2.34, 2.81]	2.77 [2.36, 3.17]	3.44 [2.91, 3.97]
Working age population unemployed (%)	1.36 [1.17, 1.55]	1.73 [1.21, 2.25]	1 [0.64, 1.37]	0.89 [0.62, 1.16]
Population female (%)	0.48 [0.10, 0.86]	1.03 [0.64, 1.42]	0.45 [-0.06, 0.95]	-0.36 [-0.99, 0.26]
Group (intervention = 1, control = 0)	14.35 [9.92, 18.79]	14.14 [9.76, 18.52]	28.96 [24.06, 33.85]	15.49 [10.05, 20.93]
Period (postintervention = 1, preintervention = 0)	-7.14 [-8.08, -6.20]	-4.74 [-6.00, -3.48]	-5.2 [-6.72, -3.67]	-10.05 [-11.99, -8.11]
DiD estimator (Group * Period)	-19.4 [-21.34, -17.47]	-17.2 [-19.67, -14.73]	-19.4 [-22.68, -16.11]	-21.6 [-25.34, -17.87]

Note. Models include random intercept for LSOA. CIs calculated using cluster robust estimation. Model 1 based on equation shown in the online supplementary file and based on 298 intervention and 1,490 control LSOAs, 21,456 LSOA-years in total. Model 2 based on 100 intervention and 500 control LSOAs, 7,200 LSOA-years in total. Model 3 based on 99 intervention and 495 control LSOAs, 7,128 LSOA-years in total. Model 4 based on 99 intervention and 495 control LSOAs, 7,128 LSOA-years in total. CI = confidence interval; DiD = difference-in-differences; LSOA = Lower Layer Super Output Area.

Our data enabled a follow-up period of 6 years. This allowed us to investigate whether effects were sustained. Third, a combination of quasi-experimental methods (propensity score matching and difference-in-differences) was applied, which can lead to causal estimates of LQIS if the trends in outcomes would have been parallel in the absence of this intervention. Fourth, our reasonably large effective sample size of 21,456 observations provided sufficient power to identify relatively small effects.

Some limitations however are of importance. It is difficult to rule out the possibility that different trends in unobserved confounding factors between the two groups could have influenced the results. Although there were differences between the intervention and control groups, time-invariant differences between the two groups could not bias the results due to the difference-in-differences methods (Angrist & Pischke, 2009). Propensity score matching identified control populations that followed similar trends in the outcome over time prior to the intervention, confirmed by the parallel nature of the trend in emergency admission rates before the intervention. In addition, we controlled for a number of observed confounders. Unobserved confounders therefore could only lead to bias in the results if they followed different time trends over time between the intervention and control groups. We were only able to assess the impact of the intervention on emergency hospital admissions and this may not represent health benefits to the users of these services. Data of other outcomes, such as mortality, was not available at the geographical level required for this analysis and therefore could not be included. The ecological nature of this study limits the conclusions that can be drawn about individual-level factors, and the results represent the population-level impact of the LQIS.

Comparison With Existing Literature

There are number of reasons that may potentially explain why our findings suggest that LQIS had a positive effect (illustrated in our conceptual framework in Figure 1) whilst the evidence for other similar schemes has been more mixed (Harrison et al., 2014; Ryan et al., 2016; Roland & Guthrie, 2016). First, the LQIS included a significant increase in overall funding for primary care, rather than just reallocating a portion of current funding into a performance incentive scheme, this allowed for increased capacity and access and as well as changes in practice (Doran et al., 2014). Second, the LQIS led to a more equitable distribution of resources between practices relative to need. This could have led to greater benefits by addressing unmet needs in those practices that had previously been under resourced (Doran et al., 2017). Third, the performance metrics were developed to specifically address identified local needs and focused on outcome as well as process measures. Concerns have been raised that although process measures are more easily achievable, they may not lead to the practice change that is needed to improve outcomes (Doran et al., 2017). Fourth, the scheme was developed by GPs and they were involved in its governance, this could have led to better design of performance indicators and greater sense of ownership among the practices, leading to greater changes in practice (Roland & Dudley, 2015; Begum et al., 2013).

Implications for Practice

Our study has important implications for practice. It indicates that local primary care quality improvement schemes can be effective at reducing demand on secondary care. The study

indicates that to be effective they should include increased equitable investment in primary care linked to key outcome based performance measures and involve GPs in their design and governance. Future research should investigate whether these findings are replicated in a trial context investigating overall health benefits and potential adverse effects in areas of care not covered by performance indicators.

Conclusions

In contrast to other evidence on national incentive schemes, we found that a local scheme developed in collaboration with GPs was effective at reducing rates and inequalities in emergency admissions. Similar approaches could be an effective component of strategies to reduce unplanned hospital admissions elsewhere.

Acknowledgments

The authors wish to thank Clinical Commissioning Group for proposing and supporting this study.

Authors' Contributions

EKM and TR are joint first authors. LB, CM, SP, MGo, EKM, TR, and BB conceptualized the study. EKM conducted the analysis and wrote the drafts. BB commented on the drafts, supervised the analysis, performed supplementary analyses, rewrote the final draft, and revised the manuscript following peer review. TR commented on the drafts, checked the analysis results, performed the supplementary analyses, and revised the manuscript following peer review. MGa, MGo, and RB read the drafts and provided their comments throughout.

Declaration of Conflicting Interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: SP, LB, and CM are employed by Liverpool Clinical Commissioning Group (CCG), the commissioner of LQIS. RB is the secretary of the Liverpool Local Medical Committee (LLMC), an elected body that represents all the GPs in Liverpool and was involved in negotiating the contract for LQIS. He is also a GP in one of the practices that provided the LQIS. MGa is an honorary consultant at NHS Liverpool CCG, member of the LLMC, and academic associate GP in one of the GP practices that provided LQIS. SP, LB, CM, RB, and MGa, were all involved in providing information about the nature of the intervention, its contracting, budget and performance management, as well as providing contextual information for the interpretation of the results. They had no role in the analysis or presentation of the results. The views expressed in this article are those of the authors and do not represent the views of Liverpool CCG, LLMC, or any of the participating GP practices. There are no other relationships or activities that could appear to have influenced the submitted work.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This

report is independent research funded by the National Institute for Health Research Applied Research Collaboration North West Coast (ARC NWC). The views expressed in this publication are those of the author(s) and not necessarily those of the National Institute for Health Research or the Department of Health and Social Care.

Ethics Approval and Consent to Participate

No ethical approval was required for this study, as it involved the use of anonymous aggregate secondary health service data and openly available data.

ORCID iD

Tanith C. Rose  <https://orcid.org/0000-0001-5338-0359>

Availability of data and materials

The Hospital Episode Statistics data that support the findings of this study are available from NHS Digital but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. All other data used in the study are publicly available, detailed in the online supplementary file.

Supplemental Material

Supplemental material for this article is available online.

References

- Alshamsan, R., Lee, J. T., Majeed, A., Netuveli, G., & Millett, C. (2012). Effect of a UK pay-for-performance program on ethnic disparities in diabetes outcomes: Interrupted time series analysis. *Annals of Family Medicine, 10*(3), 228-234. <https://doi.org/10.1370/afm.1335>
- Alshamsan, R., Majeed, A., Ashworth, M., Car, J., & Millett, C. (2010). Impact of pay for performance on inequalities in health care: systematic review. *Journal of Health Services Research & Policy, 15*(3), 178-184. <https://doi.org/10.1258/jhsrp.2010.009113>
- Angrist, J. D., & Pischke, J. S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press. <https://doi.org/10.1515/9781400829828>
- Ashworth, M., & Gulliford, M. (2017). Funding for general practice in the next decade: Life after QOF. *British Journal of General Practice, 67*(654), 4-5. <https://doi.org/10.3399/bjgp17X688477>
- Bardach, N. S., Wang, J. J., De Leon, S. F., Shih, S. C., Boscardin, W. J., Goldman, L. E., & Dudley, R. A. (2013). Effect of pay-for-performance incentives on quality of care in small practices with electronic health records: A randomized trial. *JAMA, 310*(10), 1051-1059. <https://doi.org/10.1001/jama.2013.277353>
- Begum, R., Ryan, M. S., Winther, C. H., Wang, J. J., Bardach, N. S., Parsons, A. H., Shi, S. C., & Dudley, R. A. (2013). Small practices' experience with EHR, quality measurement, and incentives. *American Journal of Managed Care, 19*(10), eSP12-8. <https://pubmed.ncbi.nlm.nih.gov/24511883/>
- British Medical Association. (2020). *Global sum allocation formula*. <https://www.bma.org.uk/advice-and-support/gp-practices/funding-and-contracts/global-sum-allocation-formula>

- Brosig-Koch, J., Hennig-Schmidt, H., Kairies-Schwarz, N., & Wiesen, D. (2016). Using artefactual field and lab experiments to investigate how fee-for-service and capitation affect medical service provision. *Journal of Economic Behavior & Organization*, 131(Pt B), 17-23. <https://doi.org/10.1016/j.jebo.2015.04.011>
- Campbell, S. M., Reeves, D., Kontopantelis, E., Sibbald, B., & Roland, M. (2009). Effects of pay for performance on the quality of primary care in England. *New England Journal of Medicine*, 361(4), 368-378. <https://doi.org/10.1056/NEJMsa0807651>
- Conrad, D. A., & Perry, L. (2009). Quality-based financial incentives in health care: Can we improve quality by paying for it? *Annual Review of Public Health*, 30, 357-371. <https://doi.org/10.1146/annurev.publhealth.031308.100243>
- Curtis, L., & Burns, A. (2016). *Unit costs of health and social care 2016 UK: Personal social services research unit, Cornwallis Building*, The University of Kent.
- Damberg, C. L., Sorbero, M. E., Lovejoy, S. L., Martsolf, G. R., Raen, L., & Mandel, D. (2014). *Measuring success in health care value-based purchasing programs: Summary and recommendations*. RAND Corporation
- Daw, J., & Hatfield, L. (2018). Matching and regression to the mean in difference-in-differences analysis. *Health Services Research*, 53(6), 4138-4156. <https://doi.org/10.1111/1475-6773.12993>
- Department for Communities and Local Government. (2015). *The English indices of deprivation 2015: Statistical release*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/465791/English_Indices_of_Deprivation_2015_-_Statistical_Release.pdf
- Dimick, J. B., & Ryan, A. M. (2014). Methods for evaluating changes in health care policy: The difference-in-differences approach. *JAMA*, 312(22), 2401-2402. <https://doi.org/10.1001/jama.2014.16153>
- Dixon, A., & Khachatryan, A. (2010). A review of the public health impact of the Quality and Outcomes Framework. *Quality in Primary Care*, 18(2), 133-138. <https://primarycare.imedpub.com/a-review-of-the-public-health-impact-of-the-quality-and-outcomes-framework.pdf>
- Doran, T., Fullwood, C., Gravelle, H., Reeves, D., Kontopantelis, E., Hiroeh, U., & Roland, M. (2006). Pay-for-performance programs in family practices in the United Kingdom. *New England Journal of Medicine*, 355(4), 375-384. <https://doi.org/10.1056/NEJMsa055505>
- Doran, T., Fullwood, C., Kontopantelis, E., & Reeves, D. (2008). Effect of financial incentives on inequalities in the delivery of primary clinical care in England: Analysis of clinical activity indicators for the quality and outcomes framework. *The Lancet*, 372(9640), 728-736. [https://doi.org/10.1016/S0140-6736\(08\)61123-X](https://doi.org/10.1016/S0140-6736(08)61123-X)
- Doran, T., Kontopantelis, E., Reeves, D., Sutton, M., & Ryan, A. M. (2014). Setting performance targets in pay for performance programmes: What can we learn from QOF? *British Medical Journal*, 348, Article g1595. <https://doi.org/10.1136/bmj.g1595>
- Doran, T., Kontopantelis, E., Valderas, J. M., Campbell, S., Roland, M., Salisbury, C., & Reeves, D. (2011). Effect of financial incentives on incentivised and non-incentivised clinical activities: Longitudinal analysis of data from the UK Quality and Outcomes Framework. *British Medical Journal*, 342, Article d3590. <https://doi.org/10.1136/bmj.d3590>
- Doran, T., Maurer, K. A., & Ryan, A. M. (2017). Impact of provider incentives on quality and value of health care. *Annual Review of Public Health*, 38, 449-465. <https://doi.org/10.1146/annurev-publhealth-032315-021457>
- Eijkenaar, F., Emmert, M., Scheppach, M., & Schoffski, O. (2013). Effects of pay for performance in health care: A systematic review of systematic reviews. *Health Policy*, 110(2-3), 115-130. <https://doi.org/10.1016/j.healthpol.2013.01.008>
- Emmert, M., Eijkenaar, F., Kemter, H., Esslinger, A. S., & Schöffski, O. (2012). Economic evaluation of pay-for-performance in health care: A systematic review. *European Journal of Health Economics*, 13(6), 755-767. <https://doi.org/10.1007/s10198-011-0329-8>
- Fleetcroft, R., Parekh-Bhurke, S., Howe, A., Cookson, R., Swift, L., & Steel, N. (2010). The UK pay-for-performance programme in primary care: Estimation of population mortality reduction. *British Journal of General Practice*, 60(578), e345-e352. <https://doi.org/10.3399/bjgp10X515359>
- Flodgren, G., Eccles, M. P., Shepperd, S., Scott, A., Parmelli, E., & Beyer, F. R. (2011). An overview of reviews evaluating the effectiveness of financial incentives in changing health-care professional behaviours and patient outcomes. *Cochrane Database of Systematic Reviews*, 7, Article 9255. <https://doi.org/10.1002/14651858.CD009255>
- Forbes, L. J., Marchand, C., Doran, T., & Peckham, S. (2017). The role of the Quality and Outcomes Framework in the care of long-term conditions: A systematic review. *British Journal of General Practice*, 67(664), e775-e784. <https://doi.org/10.3399/bjgp17X693077>
- Gillam, S. J., Siriwardena, A. N., & Steel, N. (2012). Pay-for-performance in the United Kingdom: Impact of the Quality and Outcomes Framework—A systematic review. *Annals of Family Medicine*, 10(5), 461-468. <https://doi.org/10.1370/afm.1377>
- Girault, A., Minvielle, E., Loirat, P., Jiang, S., Lalloué, B., & Ferrua, M. (2017). Evaluation of the effects of the French pay-for-performance program—IFAQ pilot study. *International Journal for Quality in Health Care*, 29(6), 833-837. <https://doi.org/10.1093/intqhc/mzx111>
- Glidewell, L., West, R., Hackett, J. E., Carder, P., Doran, T., & Foy, R. (2015). Does a local financial incentive scheme reduce inequalities in the delivery of clinical care in a socially deprived community? A longitudinal data analysis. *BMC Family Practice*, 16, Article 61. <https://doi.org/10.1186/s12875-015-0279-9>
- Green, E. P. (2014). Payment systems in the healthcare industry: An experimental study of physician incentives. *Journal of Economic Behavior & Organization*, 106(October), 367-378. <https://doi.org/10.1016/j.jebo.2014.05.009>
- Grigoroglou, C., Munford, L., Webb, R. T., Kapur, N., Doran, T., Ashcroft, D. M., & Kontopantelis, E. (2018). Association between a national primary care pay-for-performance scheme and suicide rates in England: Spatial cohort study. *British Journal of Psychiatry*, 213(4), 600-608. <https://doi.org/10.1192/bjp.2018.143>
- Grimaldi, A. (2018). Faut-il financer la qualité des soins et adopter le « pay for performance » (P4P) des anglo-saxons? [Should we

- finance the quality of care and adopt the pay for performance (P4P) of Anglo-Saxon countries?] *Médecine des Maladies Métaboliques*, 12(3), 284-290. [https://doi.org/10.1016/S1957-2557\(18\)30060-9](https://doi.org/10.1016/S1957-2557(18)30060-9)
- Guthrie, B., McLean, G., & Sutton, M. (2006). Workload and reward in the Quality and Outcomes Framework of the 2004 general practice contract. *British Journal of General Practice*, 56(532), 836-841. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1927091/>
- Hackett, J., Glidewell, L., West, R., Carder, P., Doran, T., & Foy, R. (2014). Just another incentive scheme: A qualitative interview study of a local pay-for-performance scheme for primary care. *BMC Family Practice*, 15(1), Article 168. <https://doi.org/10.1186/s12875-014-0168-7>
- Hamel, M. B., Roland, M., & Campbell, S. (2014). Successes and failures of pay for performance in the United Kingdom. *New England Journal of Medicine*, 370(20), Article 1944. <https://doi.org/10.1056/NEJMhpr1316051>
- Harrison, M. J., Dusheiko, M., Sutton, M., Gravelle, H., Doran, T., & Roland, M. (2014). Effect of a national primary care pay for performance scheme on emergency hospital admissions for ambulatory care sensitive conditions: Controlled longitudinal study. *British Medical Journal*, 349, Article g6423. <https://doi.org/10.1136/bmj.g6423>
- Hennig-Schmidt, H., Seltner, R., & Wiesen, D. (2011). How payment systems affect physicians' provision behaviour—An experimental investigation. *Journal of Health Economics*, 30(4), 637-646. <https://doi.org/10.1016/j.jhealeco.2011.05.001>
- Herbst, T., Foerster, J., & Emmert, M. (2018). The impact of pay-for-performance on the quality of care in ophthalmology: Empirical evidence from Germany. *Health Policy*, 122(6), 667-673. <https://doi.org/10.1016/j.healthpol.2018.03.013>
- Houle, S. K., McAlister, F. A., Jackevicius, C. A., Chuck, A. W., & Tsuyuki, R. T. (2012). Does performance-based remuneration for individual health care practitioners affect patient care? A systematic review. *Annals of Internal Medicine*, 157(12), 889-899. <https://doi.org/10.7326/0003-4819-157-12-201212180-00009>
- Kondo, K. K., Damberg, C. L., Mendelson, A., Motu'apuaka, M., Freeman, M., O'Neil, M., Relevo, R., Low, A., & Kansagara, D. (2016). Implementation processes and pay for performance in healthcare: A systematic review. *Journal of General Internal Medicine*, 31(Suppl. 1), 61-69. <https://doi.org/10.1007/s11606-015-3567-0>
- Kontopantelis, E., Springate, D. A., Ashworth, M., Webb, R. T., Buchan, I. E., & Doran, T. (2015). Investigating the relationship between quality of primary care and premature mortality in England: A spatial whole-population study. *British Medical Journal*, 350, Article h904. <https://doi.org/10.1136/bmj.h904>
- Lee, J. T., Netuveli, G., Majeed, A., & Millett, C. (2011). The effects of pay for performance on disparities in stroke, hypertension, and coronary heart disease management: Interrupted time series study. *PLOS ONE*, 6(12), Article e27236. <https://doi.org/10.1371/journal.pone.0027236>
- Lester, H., Schmittiel, J., Selby, J., Fireman, B., Campbell, S., Lee, J., Whippy, A., & Madvig, P. (2010). The impact of removing financial incentives from clinical quality indicators: Longitudinal analysis of four Kaiser Permanente indicators. *British Medical Journal*, 340, Article c1898. <https://doi.org/10.1136/bmj.c1898>
- Mandavia, R., Mehta, N., Schilder, A., & Mossialos, E. (2017). Effectiveness of UK provider financial incentives on quality of care: A systematic review. *British Journal of General Practice*, 67(664), e800-e815. <https://doi.org/10.3399/bjgp17X693149>
- Markovitz, A. A., & Ryan, A. M. (2017). Pay-for-performance: Disappointing results or masked heterogeneity? *Medical Care Research Review*, 74(1), 3-78. <https://doi.org/10.1177/1077558715619282>
- Marshall, M., & Roland, M. (2017). The future of the Quality and Outcomes Framework in England. *British Medical Journal*, 359, Article j4681. <https://doi.org/10.1136/bmj.j4681>
- Mendelson, A., Kondo, K., Damberg, C., Low, A., Motuapuaka, M., Freeman, M., O'Neil, M., Relevo, R., & Kansagara, D. (2017). The effects of pay-for-performance programs on health, health care use, and processes of care: A systematic review. *Annals of Internal Medicine*, 166(5), 341-353. <https://doi.org/10.7326/M16-1881>
- Minchin, M., Roland, M., Richardson, J., Rowark, S., & Guthrie, B. (2018). Quality of care in the United Kingdom after removal of financial incentives. *New England Journal of Medicine*, 379(10), 948-957. <https://doi.org/10.1056/NEJMsa1801495>
- Pandya, A., Doran, T., Zhu, J., Walker, S., Arntson, E., & Ryan, A. M. (2018). Modelling the cost-effectiveness of pay-for-performance in primary care in the UK. *BMC Medicine*, 16(1), Article 135. <https://doi.org/10.1186/s12916-018-1126-3>
- Petersen, L. A., Simpson, K., Pietz, K., Urech, T. H., Hysong, S. J., Profit, J., Conrad, D. A., Adams Dudley, R. A., & Woodard, L. D. (2013). Effects of individual physician-level and practice-level financial incentives on hypertension care: A randomized trial. *JAMA*, 310(10), 1042-1050. <https://doi.org/10.1001/jama.2013.276303>
- Petersen, L. A., Woodard, L. D., Urech, T., Daw, C., & Sookanan, S. (2006). Does pay-for-performance improve the quality of health care? *Annals of Internal Medicine*, 145(4), 265-272. <https://doi.org/10.7326/0003-4819-145-4-200608150-00006>
- Raleigh, V. S. (2018). Quality and Outcomes Framework at the fork in the road. *British Medical Journal*, 362, Article k3645. <https://doi.org/10.1136/bmj.k3645>
- Robbins, M., & Davenport, S. (2021). Microsynth: Synthetic Control Methods for Disaggregated and Micro-Level Data in R. *Journal of Statistical Software*, 97(2), 1-31. <https://doi.org/10.18637/jss.v097.i02>
- Roland, M. (2016). Does pay-for-performance in primary care save lives? *Lancet*, 388(10041):217-218. [https://doi.org/10.1016/S0140-6736\(16\)00550-X](https://doi.org/10.1016/S0140-6736(16)00550-X)
- Roland, M., & Dudley, R. A. (2015). How financial and reputational incentives can be used to improve medical care. *Health Services Research*, 50(S2), 2090-2115. <https://doi.org/10.1111/1475-6773.12419>
- Roland, M., & Guthrie, B. (2016). Quality and Outcomes Framework: What have we learnt? *British Medical Journal*, 354, Article i4060. <https://doi.org/10.1136/bmj.i4060>
- Roland, M., & Olesen, F. (2016). Can pay for performance improve the quality of primary care? *British Medical Journal*, 354, Article i4058. <https://doi.org/10.1136/bmj.i4058>

- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41-55. <https://doi.org/10.1093/biomet/70.1.41>
- Ryan, A. M., Krinsky, S., Kontopantelis, E., & Doran, T. (2016). Long-term evidence for the effect of pay-for-performance in primary care on mortality in the UK: A population study. *Lancet*, 388(10041), 268-274. [https://doi.org/10.1016/S0140-6736\(16\)00276-2](https://doi.org/10.1016/S0140-6736(16)00276-2)
- Scott, A., Sivey, P., Ait Ouakrim, D., Willenberg, L., Naccarella, L., Furler, J., & Young, D. (2011). The effect of financial incentives on the quality of health care provided by primary care physicians. *Cochrane Database Systematic Reviews*, 9, Article Cd008451. <https://doi.org/10.1002/14651858.CD008451.pub2>
- Soucat, A., Dale, E., Mathauer, I., & Kutzin, J. (2017). Pay-for-performance debate: Not seeing the forest for the trees. *Health Systems & Reform*, 3(2), 74-79. <https://doi.org/10.1080/23288604.2017.1302902>
- Steel, N., & Shekelle, P. (2016). After 12 years, where next for QOF? *British Medical Journal*, 354, Article i4103. <https://doi.org/10.1136/bmj.i4103>
- Thorne, T. (2016). How could the Quality and Outcomes Framework (QOF) do more to tackle health inequalities? *London Journal of Primary Care*, 8(5), 80-84. <https://doi.org/10.1080/17571472.2016.1215370>
- Van Herck, P., De Smedt, D., Annemans, L., Remmen, R., Rosenthal, M. B., & Sermeus, W. (2010). Systematic review: Effects, design choices, and context of pay-for-performance in health care. *BMC Health Services Research*, 10(1), Article 247. <https://doi.org/10.1186/1472-6963-10-247>
- Walker, S., Mason, A. R., Claxton, K., Cookson, R., Fenwick, E., Fleetcroft, R., & Sculpher, M. (2010). Value for money and the Quality and Outcomes Framework in primary care in the UK NHS. *British Journal of General Practice*, 60(574), e213-e220. <https://doi.org/10.3399/bjgp10X501859>
- Witter, S., Fretheim, A., Kessy, F. L., & Lindahl, A. K. (2012). Paying for performance to improve the delivery of health interventions in low- and middle-income countries. *Cochrane Database Systematic Reviews*, 2, Article Cd007899. <https://doi.org/10.1002/14651858.CD007899.pub2>
- Xi, X., Wang, E., Lu, Q., Chen, P., Wo, T., & Tang, K. (2019). Does an economic incentive affect provider behavior? Evidence from a field experiment on different payment mechanisms. *Journal of Medical Economics*, 22(1), 35-44. <https://doi.org/10.1080/13696998.2018.1539399>
- Yuan, B., He, L., Meng, Q., & Jia, L. (2017). Payment methods for outpatient care facilities. *Cochrane Database Systematic Review*, 3, Article Cd011153. <https://doi.org/10.1002/14651858.CD011153.pub2>