

Quantifying digital health inequality across a national healthcare system

SUPPLEMENTARY MATERIALS

Joe Zhang^{1,2,*}, Jack Gallifant³, Robin L Pierce⁴, Aoife Fordham⁵, James T Teo^{6,7}, Leo A Celi^{8,9,10}, Hutan Ashrafian¹

*Corresponding author: jzhang@nhs.net / joe.zhang@imperial.ac.uk

1. Institute of Global Health Innovation, Imperial College London, London, UK
2. Department of Critical Care, Guy's and St. Thomas' Hospital, London, UK
3. Department of Critical Care, Imperial College Healthcare NHS Trust, London, UK
4. University of Exeter Law School, University of Exeter, Exeter, UK
5. Transformation Directorate, NHS England, UK
6. Department of Neurology, King's College Hospital NHS Foundation Trust, London, UK
7. London Medical Imaging & AI Centre, Guy's and St. Thomas' Hospital, London, UK
8. Institute for Medical Engineering and Science, Massachusetts Institute of Technology, Cambridge, USA
9. Division of Pulmonary, Critical Care and Sleep Medicine, Beth Israel Deaconess Medical Center, Boston, USA
10. Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, USA

Results

Full results are reported in **Tables & Figures** below. Table A shows results for multivariable and univariate analyses across both interventions. Table B summarises the primary care provider population.

Methods

Study setting and data sources

We performed a retrospective observational study across populations of primary care practices in NHS England, providing general health services to 55 million patients. Data were obtained from NHS Digital and NHS Improvement who host administrative datasets and systems metadata. Small area ethnicity and deprivation data were from the Office of National Statistics. Our study took place from November 2021 to October 2022 inclusive, a period reflecting relative stability in COVID-19 activity[1]. At the start of the study period, the observed digital health interventions were nationally available.

Inclusion, exclusion, and missing data

An inclusion flow chart is shown in *Figure A*. We included all active primary care centres in England. This designation includes general practices, but also other specialist services that are cost centres for prescribing. We therefore excluded registered companies for remote, app-based services, and specialist clinics for drug, alcohol, and homeless support. We excluded practices that closed during the study period. A small number of practices did not submit administrative data in the study period (0.7%) or exhibited missing outcomes metadata (<0.5%) and were excluded.

Outcome variables

Outcome variables were: (1) percentage of practice population who had activated the NHS App by October 2022; (2) percentage of practice population who had activated online services by October 2022. These actions reflect active engagement on the part of patients and are therefore taken as indicators for digital health uptake.

Deprivation and ethnicity

We used geographic Indices of Multiple Deprivation (IMD), a standardised, aggregate index of income, employment, education, health, crime, housing, and local environment[2]. Patients are assigned to IMD quintiles based on their residence (with the first quintile representing patients from the 20% most deprived areas). We took per-practice percentage of patients from each IMD quintile.

Gold standard ethnicity data consists of small area data from the 2021 national census. We assumed that proportion of ethnicities in small area subpopulations remains consistent across registrations from each small area to each primary care practice. We took per-practice percentage of Black, Asian, and White patients as covariates.

Other covariates

Per-practice covariates included the following: percentage of patients within age categories, percentage of patients on chronic disease registers, average number of unique prescription items administered per patient, full time equivalent general practitioners and non-clinical staff per patient, and percentage of patients from nationally defined urban or rural residence. We used a dummy variable to represent university-based practices where all registered patients were students. For analysis of online services usage, we also included the software supplier as a covariate. Full descriptions are found in *Table C*.

Statistical analyses

Univariate and multivariable linear regression models were constructed. Covariates were removed from analysis if Variance Inflation Factors exceeded 6, indicating troublesome multicollinearity. Linearity, homoscedasticity, and normality of residuals were assessed visually (*Figure B*). Coefficients for the lowest two deprivation quintiles were used to estimate an associated population effect for each practice.

Patient and Public involvement

While no patients were directly involved in this study, our research relates to the national CORE20PLUS5 programme focusing on patient inequalities related to the first IMD quintile[3].

Limitations

We note several limitations. Results reflect population activity and do not directly inform inference for individual patients, or their propensity to adopt digital technology. We are unable to individually adjust or generate inference for unmeasured patient-level factors such as digital literacy, or device availability, or other social determinants of digital access. This is a trade-off of a population-level analysis. Additionally, we are only able to measure activation of digital services, due to data quality issues for “click-through” metadata. Ethnicity co-variables are derived from the 2021 census, as practice/patient-level ethnicity are unavailable for analysis (see Methods). Finally, our analysis does not directly account for per-practice initiatives to promote digital health uptake in registered populations.

Ethics

As an observational analysis of non-personal metadata and aggregate population datasets, this study did not qualify as research as research under Health Research Authority definitions. Data was taken from centralised NHS data collections, and, as such, is deemed representative of the UK population. Population analysis of metadata generally does not raise questions of privacy or informed consent. However, given the topic of this study, it is important to scrutinize the methods to examine whether and how structural factors may have impacted various aspects of the data set and findings. We found no cause for concern in this regard in this study.

Data & Code Availability

Links to all analysed datasets are found in *Table C*. Analysis performed using the *statsmodels* package in Python 3.8. Geo-visualisations created using ArcGIS 3.0.2.

References

- 1 Office of National Statistics. Coronavirus (COVID-19) latest insights dashboard. 2022. <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/coronaviruscovid19latestinsights/overview>
- 2 Ministry of Housing, Communities, & Local Government. English Indices of Deprivation 2019: technical report. 2019.
- 3 NHS Improvement. Core20Plus5: An approach to reducing health inequalities. 2021. <https://www.england.nhs.uk/about/equality/equality-hub/national-healthcare-inequalities-improvement-programme/core20plus5/>

Tables & Figures

	Univariate				Multivariable				
NHS App	coeff	pval	ci_lower	ci_upper	coeff	pval	ci_lower	ci_upper	VIF
const					0.504	0.000	0.484	0.524	
IMD Q1	-0.248	0.000	-0.255	-0.242	-0.223	0.000	-0.232	-0.213	3.029
IMD Q2	-0.181	0.000	-0.194	-0.168	-0.117	0.000	-0.128	-0.106	1.846
IMD Q5	0.281	0.000	0.273	0.290	0.121	0.000	0.111	0.131	2.315
Age 16 to 25	0.004	0.872	-0.046	0.054	0.293	0.000	0.254	0.331	1.673
Age 26 to 35	-0.310	0.000	-0.359	-0.261	0.241	0.000	0.191	0.291	2.882
Age 56 to 65	0.685	0.000	0.606	0.764	0.585	0.000	0.485	0.685	4.355
Age 76 to 85	1.027	0.000	0.937	1.117	-0.177	0.011	-0.312	-0.041	5.939
Admin staff per 1000pt	-0.010	0.001	-0.017	-0.004	0.010	0.000	0.005	0.014	1.334
GPs per 1000pt	0.039	0.000	0.030	0.048	0.025	0.000	0.019	0.030	1.105
Black ethnicity	-0.312	0.000	-0.355	-0.269	0.016	0.388	-0.021	0.054	2.077
Asian ethnicity	-0.172	0.000	-0.193	-0.150	0.016	0.095	-0.003	0.034	2.014
Urban residence	-0.048	0.000	-0.056	-0.041	0.043	0.000	0.037	0.049	1.754
Student practice	0.242	0.000	0.178	0.305	0.105	0.000	0.063	0.147	1.231
COPD	-3.044	0.000	-3.273	-2.815	0.535	0.000	0.289	0.782	2.959
Heart failure	0.015	0.904	-0.226	0.255	0.394	0.000	0.199	0.589	1.862
Dementia	2.167	0.000	1.550	2.785	-0.751	0.002	-1.233	-0.269	1.717
Mental health disorder	-1.890	0.000	-2.000	-1.780	-0.123	0.008	-0.214	-0.032	1.643
Any long-term condition	-0.157	0.000	-0.172	-0.142	-0.091	0.000	-0.110	-0.071	4.599
Prescription items per pt	-0.037	0.000	-0.042	-0.033	0.009	0.000	0.004	0.014	3.475
Primary Care Online	coeff	pval	ci_lower	ci_upper	coeff	pval	ci_lower	ci_upper	VIF
const					3.799	0.000	3.381	4.217	
IMD Q1	-2.338	0.000	-2.459	-2.216	-2.047	0.000	-2.247	-1.847	3.034
IMD Q2	-1.625	0.000	-1.823	-1.427	-1.114	0.000	-1.348	-0.880	1.846
IMD Q5	2.567	0.000	2.416	2.719	1.269	0.000	1.055	1.482	2.321
Age 16 to 25	-1.842	0.000	-2.585	-1.098	0.184	0.661	-0.638	1.005	1.674
Age 26 to 35	-2.225	0.000	-2.954	-1.496	5.819	0.000	4.762	6.876	2.882
Age 56 to 65	8.491	0.000	7.317	9.666	8.872	0.000	6.752	10.992	4.351
Age 76 to 85	10.832	0.000	9.476	12.189	-0.889	0.540	-3.731	1.952	4.816
Admin staff per 1000pt	-0.141	0.000	-0.213	-0.070	-0.059	0.157	-0.141	0.023	1.796
GPs per 1000pt	0.537	0.000	0.405	0.670	0.360	0.000	0.241	0.479	1.098
Black ethnicity	-3.017	0.000	-3.657	-2.377	0.034	0.934	-0.765	0.833	2.116
Asian ethnicity	-1.469	0.000	-1.792	-1.145	0.973	0.000	0.579	1.367	2.013
Urban residence	-0.538	0.000	-0.647	-0.430	0.298	0.000	0.174	0.421	1.752
Student practice	2.555	0.000	1.611	3.499	2.882	0.000	1.988	3.776	1.231
COPD	-23.802	0.000	-27.334	-20.270	7.377	0.006	2.131	12.624	2.957
Heart failure	7.513	0.000	3.961	11.065	10.801	0.000	6.670	14.933	1.858
Dementia	6.165	0.028	0.680	11.649	-1.488	0.676	-8.469	5.492	2.229
Mental health disorder	-14.125	0.000	-15.766	-12.485	1.387	0.161	-0.554	3.328	1.844
Any long-term condition	-1.111	0.000	-1.333	-0.888	-1.161	0.000	-1.580	-0.742	4.823
Prescription items per pt	-0.237	0.000	-0.303	-0.171	0.173	0.002	0.065	0.281	3.648
EMIS vendor	-0.209	0.000	-0.282	-0.136	-0.188	0.000	-0.252	-0.124	1.064

Table A – results of univariate and multivariable linear regression analyses for association of primary care practice population characteristics with digital intervention uptake. IMD Q1 = index of multiple deprivation, quintile 1; pt = patient; GP = general practitioner; COPD = chronic obstructive pulmonary disease; EMIS = EMIS Health, Leeds, UK. Covariates excluded due to multicollinearity include: age brackets 36 to 45, 46 to 55, 66 to 75, over 85; IMD brackets Q3 and Q4; white ethnicity; ischaemic heart disease, diabetes, and hypertension. Full description of covariates in *Appendix Table B*.

Primary care provider (n=6356) covariates	Summary statistics
Practice list size	8308 [5427-12019]
IMD Q1 % population	11.8 [0.6-39.3]
IMD Q2 % population	17.7 [7.4-30.5]
IMD Q3 % population	16.7 [8.1-27.7]
IMD Q4 % population	15.6 [5.2-27.8]
IMD Q5 % population	6.9 [0.8-26.9]
Age 16 to 25 % population	10.6 [9.6-11.9]
Age 26 to 35 % population	13.7 [11.5-16.1]
Age 36 to 45 % population	13.5 [11.9-15.7]
Age 46 to 55 % population	13.4 [12.6-14.2]
Age 56 to 65 % population	12.7 [10.7-14.3]
Age 76 to 85 % population	9.4 [6.8-11.7]
Age 86+ % population	5.5 [3.7-7.3]
Admin staff per 1000pt	1.9 [1.2-2.5]
GPs per 1000pt	1.2 [0.9-1.4]
Black ethnicity % population	0.5 [0.4-0.7]
Asian ethnicity % population	2.3 [0.7-7.3]
White ethnicity % population	6 [2.1-15]
Percentage in Urban residence	85.6 [67.8-94.7]
University-based practices	9 (0.14%)
Percentage with COPD	1.9 [1.3-2.6]
Percentage with heart failure	1.5 [1-2.2]
Percentage with dementia	0.6 [0.4-0.8]
Percentage with IHD	5.9 [4.5-7.3]
Percentage with mental health disorder	4 [3-5.4]
Percentage with hypertension	13.7 [11.3-15.9]
Prescription items per patient	156.9 [121.6-193.6]
Medicines cost per patient	1247.1 [1018.8-1446.2]
Providers using EMIS EHR	3815 (60%)
Providers using TPP EHR	2492 (39.2%)
Providers using Vision EHR	21 (0.33%)
Percentage online services activation	54.9 [45.3-63.3]
Percentage NHS App activation	60.6 [53.9-67.3]

Table B – Summary characteristics of primary care provider analysis population. Continuous variables represented with median and interquartile range (median [IQR]). Categorical variables represented with count and percentage of total population (n (%)). “List size” refers the official registered list of patients per provider. IMD = Index of Multiple Deprivation; GP = General Practitioner; COPD = Chronic Obstructive Pulmonary Disease; IHD = Ischaemic Heart Disease; EMIS = Egton Medical Information Systems; TPP = The Phoenix Partnership; EHR = Electronic Health Record; NHS = National Health Service

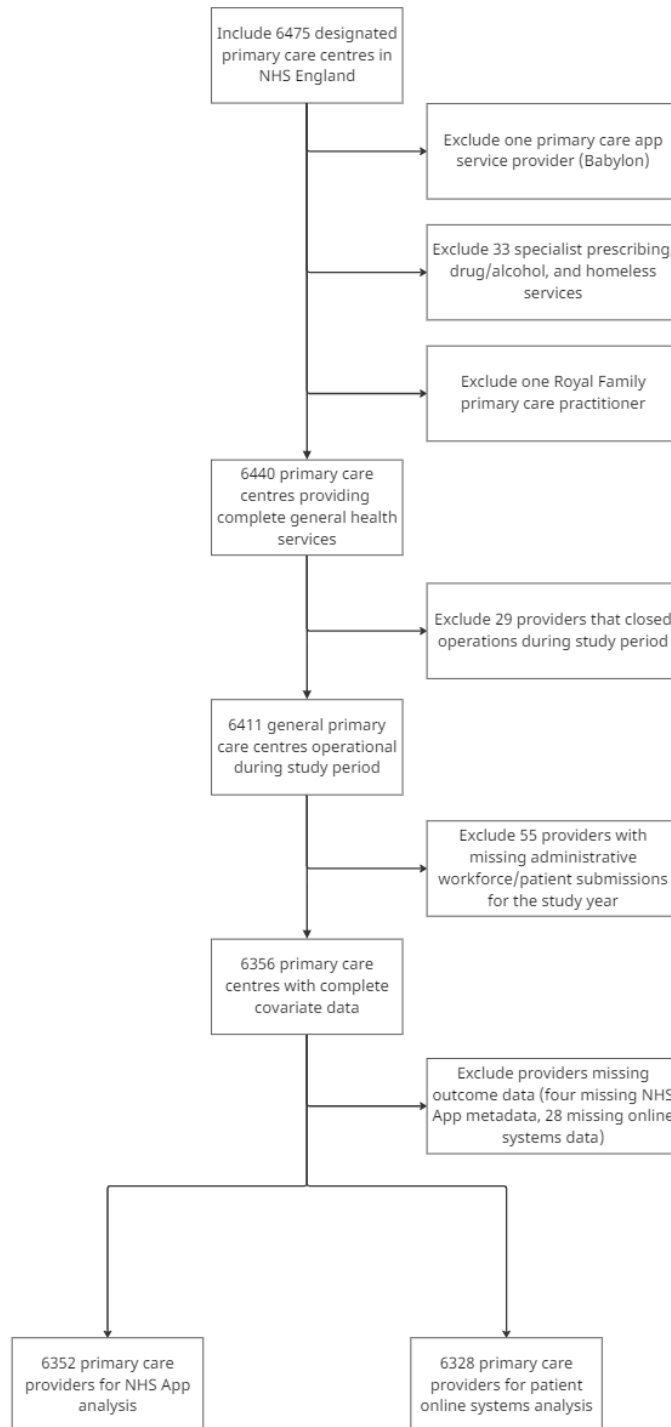


Figure A – Population inclusion flow chart

VARIABLES	DATA SOURCE
Covariates	
Primary care providers and list size	Quarterly and historical publications for active primary care practices as prescribing cost centres from NHS Digital, and the NHS Digital Organisation Data Service [Data 1 / Data 2]
Primary care provider list sizes	Monthly and historical list sizes for primary care practices from NHS England information services [Data files]
Deprivation of registered population	NHS Digital “patients registered at a GP practice”, by deprivation [Data files]
Primary care practice population age	NHS Digital “patients registered at a GP practice”, by age in years [Data files]
Primary care staffing data	NHS Digital general practice workforce datasets [Data files]
Gold-standard small area geographic ethnicity data	Results of 2021 national census, small area ethnicity dataset [Data files]
Population chronic disease prevalence	Chronic disease lists for Quality and Outcomes Framework prevalence and exceptions [Data files]
Practice-level prescribing summary data	Extract of practice level data for study period from the OpenSAFELY OpenPrescribing portal [Openprescribing].
Outcomes	
NHS App activations	Practice-level metadata for NHS App activations and usage. Data accessible through application to NHS England. [Data files]
Patient online services activations	NHS Digital Patient Online Management Information systems metadata [Data files]

Table C – Description of covariate datasets with links to data files. Data is owned by respective publishing organisations. NHS App metadata accessible through application to NHS England via any NHS organisation.

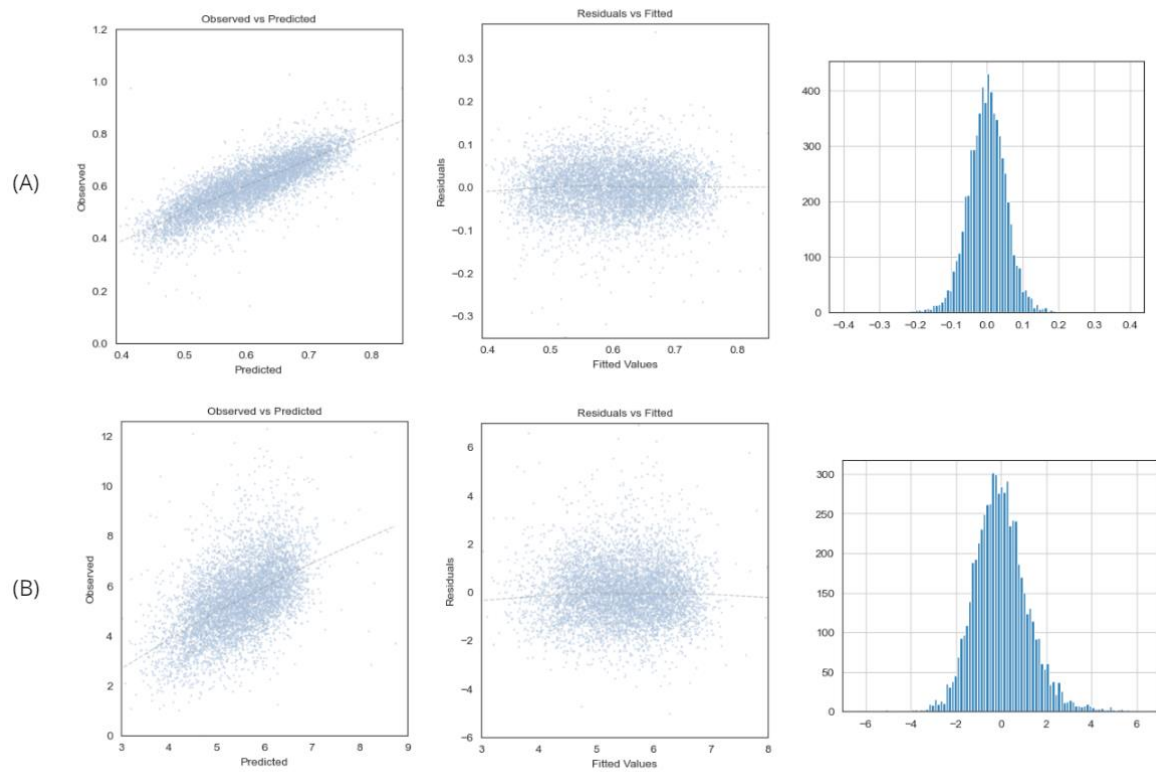


Figure B – Assessment of linearity, homoscedasticity, and normality for multivariable models: NHS App activation (A), and primary care online services (B)