Supplementary File

Appendix 1 – Matching variables

The following variables were included in a propensity score model to match Knowsley to control areas in the time period before the introduction of the intervention (2005–10).

Matching variable	Details
Age and gender profile of the population	Annual data on the size of the female population and the population aged 50+ years per lower super output area (LSOA) were derived from mid-year population estimates provided by the Office for National Statistics (ONS).
Unemployment rate	Annual unemployment rates were calculated using claimant data provided by the ONS. Unemployment was measured as the proportion of people aged 16–64 years claiming Jobseeker's Allowance or Universal Credit principally for the reason of being unemployed.
Chronic obstructive pulmonary disease (COPD) emergency admission rate	Emergency admissions for COPD were defined using ICD-10 codes: J40–J44. Annual COPD emergency admission rates per 100,000 population were calculated using Hospital Episode Statistics (HES), with population data obtained from the ONS. Continuous inpatient (CIP) spells were used to calculate emergency admissions per calendar year.
Indices of Multiple Deprivation	Index of Multiple Deprivation 2015 data were provided by the Department for Communities and Local Government.
Quality and Outcomes Framework (QOF) indicators	QOF indicator data for the prevalence of COPD and smoking, and the percentage of patients with COPD receiving inhaled treatment whose inhaler technique had been checked within the previous 15 months were included in the propensity score model. Weighted averages of QOF indicators per LSOA were calculated using data provided by NHS Digital on the number of patients registered per general practice per LSOA.
Numbers of general practitioners (GPs) per capita serving the population	Weighted averages of the number of full-time employed GPs per 1000 population were calculated using data provided by NHS Digital on the number of GPs and patients registered per general practice per LSOA.
Distance to the nearest general practice and hospital	The Consumer Data Research Centre provided data per LSOA on the average road network distance to the nearest hospital with an Accident and Emergency (A&E) department, and the nearest general practice. Road network distances in kilometres were calculated by deriving the fastest route by car to travel from each postcode within an LSOA to the nearest health service.

Appendix 2 – Outline of the Difference-in-Differences analysis used in the study

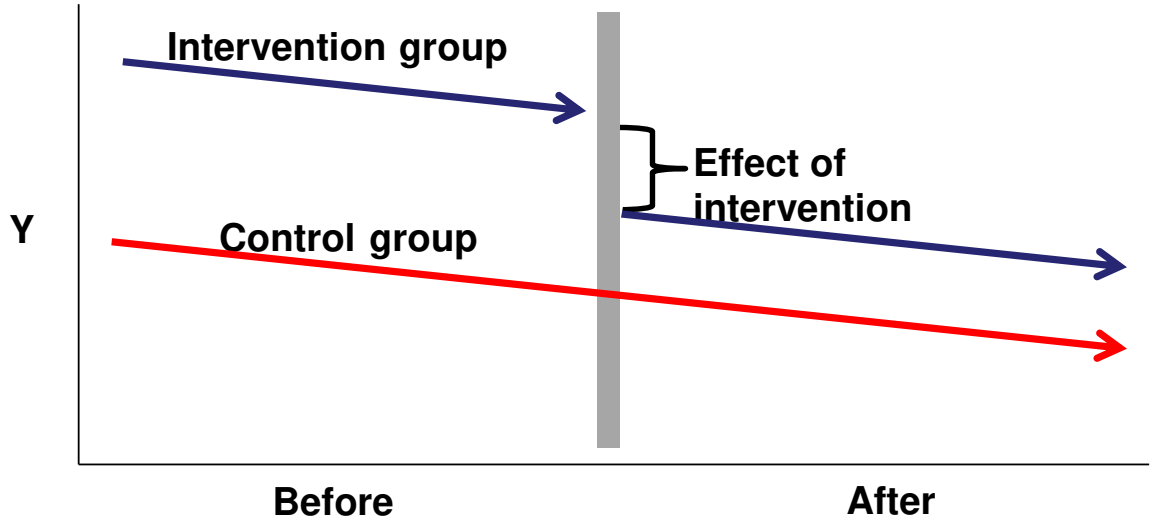
Difference-in-differences (DiD) analyses are an established approach used in econometrics[1,2] and increasingly in health research[3,4] for evaluating the impact of interventions, where the researcher has not manipulated the assignment of the intervention, sometimes known as "natural experiments".

In DiD analyses, outcomes are observed for two groups before and after an intervention and one of the groups is exposed to a treatment in the second period but not in the first period, and the second group is not exposed to the treatment during the full time frame. The average change in outcomes in the second (control) group is then subtracted from the average change in outcomes the first (treatment) group. This removes biases in second period comparisons between the treatment and control group that could be the result from permanent differences between those groups, as well as biases from comparisons over time in the treatment group that could be the result of trends.

Thus the differences-in-differences estimator is therefore:

$$\hat{\gamma} = \left(\bar{Y}_{Treatment,AFTER} - \bar{Y}_{Treatment,BEFORE}\right) - \left(\bar{Y}_{COMPARATOR,AFTER} - \bar{Y}_{COMPARATOR,BEFORE}\right)$$

$\bar{Y}$ is the mean of the outcome variable in the intervention areas after the start of the intervention ($\bar{Y}_{Treatment,AFTER}$) and before the start of the intervention ($\bar{Y}_{Treatment,BEFORE}$) and in the comparator areas after the start of the intervention ($\bar{Y}_{COMPARATOR,AFTER}$) and before the start of the intervention ($\bar{Y}_{COMPARATOR,BEFORE}$). A representation of this is given in the figure below:



To estimate the DiD estimator we can run the following regression:

$$Y_{at} = \beta_1 Treatment_a + \beta_2 AFTER_t + \beta_3 AFTER_t * Treatment_a + \varepsilon_{at}$$

where Y_{at} is the outcome variable in area a at time t, $Treatment_a$ is a dummy variable taking the value 1 for the intervention area and the value 0 for the comparator areas and $AFTER_t$ is a dummy variable taking the value 1 for time periods after the start of the intervention and 0 before. The coefficient of interest is β_3, the coefficient on the interaction term $AFTER_t * BL_a$, this is the differences-in-differences parameter.

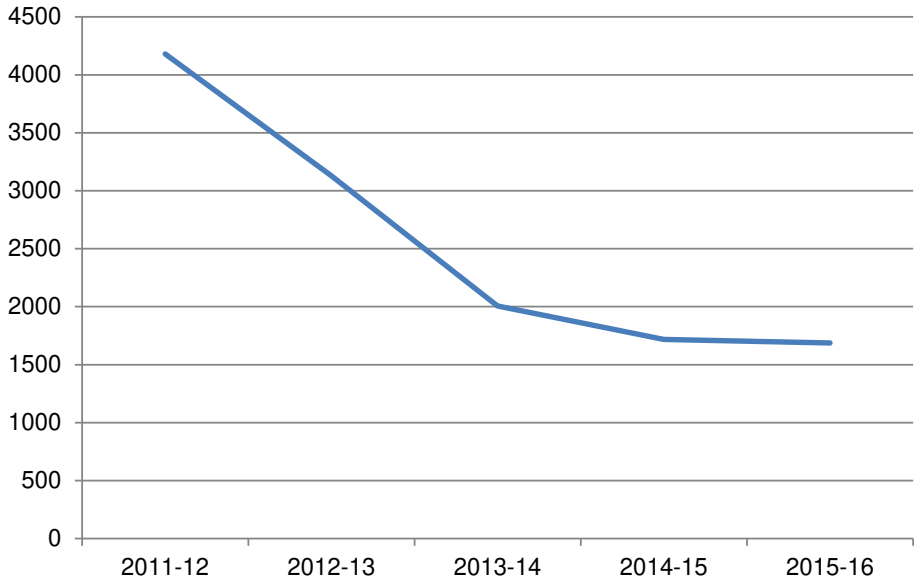
Whilst this analysis cannot be biased due to time invariant differences between the intervention and control groups, it could be biased by trends in other predictors of the outcome if these diverged between treatment and intervention groups after the intervention. We therefore additionally include potential time varying confounders in this model, giving the following model used in the analysis.

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 between the two groups provides an unbiased estimate of the interventions effect.[5]

We use a linear regression model, as the interaction term can only be interpreted as the difference-in-differences estimator in a linear probability model, and linear models provide a robust estimate of the difference-in-difference estimator even when the data are not normally distributed.[6]

Appendix 3 – Referrals to the intervention.

Total number of general practice referrals to the Knowsley COPD intervention per year, 2011–15



Appendix 4 - Power calculation

Simulation was used to estimate the power of the difference-in-differences analysis, based on the model formula given above. Power, i.e. the probability of rejecting a false null hypothesis, was estimated at specific effect sizes by running the model using multiple simulated datasets each with a specific relative effect size plus a random error drawn from a normal distribution in each iteration. Power was then estimated as the proportion of the models which falsely rejected the null hypothesis. For each specified effect size, 100 simulated datasets were tested. The results are shown below and indicate that the effect size that the study had 80% power to detect, lay somewhere between 10-11% decline in emergency admission rates associated with the intervention. In absolute terms, this is the equivalent of a reduction of 52 to 57 emergency COPD admissions per 100,000 population per year.

Effect size	Proportion of models which rejected null hypothesis %
0.83	100
0.84	100
0.85	98
0.86	96
0.87	90
0.88	91
0.89	83
0.90	74
0.91	62
0.92	62
0.93	45
0.94	31
0.95	22
0.96	18
0.97	16
0.98	8
0.99	6

Appendix 5. Analysing the differential effects of the intervention by deprivation and by gender

Low income deprivation

Result of difference-in-differences analysis showing the change in COPD emergency admissions per 100,000 population in Knowsley following the intervention relative to the control group, for areas with low income deprivation, 2005–16

	Coefficient	Lower 95% CI	Upper 95% CI	p-value
Spline 1	-10.98	-18.30	-3.67	0.003
Spline 2	21.31	10.04	32.59	<0.001
Population aged 50+ years (%)	6.05	4.19	7.90	<0.001
Population female (%)	1.21	-7.07	9.49	0.775
Working age population unemployed (%)†	155.62	39.47	271.76	0.009
Treatment (Knowsley = 1; control = 0)	-15.78	-65.11	33.54	0.528
Period (post-intervention = 1; pre-intervention = 0)	-25.70	-57.65	6.25	0.115
DiD estimator (treatment*period)	**29.99**	**-9.88**	**69.86**	**0.140**

† Variable entered into model in units of 10% points
Model includes random intercept for LSOA
Model based on 29 Knowsley and 135 control LSOAs, and 1968 observations
CI = confidence interval; COPD = Chronic Obstructive Pulmonary Disease; DiD = Difference-in-Differences; LSOA = Lower-layer Super Output Area

Medium income deprivation

Result of difference-in-differences analysis showing the change in COPD emergency admissions per 100,000 population in Knowsley following the intervention relative to the control group, for areas with medium income deprivation, 2005–16

	Coefficient	Lower 95% CI	Upper 95% CI	p-value
Spline 1	-30.84	-42.82	-18.87	<0.001
Spline 2	41.82	22.43	61.22	<0.001
Population aged 50+ years (%)	7.02	2.58	11.47	0.002
Population female (%)	-6.20	-17.80	5.40	0.294
Working age population unemployed (%)†	40.26	-70.45	150.96	0.476
Treatment (Knowsley = 1; control = 0)	18.88	-60.48	98.25	0.639
Period (post-intervention = 1; pre-intervention = 0)	19.78	-32.27	71.83	0.456
DiD estimator (treatment*period)	**-64.33**	**-126.91**	**-1.76**	**0.044**

† Variable entered into model in units of 10% points
Model includes random intercept for LSOA
Model based on 32 Knowsley and 132 control LSOAs, and 1968 observations
CI = confidence interval; COPD = Chronic Obstructive Pulmonary Disease; DiD = Difference-in-Differences; LSOA = Lower-layer Super Output Area

High income deprivation

Result of difference-in-differences analysis showing the change in COPD emergency admissions per 100,000 population in Knowsley following the intervention relative to the control group, for areas with high income deprivation, 2005–16

	Coefficient	Lower 95% CI	Upper 95% CI	p-value
Spline 1	-38.28	-51.98	-24.58	<0.001
Spline 2	51.77	27.65	75.89	<0.001
Population aged 50+ years (%)	26.84	21.57	32.12	<0.001
Population female (%)	2.49	-9.11	14.08	0.674
Working age population unemployed (%)†	23.76	-73.59	121.12	0.632
Treatment (Knowsley = 1; control = 0)	43.22	-36.52	122.95	0.286
Period (post-intervention = 1; pre-intervention = 0)	-50.36	-112.14	11.41	0.110
DiD estimator (treatment*period)	**-49.57**	**-119.48**	**20.33**	**0.164**

† Variable entered into model in units of 10% points
Model includes random intercept for LSOA
Model based on 37 Knowsley and 125 control LSOAs, and 1944 observations
CI = confidence interval; COPD = Chronic Obstructive Pulmonary Disease; DiD = Difference-in-Differences; LSOA = Lower-layer Super Output Area

Women.

Result of difference-in-differences analysis showing the change in COPD emergency admissions per 100,000 women in Knowsley following the intervention relative to the control group, 2005–16

	Coefficient	Lower 95% CI	Upper 95% CI	p-value
Spline 1	-39.91	-48.86	-30.96	<0.001
Spline 2	83.09	69.33	96.85	<0.001
Population aged 50+ years (%)	5.95	3.05	8.85	<0.001
Population female (%)	5.81	-3.25	14.87	0.209
Working age population unemployed (%)†	405.69	340.14	471.23	<0.001
Treatment (Knowsley = 1; control = 0)	45.48	-17.81	108.77	0.159
Period (post-intervention = 1; pre-intervention = 0)	-16.43	-57.40	24.54	0.432
DiD estimator (treatment*period)	**6.09**	**-42.67**	**54.84**	**0.807**

† Variable entered into model in units of 10% points
Model includes random intercept for LSOA
Model based on 98 Knowsley and 392 control LSOAs, and 5880 observations
CI = confidence interval; COPD = Chronic Obstructive Pulmonary Disease; DiD = Difference-in-Differences; LSOA = Lower-layer Super Output Area

Men.

Result of difference-in-differences analysis showing the change in COPD emergency admissions per 100,000 men in Knowsley following the intervention relative to the control group, 2005–16

	Coefficient	Lower 95% CI	Upper 95% CI	p-value
Spline 1	-37.47	-46.09	-28.84	<0.001
Spline 2	62.42	49.26	75.58	<0.001
Population aged 50+ years (%)	6.22	3.65	8.80	<0.001
Population female (%)	18.44	10.10	26.78	<0.001
Working age population unemployed (%)†	290.86	230.85	350.88	<0.001
Treatment (Knowsley = 1; control = 0)	18.50	-37.07	74.07	0.513
Period (post-intervention = 1; pre-intervention = 0)	-22.72	-62.63	17.19	0.264
DiD estimator (treatment*period)	**-59.80**	**-107.29**	**-12.32**	**0.014**

† Variable entered into model in units of 10% points
Model includes random intercept for LSOA
Model based on 98 Knowsley and 392 control LSOAs, and 5880 observations
CI = confidence interval; COPD = Chronic Obstructive Pulmonary Disease; DiD = Difference-in-Differences; LSOA = Lower-layer Super Output Area

Appendix 6. Robustness tests

Checking difference in trends in the pre-intervention period.

Analysis showing the change in COPD emergency admissions per 100,000 population per year in Knowsley relative to the control group in the pre-intervention period only (2005–10)

	Coefficient	Lower 95% CI	Upper 95% CI	p-value
Population aged 50+ years (%)	8.37	5.39	11.34	<0.001
Population female (%)	23.29	13.74	32.85	<0.001
Working age population unemployed (%)†	406.46	340.60	472.32	<0.001
Treatment (Knowsley = 1; control = 0)	12110.99	-16484.08	40706.06	0.406
Year	-41.12	-48.54	-33.70	<0.001
Treatment*year	-6.02	-20.24	8.20	0.406

† Variable entered into model in units of 10% points
Model includes random intercept for LSOA
Model based on 98 Knowsley and 392 control LSOAs, and 2940 observations
CI = confidence interval; COPD = Chronic Obstructive Pulmonary Disease; LSOA = Lower-layer Super Output Area

Checking effect when running the analysis using an outcome (emergency admissions for GI infections) that would not plausibly be influenced.

Result of difference-in-differences analysis showing the change in gastrointestinal infection emergency admissions per 100,000 population in Knowsley following the intervention relative to the control group, 2005–15

	Coefficient	Lower 95% CI	Upper 95% CI	p-value
Spline 1	14.416	10.80	18.03	<0.001
Spline 2	-10.349	-16.76	-3.94	0.002
Population aged 50+ years (%)	-1.164	-2.02	-0.31	0.008
Population female (%)	6.476	3.94	9.02	<0.001
Working age population unemployed (%)†	127.688	106.28	149.10	<0.001
Treatment (Knowsley = 1; control = 0)	22.536	2.02	43.05	0.031
Period (post-intervention = 1; pre-intervention = 0)	-97.097	-115.99	-78.21	<0.001
DiD estimator (treatment*period)	**-17.438**	**-40.40**	**5.53**	**0.137**

† Variable entered into model in units of 10% points
Model includes random intercept for LSOA
Model based on 98 Knowsley and 490 control LSOAs, and 6468 observations
CI = confidence interval; DiD = Difference-in-Differences; LSOA = Lower-layer Super Output Area

Checking effect when removing a spline term allowing for a change in trend across both groups after the intervention

Result of difference-in-differences analysis showing the change in COPD emergency admissions per 100,000 population in Knowsley following the intervention relative to the control group, 2005–16

	Coefficient	Lower 95% CI	Upper 95% CI	p-value
Spline 1	-4.55	-8.84	-0.27	0.037
Population aged 50+ years (%)	4.76	2.23	7.29	<0.001
Population female (%)	9.95	2.74	17.16	0.007
Working age population unemployed (%)†	31.77	-10.22	73.76	0.138
Treatment (Knowsley = 1; control = 0)	47.99	-11.78	107.76	0.115
Period (post-intervention = 1; pre-intervention = 0)	-36.77	-65.74	-7.81	0.013
DiD estimator (treatment*period)	-25.95	-60.62	8.72	0.142

† Variable entered into model in units of 10% points
Model includes random intercept for LSOA
Model based on 98 Knowsley and 392 control LSOAs, and 5880 observations
CI = confidence interval; COPD = Chronic Obstructive Pulmonary Disease; DiD = Difference-in-Differences; LSOA = Lower-layer Super Output Area

References

1. Angrist JD, Pischke J-S. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton: Princeton University Press 2009.
2. Wooldridge JM. Econometric Analysis of Cross Section and Panel Data. 2nd Revised ed. Cambridge: Mass: MIT Press 2010.
3. Craig P, Dieppe P, Macintyre S, et al. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008;337:a1655. doi: 10.1136/bmj.a1655 [published Online First: 2008/10/01]
4. Craig P, Katikireddi SV, Leyland A, et al. Natural Experiments: An Overview of Methods, Approaches, and Contributions to Public Health Intervention Research. *Annu Rev Public Health* 2017;38:39-56. doi: 10.1146/annurev-publhealth-031816-044327 [published Online First: 2017/01/27]
5. Dimick JB, Ryan AM. Methods for evaluating changes in health care policy: the difference-in-differences approach. *JAMA* 2014;312(22):2401-2. doi: 10.1001/jama.2014.16153 [published Online First: 2014/12/10]
6. Puhani PA. The treatment effect, the cross difference, and the interaction term in nonlinear "difference-in-differences" models. *Economics Letters* 2012;115(1):85-87. doi: 10.1016/j.econlet.2011.11.025

STROBE Statement—Checklist of items that should be included in reports of ***case-control studies***

	Item No	Recommendation	Page
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	3
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	6
Objectives	3	State specific objectives, including any prespecified hypotheses	6
Methods			
Study design	4	Present key elements of study design early in the paper	7-8
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	7
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls	7-8
		(*b*) For matched studies, give matching criteria and the number of controls per case	8-9
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	6-7
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	9-10
Bias	9	Describe any efforts to address potential sources of bias	12-13
Study size	10	Explain how the study size was arrived at	11-12
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	6-7
Statistical methods	12	(*a*) Describe all statistical methods, including those used to	7-8

		control for confounding	
		(*b*) Describe any methods used to examine subgroups and interactions	8-9
		(*c*) Explain how missing data were addressed	12-13
		(*d*) If applicable, explain how matching of cases and controls was addressed	7-8
		(*e*) Describe any sensitivity analyses	12-13
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	7-8
		(b) Give reasons for non-participation at each stage	NA
		(c) Consider use of a flow diagram	NA
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	Table 1
		(b) Indicate number of participants with missing data for each variable of interest	No missing data
Outcome data	15*	Report numbers in each exposure category, or summary measures of exposure	Table 1, Appendix 7
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	Table 2
		(*b*) Report category boundaries when continuous variables were categorized	Table 2
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	NA
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses Tables 3-7, Appendices 5-7	
Discussion			
Key results	18	Summarise key results with reference to study objectives 15	

Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias 16-17
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence 17-19
Generalisability	21	Discuss the generalisability (external validity) of the study results 16-17
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based 20

*Give information separately for cases and controls.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.